International Studies in Entrepreneurship

Per Davidsson

Researching Entrepreneurship

Conceptualization and Design

Second Edition

Springer

International Studies in Entrepreneurship

Volume 33

Series editors

Zoltan J. Acs, George Mason University, Fairfax, VA, USA
David B. Audretsch, Indiana University, Bloomington, IN, USA

More information about this series at http://www.springer.com/series/6149

Per Davidsson

Researching Entrepreneurship

Conceptualization and Design

Second Edition

Springer

Per Davidsson
QUT Business School
Queensland University of Technology
Brisbane, Queensland, Australia

Jönköping International Business School
Jönköping, Sweden

ISSN 1572-1922 ISSN 2197-5884 (electronic)
International Studies in Entrepreneurship
ISBN 978-3-319-26691-6 ISBN 978-3-319-26692-3 (eBook)
DOI 10.1007/978-3-319-26692-3

Library of Congress Control Number: 2015958887

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Before We Begin…

Thank you so much for showing an interest in this new edition of *Researching Entrepreneurship*! It's going on 12 years since the first edition was published and 3 years—more when you read this—since I started working on this new edition. During that time, entrepreneurship research has undergone tremendous development. The qualitative and quantitative growth that the field has undergone is just amazing. Understandably, this plus the growing ambition that always tends to set in during a project like this mean that writing this new edition has been a much greater challenge than I first thought. But as I state repeatedly in the book: challenges are fun! If being an academic were easy, it would be boring. As a result, this book is just as much an entirely new book as it is a new edition of an existing one. As a case in point, more than half the figures and tables are new.

The Umpire Strikes Back! was my spontaneous first suggestion for the subtitle when the publisher suggested I come up with one for this new edition. I thought it would be kind of a fun and fitting pun for a sequel where someone takes on the outrageously self-aggrandizing role of telling others how to do their research, especially as entrepreneurship research spans an impossible range of topics, theoretical angles, types of data, and analysis techniques. In the end, I settled for *conceptualization and design*. I think it is still true that "This is a methods book. Of sorts" as I put it in the preface to the first edition. This said, I think the chosen subtitle adequately captures the drift in emphasis. This new edition has less emphasis on data and technique and more on fundamental thinking about what we are really trying to do, when we are doing "research." I hope the subtitle I finally chose is at least as fair a description of the contents as was my first idea.

So what have I done in more detail? Chapters 1 and 2 present essentially the same argument as before about what entrepreneurship and entrepreneurship research are, but have been thoroughly updated to reflect recent developments, as have all chapters. Chapter 3 has undergone greater changes, not least because the field of entrepreneurship research has become so much more theory-driven since 2004. Chapter 4, on general design issues, has tripled in length from 10 to over 30 pages and now provides a much better introduction to design and methods issues, if I may say so myself (and I may, because this is my book!). Although I sprinkle it with entrepreneurship-specific comments, it is actually an "introduction, but with a twist" that can be applied in other fields of research as well and to other phenomena.

Chapter 5 (on sampling and case selection) has been given a general overhaul and updating while most of its message stays intact. Chapter 6 keeps its title ("Operationalization Issues") but not much more. So much has happened in this area over the past decade that I decided to essentially start from scratch. Well, maybe 25 % stays essentially the same. The old Chaps. 7 and 8 have been scrapped, not because those topics have become obsolete but because my expertise on them has not been much updated since I wrote the first edition. Instead, I have added chapters on topics where I have done some hard and fun intellectual work in recent years, namely, the dependent variables in entrepreneurship research (Chap. 7) and "entrepreneurial opportunities" and their role in the "entrepreneurship nexus" (Chap. 8). Chapter 9 has doubled in length, and its front part is now the crescendo of a theme that runs through the book: the insufficiency of statistical significance testing and how we need to start to embrace replication and reproducibility if we really want to take seriously our role as developers of solid knowledge. I have also expanded the second half of the chapter with a couple of new replication examples. Chapter 10 on analysis approaches is short as before, but has been duly updated.

So what have I retained from the first edition? In its preface, I wrote "while hopefully retaining enough seriousness and credibility, I will try to refrain from dull academic jargon and unnecessarily heavy style." In the preface to the paperback edition, I similarly signaled that I tried to avoid "unnecessarily dry style and impenetrable academic jargon." A formal reviewer of that text noted that some might like this choice of style, others not. I have bet my money on the former and kept the informal and sometimes even joking style, while remaining dead serious about the message. The book will essentially continue in the chatty style of this preface, so if that makes you want to puke, don't say I didn't warn you! By the way, when I come to think of it, I realize it is not completely inconceivable that the stretch of long, core Chaps. 4–6 will not exactly come across as an example of lightweight, bedtime reading.

One reviewer of the first edition appreciated as a great strength that the book "provokes reflection and debate, rather than setting out rules to follow and comply with." I think I kept that as well. I want to encourage you to think—from a somewhat more enlightened position—rather than providing you with firm, authoritative rules of admonition. If you like the latter better, I'm sorry, but (research) life just isn't that easy/boring. As a consequence, the alert reader will spot some ambivalence as to where I stand on some topics, such as the merits of exploratory research. If you now think you're up for an orgy in indecision and cowardice, I should rush to ensure you that you will also find a greater number of frank statements about some of our research practices than you are likely to find in print elsewhere.

Another observation made by a formal reviewer of the first edition was that the book had "an element of autobiography about it." This is still true; in terms of the relative space allotted, the book is biased in favor of the types of topics and research approaches with which I am most familiar, and many examples are drawn from my own journey as a researcher. This is simply because it is with such a focus that there is any hope that I can provide insights you could not just as easily get elsewhere.

However, the enormous growth of entrepreneurship research has forced me to build much more on the work of others this time around, and as a consequence you can "look forward to" a rather long list of references. But sure enough, you will get a dose or two of my pet peeves!

Oh, another thing: even though the entrepreneurship research community has grown a lot, it is still the case that "the world market for a book by the title *Researching Entrepreneurship* is so limited that one can guarantee that it won't make its author rich. The upside of that is that you can trust it is an honest book. I write it because I want to share my experiences and not with the intent to maximize profits; hence I do not have to compromise with my convictions in order to reach my goals." One wealthy colleague used the expression "glutton for punishment" to describe my work on this book, while another likewise wealthy one used the term "labour of love" (sorry, he's Australian, so he did not say "labor"). I accept both characterizations.

For whom is this book intended? I like to think of it as a dialog with both emerging and established peers. Research students, doctoral programs, and courses focusing on entrepreneurship and entrepreneurship research are obvious primary targets. Those oriented toward related topics like small business, innovation, regional development, or organization/management/strategy more broadly may also find many aspects of the book useful. I think established colleagues might enjoy parts of the book, too. For example, the first two chapters if they are not already familiar with the argument, as well as the new, more conceptually oriented Chaps. 7 and 8. I would very much want every social science colleague there is to take Chap. 9 to heart. Other parts would offer fewer new content ideas, but may have value as refresher as well as inspiration for how to approach various issues in doctoral courses and supervision. This said, there are passages where I blush at the thought of a colleague thinking that *I* am thinking that *they* need *me* to tell *them* this or that basic point about research. No, that was not my intention.

Research students and fellow academics aren't the only target groups for this book, though. Analysts in market research firms and among opinion pollsters, in policy-making and policy-preparing offices, consultancy firms, statistical agencies, and business associations can also find large parts of the contents to be valuable, even if not every page and every sentence are a perfect fit for their needs. This is particularly the case if they have an interest in enhancing their professional competence at telling good evidence from bad and if their topical interest is directed toward entrepreneurship, small business, innovation, or economic and regional development issues.

Returning to pet peeves, when you write up a work like this over an extended period of time, you tend to somehow drift into discussing the same issues regardless of the main theme of the current chapter. I have weeded out some repetition of that nature, but deliberately let some stay. This is for the following reasons. First, some messages deserve and need some repetition—back in the days I taught marketing, they used to say you need to be hit by a message at least three times for it to have any effect at all. Second, although I have maintained the ambition (fantasy?) that it should be "a bearable experience to read it from cover to cover," the fact remains that books like this one "tend to be used rather than read." The repeated points belong in various contexts, and many users will only use a chapter or two at a time (or at all). Hence, please try to patiently endure the instances of repetition. I hope you will appreciate that at least I vary the phrasing; you won't find any sloppy cut-and-paste jobs.

Something else you will not find much in this book is philosophy of science arguments or references. This said, I have sneaked in the odd reference to ontology and epistemology in this new edition (and even discuss ancient Greek philosophy in one of the many footnotes—don't miss them; they're where half of the gems are!). Philosophy of science has its place and its points, but it rarely gives you much to really hold on to when conducting empirical research (because that is not what philosophers of science do and therefore they simply don't have that experience and expertise). I stay far short of being an expert on philosophies of science, but I'm not completely ignorant, and I agree that reflection on the foundations of knowledge production is both important and admittedly lacking in a lot of mainstream research. Although I don't find it totally convincing in terms of logical coherence, *scientific realism* is the school of thought that probably comes closest to the practically workable middle ground that I find most useful for guiding empirical research. This said, I do not believe in having a *faith* when it comes to philosophy of science (see, non-believing *is* obviously my faith). I can think of no more narrow-minded and unacademic attitude than thinking that "all the good guys think like us." So I tend to be an eclectic and pragmatic skeptic, accepting and refuting arguments from several camps.

Many people and organizations have contributed to this book. Far too many, in fact, for it to be possible to mention them all individually. To those mentioned in the preface to the first edition, I need to add at least the Australian Research Council and the Talbot Family Foundation for financial support. Many wonderful research students, postdocs, and other colleagues at QUT/ACE, JIBS, University of Louisville, the AoM Entrepreneurship Division, and beyond have inspired, critiqued, and in other ways contributed to this work. Among the most important recent research collaborators not mentioned in the first edition we find Scott R. Gordon, Lucia Naldi, and Paul Steffens. Although we do not publish much together, I also need to specifically mention two of the giants in the field, who have had great influence on my scholarship as well as on our entire field, in different but equally important ways: Paul Reynolds and Dean Shepherd. Throughout the book itself and its reference list, I show my appreciation to many others not specifically mentioned here or in the original preface. I would also like to acknowledge Robyn Denton's help with a couple of figures.

I will this time change my habit of not dedicating my books to someone near and dear, which I have followed because of the awkwardness of "giving" people something that does not interest them. My change of mind is due to the fact that the delightful human being who is now my wonderful wife, Thu Nguyen, is herself such an oddball that she voluntarily read and commented not only on this edition but also on drafts of the 2005 paperback edition back in the days when our relationship was only professional. Honey, this one's for you! It took a couple of periods of separation to get the damn thing finished, but it'll make the G & T taste even better, and I'm sure we'll continue to have good conversations and laughs about this book, just like we have about everything else. Thanks for making it sheer bliss to wake up every morning. You know you'll always have my unconditional love.

Brisbane, Australia Per Davidsson

Contents

What Is Entrepreneurship?

1

Abstract

What is entrepreneurship? To do research on entrepreneurship, we first need to decide what we mean by that term. A challenge here is that entrepreneurship has many definitions and connotations. As a societal phenomenon, this chapter proposes that entrepreneurship be defined as *the competitive behaviors that drive the market process*, alternatively phrased as *the introduction of new economic activity that leads to change in the marketplace*. The chapter elaborates on the advantages and implications of this choice of perspective.

1.1 On the Variety of Definitions and Views of Entrepreneurship

Researching entrepreneurship is fun, fascinating, frustrating—and important, if you ask me. One of the fascinations is the richness of the phenomenon, which leads to one of the greatest frustrations, namely, the lack of a common understanding of what precisely entrepreneurship *is*. Let me put it this way: there is no shortage of suggestions as to what the phenomenon "entrepreneurship" really consists of. Here are a few examples:

– New entry (Lumpkin & Dess, 1996)
– The creation of new enterprise (Low & MacMillan, 1988)
– The creation of new organizations (Gartner, 1988)
– A purposeful activity to initiate, maintain, and aggrandize a profit-oriented business (Cole, 1949)
– The process by which individuals—either on their own or inside organizations—pursue opportunities without regard to the resources they currently control (Stevenson & Jarillo, 1990)
– The process of creating something different with value by devoting the necessary time and effort; assuming the accompanying financial, psychic, and social risks;

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_1

and receiving the resulting rewards of monetary and personal satisfaction and independence (Hisrisch, Peters, & Shepherd, 2008)
- The occupational choice to work for one's own account and risk (Stephan & Uhlaner, 2010)
- The junction where venturesome individuals and valuable business opportunities meet (Parker & van Praag, 2012)
- A specific effort by an existing firm or new entrant to introduce a new combination of resources (Lee, Peng, & Song, 2013)
- The act by which new firms come into existence (Bird & Wennberg, 2014)

Kirzner (1983) offered the following compilation of roles assigned to the entrepreneur by various economic theorists:

- A specific kind of labor service
- Assuming the risk
- Innovator
- Arbitrageur
- Coordinator, organizer, or gap-filler
- Providing leadership
- Exercising genuine will
- Acting as a pure speculator
- Acting as employer
- Acting as superintendent or manager
- Acting as a source of information
- Being alert to opportunities as yet overlooked in the market

Using an empirical approach to the question of what entrepreneurship is, Gartner (1990) found the following eight themes to emerge when professional users (academic and others) of the entrepreneurship concept were asked about its inherent meaning:

- The entrepreneur
- Innovation
- Organization creation
- Creating value
- Profit or nonprofit
- Growth
- Uniqueness
- The owner-manager

Similarly, a content analysis of journal articles and books performed by Morris, Lewis, and Sexton (1994) yielded the following most common definitional keywords:

- Starting, founding, creating
- New business/new venture
- Innovation, new product, new market

- Pursuit of opportunity
- Risk taking, risk management, uncertainty
- Profit seeking, personal benefit
- New combinations of resources, means of production
- Management
- Marshaling resources
- Value creation

Tired yet? Feeling most of the references are oldish? Good! We may be getting somewhere in this field! At this point it should be superfluous to point out that no one can claim to have the one, true answer to the question of what the phenomenon "entrepreneurship" truly is. So far, in the social construction game of filling the entrepreneurship concept with meaning, none of the existing and partially overlapping constructions seems to have achieved dominance over the others. Some of the issues on which the views on entrepreneurship differ are the following:

- Is entrepreneurship something that is restricted to the *commercial sector*, and is it an *economic* phenomenon or something that can present itself within any area of human endeavor?
- Is entrepreneurship restricted to *small* or *new* or *owner-managed firms*, or can it be executed by or within organizations of any age, size, and governance structure?
- Is entrepreneurship an *innate characteristic* (disposition) or a type of *behavior*, or does it involve a special type of *outcome* (e.g., is success required)?
- Does something have to be *purposeful* in order to amount to entrepreneurship, or can processes involving luck and serendipity qualify?
- Is *innovation* required, or can imitative initiatives exemplify entrepreneurship?
- Is *risk taking* a necessary requirement?
- Does entrepreneurship involve the *discovery* (or creation) of ideas for new ventures, the *exploitation* of such ideas, or both?
- Is it solely a *microlevel phenomenon*, or is entrepreneurship a meaningful concept on more aggregate levels as well?

The language games we play regarding the meaning of entrepreneurship are of the funny type of games where—unlike sports—it is totally conceivable that two opposing players both determine that they (according to their own rules) won the game, whereas the spectators, that is, the fellow researchers who read the arguments, find that both sides scored points, but since they did not play the same game on the same field, it wouldn't be very meaningful to appoint a winner. Like sports, however, those language games are something some people think are extremely interesting and important, whereas others couldn't care less.

I tend to be somewhat ambivalent on the importance of precise, inherently consistent, and agreed-upon definitions. I am pretty sure, however, that underneath the various constructions of entrepreneurship we shall find interesting and important social *phenomena*. Part of me says, "Forget definitions—let's just go and learn and tell about those important phenomena!" Another part of me, however, strongly feels that in order to do just that, a researcher must have a very clear idea of what that phenomenon is and be able to communicate that idea, be it shared or not by most of the readers. I will let that other part of me speak for a while now. Besides, as you can tell from Chap. 8, I have become more convinced over time that clarity about what we mean by our theoretical concepts is essential.

Some of the variations in entrepreneurship definitions, I believe, are relatively minor and of little significance. They largely reflect the same underlying social phenomenon, and therefore the differences in the finer nuances do not confuse communication or hinder knowledge accumulation. Other differences, however, may have such effects and therefore cannot be disregarded as easily. Over the years, I have come to the conclusion that the different entrepreneurship definitions actually address *two* relatively distinct phenomena (Davidsson, 2006).

The first equates the term with *independent business*: entrepreneurship is starting and running one's own firm (recently increasingly including social enterprises in the not-for-profit sector as well, see Austin, Stevenson, & Wei-Skillern, 2006; Mair & Marti, 2006). According to this view, entrepreneurship research studies *entrepreneurs* understood as flesh-and-blood business owner-managers. Such people remain entrepreneurs for life or at least as long as they are running their own business. Consequently, any trait, emotion, cognition, behavior, or achievement of such individuals is an issue for entrepreneurship research.

The second view regards entrepreneurship as the *creation of new economic activity* (or, in the most allowing cases, any new activity). The major underlying theme here is that the development and renewal of any society, economy, or organization requires microlevel actors who show the initiative and persistence to make change happen. According to this view, *entrepreneur* is a theoretical abstraction that refers to one or more individuals who in a particular case bring about this change as an individual feat *or* as a team/organizational effort *or* in sequence, i.e., different actors may fulfill different roles as an entrepreneurial process unfolds over time (Dimov, 2007). The focus is on the activity, on entrepreneur*ship*. While this requires individual initiative, it is not necessary to label one individual as "the entrepreneur" in an entrepreneurial process. Neither is there a particular class of people who are constantly "entrepreneurs" while others are "non-entrepreneurs." Rather, entrepreneur is a *role*, which individuals exercise on a temporary basis (Schumpeter, 1934).

To further illustrate the difference as well as the overlap, we may note that the creation or emergence of new, independent business is of central interest from both perspectives. A topic like family business succession problems falls naturally within the domain when entrepreneurship is understood as the founding *and running* of independent businesses, whereas this topic has nothing to do with entrepreneurship from the *creation of new economic activity* view—unless the research focuses specifically on, e.g., the effect of succession on the firm's ability to innovate. Corporate entrepreneurship—i.e., creation of new economic activities by large, established firms (with dispersed ownership)—is part of the domain from the latter perspective but an oxymoron when entrepreneurship is understood as starting and running an independent business.

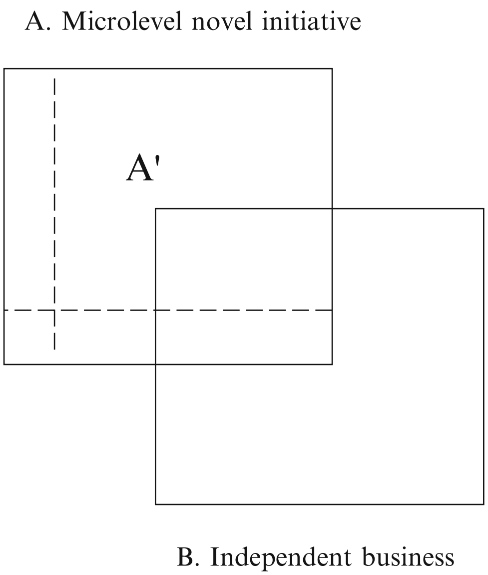
A problem with many definitions of entrepreneurship, as well as many implicit views on this phenomenon, is that they cover in fact an amalgam of the two social realities described above. This was certainly the case with the old Domain Statement of the Entrepreneurship Division of the Academy of Management. The new domain statement acknowledges both perspectives—and that they are separate:

> Specific domain: (a) the actors, actions, resources, environmental influences and outcomes associated with the emergence of entrepreneurial opportunities and/or new economic activities in multiple organizational contexts, and (b) the characteristics, actions, and challenges of owner-managers and their businesses (Entrepreneurship Division, 2011; Mitchell, 2011).

My personal development over time has certainly been a drift—like the overall tendency in the international research community—from embracing an entrepreneurship/small business view toward being more inclined to include the creation of new business ventures within any organizational context (at least conceptually) and at the same time becoming more reluctant to include just any aspects of ownership-management (Shane, 2012). I think that in order to make useful contributions to entrepreneurship research, the researcher needs to take sides here. I'd be very pleased if I could convince my readers of that point, whether or not they decide to follow the specific direction I will outline below. "Taking sides" refers to how we use the E-words, not what research interests we pursue. For example, I have undertaken quite a bit of research on small firm growth and other aspects of small- and medium-sized enterprise (SME) management that I see nothing wrong in—but today I would not necessarily apply an E-label to all of it.

The choice is actually not only between the two alternatives outlined above. There are also more restricted or refined alternatives. In order to discuss these, we take the help of Fig. 1.1.

Fig. 1.1 Possible delineations of the entrepreneurship phenomenon



One obvious alternative is to choose the entire square *B* as one's view of entrepreneurship. I personally do not see the logical or linguistic reasons for doing so. We have seen above that entrepreneurship is widely connoted with quite an array of things that are definitely not necessary characteristics of independent businesses. If one wants to reserve the concept for independent firms, the intersection *A* and *B*—entrepreneurial small business, if you like—would seem a more attractive alternative. This would include, for example, new firm formation; small firm innovation; internationalization and certain other aspects of growth of small, independent firms; and possibly the rejuvenation of family businesses as a result of ownership and management succession.

This is not my own preferred choice. Neither will I argue for the inclusion of the entire square *A*. The view of the entrepreneurship phenomenon that I am going to elaborate on below—and which was first developed and presented in Davidsson (2003)—is instead a more restricted alternative illustrated by the square A', demarcated by dashed lines at the left and bottom. That is, I propose a "microlevel novel initiative" view, but for reasons detailed below, I restrict it to *economic* endeavors—those dealing with resource utilization—in a *market or market-like context*. I thus exclude nonmarket activities such as internal, organizational change per se. Activities undertaken by existing or emerging independent businesses are certainly included in this view, as is social entrepreneurship, but only as long as they entail the introduction of new goods or services or at least new competition in the market.

1.2 My Proposed View of the Entrepreneurship Phenomenon

Hoping that the reader remembers that I have already pointed out that no one can claim to have *the* right answer to the question of what entrepreneurship really is, here is what I propose: a fruitful way to define the societal/economic phenomenon "entrepreneurship" is the functional notion in Austrian economics that entrepreneurship consists of *the competitive behaviors that drive the market process* (Kirzner, 1973, pp. 19–20). This does not imply a general admiration or preference for Kirzner's theorizing over, for example, Schumpeter's (1934) or Baumol's (1993). I favor this definition because it is succinct and gives a satisfactorily clear delineation of the role of entrepreneurship in society.

Firstly, it is based jointly on behavior and outcomes. The behavior part is necessary in order not to lose track of the fact that microlevel decisions and actions are needed for any change to be introduced. As regards the outcome part, I argue that when we think of entrepreneurship as a societal phenomenon, it is a distinctive advantage to include an outcome criterion and make clear, for example, that mere contemplation over radically new ideas or vain introduction of fatally flawed ones does not amount to "entrepreneurship." Entrepreneurship makes a difference, or else it isn't entrepreneurship. In order to "drive the market process," the activity has to have some direct or indirect success. To those readers who get itches at this stage, I can only say I hope that the next chapter will solve the problem. So please stay tuned, because I will relax outcome as a necessary criterion when discussing entrepreneurship as a scholarly domain. When defining entrepreneurship as a societal *phenomenon*, I believe it useful to portray it as microlevel behavior that has macrolevel implications.

Secondly, this Austrian notion puts entrepreneurship squarely in a market context and makes clear that it is the suppliers who exercise entrepreneurship—not customers, legislators, or natural forces that also affect outcomes in the market. When suppliers engage in entrepreneurship, they introduce new, improved, or competing offerings in an emerging or already existing market. They thereby drive the market process in one or more of the following ways:

1. They provide customers with new choice alternatives, potentially giving some of those customers more value for their money.
2. They stimulate incumbent actors to improve their market offerings in their turn, which increases efficiency and/or effectiveness of those actors.
3. If successful, they attract other new entrants to the market, thus further increasing competitive pressures toward improved efficiency and effectiveness (Holcombe, 2003; Plummer, Haynie, & Godesiabois, 2007).

Importantly, driving the market process does not require that the first mover makes a profit but refers to the suppliers as a collective. Even if it eventually loses out, the first mover contributes to driving the market process if subsequently someone else gets it right, which leads to a lasting change in the market (see Fig. 1.3).

Admitting that change-inducing microlevel initiatives are undertaken in nonmarket contexts, I believe it an advantage to restrict the use of the specific term "entrepreneurship" for the market or market-like contexts, that is, when the setting involves customers, suppliers, and (potential) competitors or very close equivalents to those. The main reason for this restriction is that I think it is valuable for the progress of entrepreneurship research to make the concept as distinct and well defined as possible. Where does this put social entrepreneurship? Well, does the social venture directly or indirectly, intentionally or not, have the effect that resources be put to better use? OK, then it is economic. Does it provide new choices—including a choice of "something" where previously there was "nothing"—for some type of client, customer, or the like? If that's another "tick," we may not even need the third, but if the activity also has the capacity to change the behavior of others—incumbents and followers providing similar services or addressing the same social problem—we should have no hesitation that it is entrepreneurship we are seeing.

Now, broad agreement in the research community is probably not to be hoped for. Some would like to define entrepreneurship more narrowly than this, while others would argue for an even more inclusive perspective. This said (and accepted), I think it important that individual researchers carrying out specific research projects at least base their use of the entrepreneurship concept on a notion as clear as the one suggested here. Moreover, those who want to include novelty through "new combinations" (Schumpeter, 1934) in *any* domain of human behavior in the concept of "entrepreneurship" may have reason to contemplate the full implications of this

		(To) market New	(To) market Old
(To) firm	New	**I** New offer: - Product/service - Bundle - Price/value relationship New competitor	**II** Organizational change: - Acquisitions - Spin-outs/Buy-outs - Internal re-organization - Management succession
	Old	**IV** Geographical market expansion (incl. internationalization)	**III** Business as usual Non-entrepreneurial growth

Fig. 1.2 Firm and market newness of economic activities

choice. This would not only allow, for example, novelty in the arts and in the organization of humanitarian aid activities into the picture, but also novelty in crime and warfare. And it would certainly make the events of September 11, 2001, an entrepreneurship masterpiece. To conceive of a fully fueled passenger jet as a missile and to combine the idea of hijacking with that of kamikaze attacks was certainly innovative, and in terms of impact—economic and otherwise—it has few parallels. However, regarding these attacks as driving *market processes* is far-fetched. This author would therefore suggest they not be regarded an instance of entrepreneurship. Put in slightly different words, entrepreneurship according to the suggested perspective consists of the *introduction of new economic activity* that leads to change in the marketplace (cf. Herbert Simon in Sarasvathy, 2000, pp. 2, 11). This is illustrated in Fig. 1.2.

Note that "new" along the market axis means either that an entirely new market emerges or that an activity is new to an existing market (Dahlqvist & Wiklund, 2012). In the latter case, "new" could mean the launch of an innovation, but merely entering as a new competitor could also qualify. Likewise, along the firm axis, "new" means that either the new activity is an independent start-up, implying that a new firm emerges as a result, or it is an internal new venture, which means that the firm has previously not been making this particular market offering.

Under the suggested definition, the left-hand side of the figure—quadrants I and IV—exemplifies entrepreneurship, whereas quadrants II and III do not. This concurs also with the argument developed at some length by Baumol (1993) in that imitative entry and internationalization are included in the concept, whereas acquisitions, for example, are excluded.

1.2.1 New Offer as Entrepreneurship

The first entry in quadrant I reads "New offer." This is when something so new is introduced that a new market is created (Bhave, 1994; Navis & Glynn, 2010; Santos & Eisenhardt, 2009; Sarasvathy, 2008) or at least no supplier has previously made the same offer in the same market. There is hardly any disagreement among scholars that this should be included in the concept of entrepreneurship, although some might want to restrict the inclusion to situations where a new and/or independent firm is behind the new offer.

The first category, *new product or service*, corresponds to Schumpeter's (1934) "new product" and Bhave's (1994) notion of "product novelty," respectively, and requires no further explanation. The second category, *new bundle*, refers to any combination of product and service components that—as a package deal—is unique relative to what has previously been offered on the market, although no individual component may be strictly new. This is what Bhave (1994) calls "new business concept" and what Amit and Zott (2001) have in mind when they talk about "new business model"—as long as the concept or model includes newness as perceived by buyers and competitors. In some cases it amounts to Schumpeter's (1934) category "reorganization of an entire industry." Illustrative cases include the market entry by IKEA and Dell. The newness they brought to the market was not so much the product in use, but in the division of labor among different actors—including the consumer—in the production and distribution of the end product. More recent successes like Uber and Airbnb can also be put in this category.

IKEA would also qualify under the third category included in "New offer" *new price/value relationship*. This does not create a new market but drives the market process because it changes consumer choices and gives other competitors reason to change their offerings. Consequently, Kirzner (1973, pp. 23-24) explicitly discusses offering the same product at a lower price as one form of entrepreneurship. A process innovation or organizational change (quadrant II) may often be the underlying cause of a new price/value relationship, but this is not *necessarily* the case. It may also represent a strategic change that relies on expected economies of scale or experience or a switch from low-volume/high-margin to high-volume/low-margin strategy.

1.2.2 New Competitor as Entrepreneurship

The second main entry in quadrant I is "New competitor." That is, I suggest that not only innovative but also imitative entry be included in the entrepreneurship concept (cf. Aldrich & Ruef, 2006). This is when a new, start-up firm enters the market or an existing firm launches a new product line in a situation where other firms already supply the market with essentially the same product. Now, the reader may wonder whether the author is incapable of seeing the difference between the entry of yet another hairdresser or mom-and-pop store on the one hand and the venture capital-backed launch of a new, high-potential biotech firm on the other. Well, let's look back and see how we have defined entrepreneurship. Does the new hairdresser provide customers with a new, potentially better alternative? If nothing else is special about the new actor, it will at least have a different location, which may be more convenient for some customers. And the closest competitors may well find reason to reduce their price or ramp up their service level in order to limit the damage caused by the new competitor. Hence, the reason for imitative entry to be included in the entrepreneurship concept is that such entry drives the market process in the sense that consumers get additional choices and incumbent firms get reason to change their behavior to meet this new competition.

Another vantage point for this argument is that studies have found that entry with complete lack of novelty tends not to appear empirically (Bhave, 1994, p. 230; Davidsson, 1986). No entrant is a perfect clone of an existing actor. Therefore, trying to include an innovativeness criterion in the definition of entrepreneurship would create problems. Rather than drawing the line at zero innovation (which would exclude no cases), one would be forced to define an arbitrary minimum limit of innovativeness across different industries and types of novelty. All in all, there are several good reasons to include imitative market entry in the entrepreneurship phenomenon. However, if the new entrant is inferior along all dimensions, it will neither succeed nor influence other actors' behavior, and then it does not constitute an instance of entrepreneurship. And yes, yes, and yes—there is a difference between a new hairdresser and a new Google or Apple. We will get to degrees of entrepreneurship shortly.

1.2.3 Geographical Market Expansion as Entrepreneurship

Defining entrepreneurship the way we have done makes it logical to also include quadrant IV—geographical market expansion—in the concept of entrepreneurship. Some readers may find this to be overextending the entrepreneurship concept. What makes the "simple" repetition of old success recipes in new contexts entrepreneurial? The answer lies in the fact that we have defined entrepreneurship from a market perspective. Although the activities may (largely) no longer be new from the firm's perspective, their introduction in new markets—if not totally unsuccessful—drives the market process in these new places. When business model innovators like McDonalds, IKEA, Dell, eBay, or the free newspaper Metro entered their *n*th country market, it may well have been as revolutionary for the consumers and competitors in that market as it was for consumers and competitors in the markets where these businesses originated. If the entry is successful, it reflects Schumpeter's (1934) "new market" category of economic development. The alternative to require newness to the firm as a criterion would lead to less desirable consequences. For example, had Southwest Airlines successfully copied their own concept in the European market, it would not constitute entrepreneurship. If instead a new actor (Ryanair) copied the concept and took it to the European market, it would count as entrepreneurship. This is less than satisfactory from any perspective, and from a market perspective it makes no sense at all.

1.2.4 Organizational and Ownership Changes Are Not Entrepreneurship

By contrast, according to our conceptualization, the organizational and ownership changes listed in quadrant II do *not* by themselves constitute entrepreneurship. At this point, after the generosity awarded to imitative start-ups and geographical expansions, some readers may be outright annoyed to find internal reorganizing, no matter how dramatic and creative, to be excluded from the entrepreneurship concept. But please bear with me a few more lines or perhaps a few more pages. Perhaps you can appreciate the internal logic of the argument, regardless of whether you are inclined or not to fully accept the definition of entrepreneurship that I develop here.

I freely admit that it is conceivable (and likely) that reorganization facilitates the creation of new economic activity by the organization. However, it is not *necessarily* the case that organizational and ownership changes lead to such effects. Actually, there are at least four cases: (a) an organizational or ownership change is intended to lead to more new market offerings by the firm and does so, (b) the same as (a) but the intended increase in new market offerings does not happen, (c) the change is undertaken for other reasons and has no effect on the firm's market offerings, and (d) the change is undertaken for other reasons but has the *unintended* effect of also making the firm more entrepreneurial in terms of introducing novelty in the marketplace. I think it is valuable not to lump together all those cases and include them in the notion of entrepreneurship. Instead, I see it as valuable to conceptually separate the organizational or ownership change from its effects. With our market-based definition of entrepreneurship, it is the (successful or influential) launching of new business activities that might follow from it, and not the organizational change itself, that constitutes entrepreneurship. Whether increased launch of novelty to the market was an intentional outcome or not does not matter.

The argument is perhaps easier to accept if we move to the level of societal organization. Politicians can decide on changes in how society is organized and introduce deregulation or other institutional changes which create room for new economic activity in market *x* and therefore an increase in competitive behaviors that drive the market process in that market. In other words, the result is more entrepreneurship in market *x*. According to my argument, it is not *the politician* but the microlevel actors in that market who exercise entrepreneurship in market *x*. The political decision *facilitates* entrepreneurship. In the same way, a manager may facilitate entrepreneurship through organizational change, but it is the market-related activities that may result, and not the organizational change per se, that constitute entrepreneurship.

This conceptual distinction is also the reason why I refrain from including Schumpeter's (1934) "new production method" and "new source of supply" or Bhave's (1994) "novelty in production technology" in the definition of the entrepreneurship phenomenon (cf. Davidsson, 2003; Kirzner, 1983, p. 288). According to my argument, it is only when these events are translated into new offers or a new price/value relation in the market that we see entrepreneurship. As we shall see in the next chapter, the study of how organizational change relates to discovery and exploitation of new venture ideas remain an important question for entrepreneurship as a scholarly domain.

1.2.5 Business as Usual and Non-entrepreneurial Growth

Turning now to quadrant IV, "Business as usual" here is, at first glance, as easy to exclude from the notion of entrepreneurship, as "New offer" in quadrant I was easy to include. But full agreement does not seem to exist even here. First, we have von Mises' denial of the existence of such a thing as "business as usual" when saying that "In any real and living economy every actor is always an entrepreneur" (Mises, 1949, p. 253). One can argue that no market action is completely void of novelty. For example, when a daily newspaper carries out the totally expected and routine actions of producing a new issue and distributing it to its subscribers and usual sales outlets, it is a *new* issue, and not yesterday's paper, that is being distributed. Competitors will equally routinely read it, and it cannot be ruled out that some part of the contents may have a twist that inspires the competitor to do something in a future issue, which it otherwise would not have done. In other words, we find an element of "competitive behaviors that drive the market process" in these routine actions. Although this seems to lead to a delimitation problem similar to the arbitrary innovation criterion discussed above, my conclusion in this case goes in the other direction. That is, there is a lot of "known products for known buyers" activity going on that is so clearly *predominantly* of a "business as usual" character that it is not very difficult to classify it as such, both conceptually and empirically, and thus exclude it from our definition of entrepreneurship.

The issue of non-entrepreneurial growth is tricky for slightly different reasons (see Davidsson, Delmar & Wiklund, 2002, for an elaborate discussion). When an economic actor exploits a venture idea, there will be no well-defined moment at which "entry" ends and "continued, routine exploitation" begins. Schumpeter (1934) held that mere volume expansion was not entrepreneurial, whereas he included the opening of new markets. It is a similar distinction I have in mind here. By "non-entrepreneurial growth," I mean passively or reactively letting existing activities grow with the market. This would not provide much cause for alarm among competitors nor give customers new choices.

1.2.6 Entrepreneurship as Microlevel Behavior with Macrolevel Implications

I pointed out in the early parts of this chapter that one important feature of the entrepreneurship definition I have chosen is that it portrays entrepreneurship as a microlevel behavior with important implications for more aggregate levels of analysis. Simplistic conceptions of venture outcomes typically classify them as successes or failures. However, when you consider multiple levels of analysis, a more complex picture emerges. It is, for example, possible that a new venture that fails miserably has important positive effects on the economy at large, because both those involved and others learn for the future and can come up with smarter solutions that would not have been within reach without the initial "failure." This is what Fig. 1.3 is getting at. "Venture" could here mean the sole activity of a new firm

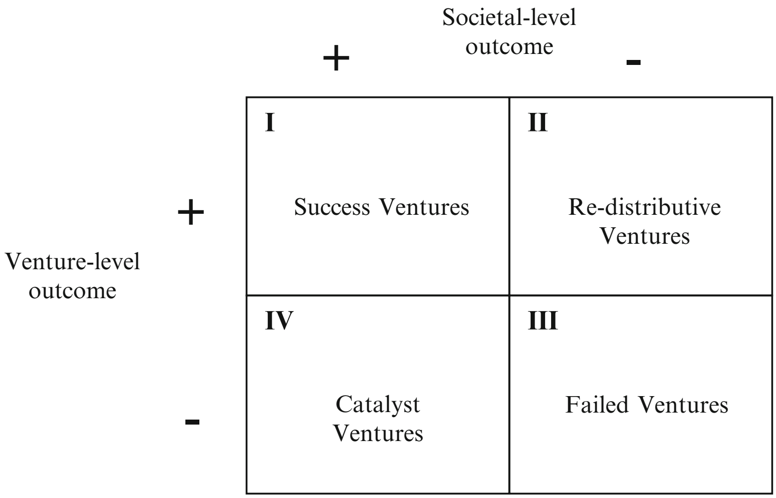


Fig. 1.3 Outcomes on different levels for new ventures (new economic activities)

or a new, additional activity by an established firm. Thus, "venture" should not be interpreted (necessarily) as new firm or company, but as a new-to-the-market activity as discussed above.

If we turn first to quadrant I, we find ventures that are successful in themselves and which produce net utility to society as well. These ventures are analytically unproblematic. Their successful entry into the market no doubt "drives the market process," and hence they exercise entrepreneurship under the definition I have suggested. Likewise, the failed ventures in quadrant III pose no trouble. They represent launching efforts that do not succeed in establishing themselves in the market, and neither do they inspire followers or incumbent firms so that the eventual net effect becomes positive on the societal level. They are, so to speak, completely vain efforts.

The catalyst ventures in quadrant IV are a more interesting category. Moreover, they are probably much more common than we might first think. Although not successful on the microlevel—perhaps because they are outsmarted by followers or retaliating incumbents—they do "drive the market process" precisely because they bring forth such behavior on the part of other actors. An unsuccessful venture that inspires more profitable successors does not *complete* the entrepreneurial process, but it no doubt contributes to the entrepreneurship phenomenon. The total effect on the economy is not necessarily smaller for catalysts than for "success ventures." Catalyst ventures may therefore be a very important category from a societal point of view, and I would be very pleased to see more research on this neglected topic. The importance of catalyst ventures should also serve as a warning against too simplistic a view on microlevel failure.

The ventures in quadrants I and IV, then, represent entrepreneurship whereas the failed ventures in quadrant III do not. What about the "redistributive" ventures in quadrant II? These are ventures that yield a surplus on the microlevel while at the same time the societal outcome is negative. Examples could be trafficking with heavy drugs or—as in an actual case in Sweden—a graffiti removal operation whose owners use nighttime and spray paint to generate demand for their business. Thus, in these cases, those involved in the venture enrich themselves at the expense of collective wealth. Does this represent entrepreneurship? I would argue that the theoretical status of "redistributive" ventures is determined by the answer to "toward what" entrepreneurship drives the market process. Schumpeter (1934) and Kirzner (1973, p. 73) give seemingly contradictory answers to that question. On closer look, however, the movement *from* Schumpeter's (local) equilibrium and the movement *toward* Kirzner's (global) equilibrium are in full agreement insofar as that entrepreneurship drives the market process toward *more effective and/or efficient use of resources*. Therefore—admitting a sense of comfort and relief—I hold that there is theoretical ground to suggest that "redistributive" ventures do *not* represent entrepreneurship[1]. Entrepreneurship leads to improved use of resources in the economic system as a whole, and the redistributive ventures in Fig. 1.3 do not fulfill that criterion. Pick your heroes carefully!

Portraying the possible outcomes the way I have done in Fig. 1.3 is, of course, still a radical simplification. Outcomes are described as dichotomous and no explicit time horizon was introduced. What is perceived as socially productive today may be seen as pure evil in the future. Further, outcomes on only two out of many possible levels (e.g., venture, firm, industry, region, nation, and world) were discussed. In practice, assessing exactly where individual ventures fit into this framework would in many cases be a daunting task, especially when it comes to aggregating utility across individuals and generations in order to determine what is and is not socially valuable. Further, if one enjoyed the luxury of a perfect and just institutional framework, it would be easy to argue that redistributive ventures equal illegal ventures. Regrettably, we will have to live with the fact that in real economies "legal yet redistributive" and "illegal yet socially beneficial" ventures are both possible. Despite these problems, I think it is useful to highlight the distinctions made here and to note that as theoretical categories not only "success ventures" but also "catalyst ventures" carry out the entrepreneurial function in the economy, whereas neither "failed ventures" nor "redistributive ventures" fulfill this role.

1.2.7 Degrees of Entrepreneurship?

I said earlier that the inclusion of imitative entry called for a discussion of "degrees" of entrepreneurship (cf. S. Carter, 2011; Davidsson & Gordon, 2012; Shane, 2009; Wong, Ho, & Autio, 2005). Realizing the variations in scope and impact of

[1] In one of the most important papers of all time in the entrepreneurship literature, Baumol (1990) has a slightly different take on this issue. He would accept redistributive ventures as instances of entrepreneurship, but emphasizes that entrepreneurship comes in productive, unproductive, and destructive varieties. The conclusion remains the same: societal institutions need to gear the energy of creative and profit-seeking people toward activities that benefit society as a whole as well.

"competitive behaviors that drive the market process," it seems natural to treat entrepreneurship not as a dichotomous variable, but to say that some ventures show more entrepreneurship than others. But what should be the criterion by which we judge the degree of entrepreneurship? There are at least three possibilities:

The degree of (direct and indirect) impact on the economic system. If we choose this criterion, we stay true to our definition of entrepreneurship as the competitive behaviors that drive the market process. In a theoretical discussion of entrepreneurship, then, this should be the preferred criterion, simply because it is the most correct one. For empirical research practice, the criterion has severe shortcomings because impact can only be assessed after the fact and not in real time and because even then it can be very difficult to obtain even roughly correct estimates of total direct and indirect effects on a complex economic system. These problems, however, should bother us in the next chapter rather than this one. A variation (or an indicator) of the "degree of impact" criterion is the criterion "amount of wealth created." Needless to say, this suffers from the same kind of assessment problems.

The degree of novelty to the market. This is intuitively appealing in the sense that what is more creative is seen as a higher degree of entrepreneurship. Although the problem of comparing very different kinds of novelty pertains to this criterion, it has the advantage that it is not totally impossible to assess in real time (see Dahlqvist & Wiklund, 2012, and Chap. 6). For the very reason of capturing more cases with a high degree of market novelty, we added a "high-potential" judgment sample to our current, large-scale study of ongoing start-up efforts, the Comprehensive Australian Study of Entrepreneurial Emergence (CAUSEE). More about that study later. The main downside with the market novelty criterion is the following: innovative new activities that are successful are likely to have greater market impact on average. However, it may actually be more difficult for an innovative venture to be successful at all (Semasinghe, 2011). There is no guarantee that a high degree of novelty ascertains market effect—history is full of weird inventions that nobody wanted! Some seemingly marginal innovations revolutionize markets and create great wealth whereas some radical innovations have marginal impact or fail altogether. Therefore, the degree of novelty to the market is at best a rough proxy for degree of entrepreneurship.

The degree of novelty to the actor. Sometimes, you hear expressions like "That was very entrepreneurial of you" or "That was a very entrepreneurial move for that firm." Presumably, this means that the action was radically different from what *that actor* has done before. The problem is that the same action was not necessarily very novel or valuable as the market sees it. Relating the degree of entrepreneurship to the history of the actor rather than to the market has highly undesirable consequences. For one thing, this type of criterion actually makes previous inactivity or conservatism increase an actor's potential for showing a high degree of entrepreneurship! Moreover, it is a criterion that regards it more entrepreneurial to do something totally unrelated to one's prior experience. Theories as well as empirical findings suggest this may not be a wise move (Barney, 1991; McMullen & Shepherd, 2006; Sarasvathy, 2001; Shane, 2000). I would therefore discourage the use of this kind of criterion for “degree of entrepreneurship.”

In all, there is a conceptual need for discussing “degrees of entrepreneurship.” Importantly, in my view this variation belongs primarily in the dependent variable and not in the definition of what is entrepreneurial. Admittedly, there is no easy or straightforward way to actually assess such variation—especially not prospectively. But if theorizing and researching were easy tasks, they wouldn’t be much fun! Of the less-than-perfect but available alternatives, the degree of impact on the economic system is the criterion that best matches the definition of entrepreneurship that I have proposed. Degree of novelty either to the market or to the actor is better regarded as a possible cause of variations in the degree of entrepreneurship (or impact of entrepreneurship) than being a direct measure of such variation. In practice, when conditions do not allow careful, retrospective assessment of economic impact, researchers will have to accept proxies that are thought to reflect potential for higher impact. One example is the development over time in the Global Entrepreneurship Monitor (GEM) project of indicators of the quality of each start-up attempt, such as whether it is necessity based or driven by perception of opportunity, employs new technology, and aims for growth, innovation, or internationalization (Kelley, Bosma, & Amoros, 2011).

1.3 Summary and Conclusion

There are almost innumerable suggestions in the literature concerning what entrepreneurship really is. Noting that no one can claim to have the correct answer, I have proposed that defining entrepreneurship as the *competitive behaviors that drive the market process* has much to commend it. I think so for the following reasons. This definition emphasizes *behavior* rather than assuming a dispositional stance that has proven largely unfruitful (Gartner, 1988; Foss & Klein, 2012). It also includes an *outcome* that is successful or influential. Jointly, this implies that the processes of *discovery* and *exploitation* are included and that mere contemplation over radical ideas is not an example of entrepreneurship and neither is or the introduction of completely vain innovations. Further, the definition restricts entrepreneurship to a *market context*, which gives a more precise and coherent characterization of this phenomenon. At the same time, the definition is permissive in that it does *not* take a restrictive stand on purposefulness, innovation, organizational context, or ownership and personal risk taking. Importantly, it *links micro to macro* by portraying entrepreneurship as a microlevel phenomenon with important effects on more aggregate levels. Finally, relative to many other alternatives, I would argue that the suggested definition has advantages in terms of being *clearly delimited*, *logically coherent*, and *easy to communicate* (Suddaby, 2010).

Of course, my arguments will not convince everybody. To those who want entrepreneurship to mean “anything that concerns independent businesses,” I can only say “I’m sorry! Our interests have a certain degree of overlap, but our views on the entrepreneurship phenomenon are fundamentally different.” Therefore, much of the remainder of this book may be of less value for such readers. Those with particular interest in social-, institutional-, sustainable/eco-/green-, or sports-related varieties of entrepreneurship—to mention a few, significant developments of recent times (see, e.g., Greenwood & Suddaby, 2006; Mair & Marti, 2006; Ratten, 2012; Shepherd & Patzelt, 2011)—do not have reason to despair just yet, however. Other aspects of the perspective I have outlined may be hard to swallow for some, for example, the inclusion of an outcome criterion and the exclusion of organizational change per se as entrepreneurship. I remain optimistic, though, and ask doubtful readers with such objections to please try to make it through the next chapter. There are good reasons to believe that our differences will be sorted out there. Stay tuned!

References

Aldrich, H. E., & Ruef, M. (2006). *Organizations evolving*. Newbury Park, CA: Sage.

Amit, R., & Zott, C. (2001). Value drivers in e-business. *Strategic Management Journal, 22*, 493–520.

Austin, J., Stevenson, H., & Wei-Skillern, J. (2006). Social and commercial entrepreneurship: Same, different, or both? *Entrepreneurship: Theory and Practice, 30*(1), 1–22.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Baumol, W. J. (1990). Entrepreneurship: Productive, unproductive and destructive. *Journal of Political Economy, 98*(5), 893–921.

Baumol, W. J. (1993). *Entrepreneurship, management and the structure of payoffs*. Cambridge, MA: MIT Press.

Bhave, M. P. (1994). A process model of entrepreneurial venture creation. *Journal of Business Venturing, 9*, 223–242.

Bird, M., & Wennberg, K. (2014). Regional influences on the prevalence of family versus non-family start-ups. *Journal of Business Venturing, 29*(3), 421–436.

Carter, S. (2011). The rewards of entrepreneurship: Exploring the incomes, wealth, and economic well-being of entrepreneurial households. *Entrepreneurship: Theory and Practice, 35*(1), 39–55.

Cole, A. H. (1949). Entrepreneurship and entrepreneurial history. *Change and the Entrepreneur*. pp. 88-107.

Dahlqvist, J., & Wiklund, J. (2012). Measuring the market newness of new ventures. *Journal of Business Venturing, 27*(2), 185–196.

Davidsson, P. (1986). *Tillväxt i små företag: En pilotstudie om tillväxtvilja och tillväxtförutsättningar i små företag* (Small Firm Growth: A Pilot Study on Growth Willingness and Opportunity for Growth in Small Firms) (Studies in Economic Psychology No. 120). Stockholm: Stockholm School of Economics.

Davidsson, P. (2003). The domain of entrepreneurship research: Some suggestions. In J. Katz & D. Shepherd (Eds.), *Advances in entrepreneurship, firm emergence and growth* (Cognitive approaches to entrepreneurship research, Vol. 6, pp. 315–372). Oxford, UK: Elsevier/JAI Press.

Davidsson, P. (2006). Method challenges and opportunities in the psychological study of entrepreneurship. In J. R. Baum, M. Frese, & R. A. Baron (Eds.), *The psychology of entrepreneurship* (pp. 287–323). Mahway, NJ: Erlbaum.

Davidsson, P., Delmar, F., & Wiklund, J. (2002). Entrepreneurship as growth; growth as entrepreneurship. In M. A. Hitt, R. D. Ireland, S. M. Camp, & D. L. Sexton (Eds.), *Strategic entrepreneurship: Creating a new mindset* (pp. 328–342). Oxford, UK: Basil Blackwell & Mott, Ltd.

Davidsson, P., & Gordon, S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics, 39*(4), 853–876.
Dimov, D. (2007). Beyond the single-person, single-insight attribution in understanding entrepreneurial opportunities. *Entrepreneurship: Theory and Practice, 31*(5), 713–731.
Entrepreneurship Division, E. (2011). Domain statement of the Entrepreneurship Division. Retrieved from http://aom.org/DIG/.
Foss, N. J., & Klein, P. G. (2012). *Organizing entrepreneurial judgment: A new approach to the firm*. Cambridge, UK: Cambridge University Press.
Gartner, W. B. (1988). "Who is an Entrepreneur?" is the wrong question. *American Small Business Journal, 12*(4), 11–31.
Gartner, W. B. (1990). What are we talking about when we are talking about entrepreneurship? *Journal of Business Venturing, 5*, 15–28.
Greenwood, R., & Suddaby, R. (2006). Institutional entrepreneurship in mature fields: The big five accounting firms. *Academy of Management Journal, 49*(1), 27–48.
Hisrisch, R. D., Peters, M. P., & Shepherd, D. A. (2008). *Entrepreneurship* (7th ed.). New York, NY: McGraw-Hill.
Holcombe, R. G. (2003). The origins of entrepreneurial opportunities. *The Review of Austrian Economics, 16*(1), 25–43.
Kelley, D., Bosma, N., & Amoros, J. E. (2011). *Global Entrepreneurship Monitor. 2010 Global Report*. London: Global Entrepreneurship Research Association (GERA).
Kirzner, I. M. (1973). *Competition and entrepreneurship*. Chicago, IL: University of Chicago Press.
Kirzner, I. M. (1983). Entrepreneurs and the entrepreneurial function: A commentary. In J. Ronen (Ed.), *Entrepreneurship*. Lexington, MA: Lexington Books.
Lee, S.-H., Peng, M. W., & Song, S. (2013). Governments, entrepreneurs, and positive externalities: A real options perspective. *European Management Journal, 31*(4), 333–347.
Low, M. B., & MacMillan, I. C. (1988). Entrepreneurship: Past research and future challenges. *Journal of Management, 14*, 139–161.
Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. *Academy of Management Review, 21*(1), 135–172.
Mair, J., & Marti, I. (2006). Social entrepreneurship research: A source of explanation, prediction, and delight. *Journal of World Business, 41*(1), 36–44.
McMullen, J. S., & Shepherd, D. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review, 31*(1), 132–152.
Mises, L. (1949). *Human action*. New Haven, CT: Yale University Press.
Mitchell, R. K. (2011). Increasing returns and the domain of entrepreneurship research. *Entrepreneurship: Theory and Practice, 35*(4), 615–629.
Morris, M. H., Lewis, P. L., & Sexton, D. L. (1994). Reconceptualizing entrepreneurship: An input-output perspective. *Advanced Management Journal, 9*(Winter), 21–31.
Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. *Administrative Science Quarterly, 55*(3), 439–471.
Parker, S. C., & van Praag, C. M. (2012). The entrepreneur's mode of entry: Business takeover or new venture start? *Journal of Business Venturing, 27*(1), 31–46.
Plummer, L. A., Haynie, J. M., & Godesiabois, J. (2007). An essay on the origins of entrepreneurial opportunity. *Small Business Economics, 28*(4), 363–379.
Ratten, V. (2012). Guest editor's introduction: Sports entrepreneurship: Towards a conceptualisation. *International Journal of Entrepreneurial Venturing, 4*(1), 1–17.
Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal, 52*(4), 643–671.
Sarasvathy, S. D. (2000). Seminar on research perspectives in entrepreneurship. *Journal of Business Venturing, 15*, 1–57.
Sarasvathy, S. D. (2001). Causation and effectuation: Towards a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–288.

Sarasvathy, S. D. (2008). *Effectuation: Elements of entrepreneurial expertise*. Cheltenham, UK: Edward Elgar Publishing.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, UK: MA: Harvard University Press.

Semasinghe, D. M. (2011). *The role of idea novelty and relatedness on venture performance.* Doctoral dissertation, Queensland University of Technology, Brisbane.

Shane, S. A. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science, 11*(4), 448–469.

Shane, S. A. (2009). Why encouraging more people to become entrepreneurs is bad public policy. *Small Business Economics, 33*(2), 141–149.

Shane, S. A. (2012). Reflections on the 2010 AMR Decade Award: Delivering on the promise of entrepreneurship as a field of research. *Academy of Management Review, 37*(1), 10–20.

Shepherd, D. A., & Patzelt, H. (2011). The new field of sustainable entrepreneurship: Studying entrepreneurial action linking "what is to be sustained" with "what is to be developed". *Entrepreneurship: Theory and Practice, 35*(1), 137–163.

Stephan, U., & Uhlaner, L. M. (2010). Performance-based vs socially supportive culture: A cross-national study of descriptive norms and entrepreneurship. *Journal of International Business Studies, 41*(8), 1347–1364.

Stevenson, H. H., & Jarillo, J. C. (1990). A paradigm of entrepreneurship: Entrepreneurial management. *Strategic Management Journal, 11*, 17–27.

Suddaby, R. (2010). Editor's comments: Construct clarity in theories of management and organization. *Academy of Management Review, 35*(3), 346–357.

Wong, P. W., Ho, Y. P., & Autio, E. (2005). Entrepreneurship, innovation and economic growth: Evidence from GEM data. *Small Business Economics, 24*, 335–350.

Entrepreneurship as a Research Domain

2

Abstract

What is entrepreneurship research? Entrepreneurship as a research domain cannot be restricted to proven cases of entrepreneurship as defined in Chap. 1. This is because in order to understand the societal phenomenon as defined in Chap. 1, the research domain needs to understand also the choice not to engage in entrepreneurship and the reasons for failure to succeed at it. Combining ideas from prior literature, this chapter develops and discusses a delineation of the entrepreneurship research domain, focusing on the process of (completed or aborted) emergence of new economic ventures across organizational contexts.

2.1 Why Distinguish Between the Phenomenon and the Domain?

Now that we have devoted an entire chapter to discussing what entrepreneurship is, there shouldn't be much need for a chapter delineating the research domain "entrepreneurship," should there? Entrepreneurship as a research domain aims at better understanding of the phenomenon we call "entrepreneurship," so now that we "know" what it is, why not just go out and study it?

Paradoxically, the research domain cannot be equated to the study of empirical cases known to qualify under the definition of entrepreneurship that we discussed in the previous chapter. How can that be? First, we cannot learn why some initiatives manage to perform the societal function of entrepreneurship while others do not by only studying the successful cases. In order to see what sets them apart, we need to study failed attempts as well. That is, although including an outcome criterion is desirable when we discuss entrepreneurship as a societal phenomenon, it becomes a burden when we think of entrepreneurship as a research domain. This is further emphasized by our need to be able to study entrepreneurial processes

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_2

concurrently, in real time, before the outcome is known. It would be awkward indeed not to know until afterward whether one was doing "entrepreneurship research" or not. It would also be a bit hard on the researcher to require that every empirical study of "entrepreneurship" await and assess the outcome on every relevant level. Researchers must be allowed to go deeply into aspects of the process without following up on the outcomes—and still be acknowledged for doing "entrepreneurship research." That is, *attempts* to offer buyers new choices should suffice, and even processes that unintentionally (at early stages) could lead to such outcomes should qualify.

A second very important reason for making a distinction between the phenomenon and the research domain is that previous and current entrepreneurship practice does not necessarily have all the answers needed to develop *normative* theory about entrepreneurship (Fiet & Patel, 2008). That is, there may be better ways to learn meaningful things about entrepreneurship than finding real cases of "average practice" or even current "best practice." To study what successful entrepreneurs *have* done is important, but an even more important and interesting question is what *could* be done. As entrepreneurship scholars, we should be able to answer such questions, too, if we are the experts at abstracted sensemaking that we claim to be (Davidsson, 2002). This implies that the research domain should also include purely theoretical development (e.g., Baron, 2008; McMullen & Shepherd, 2006) and that empirical entrepreneurship research may be well advised to study not only naturally occurring entrepreneurial behavior but also induced entrepreneurial situations, such as experiments or simulations (cf. Crawford & McKelvey, 2010; Grégoire & Shepherd, 2012).

Yet other reasons for distinguishing between the phenomenon and the research domain also deserve mentioning. The behavior-plus-outcome definition lures one into a retrospective view that compresses time and deemphasizes the process aspects of entrepreneurship (Dimov, 2007; see also Chap. 8). It may therefore be advisable to employ a domain delineation that explicitly highlights the process nature of entrepreneurship. To study the processes as they happen is important in order to avoid selection and hindsight biases, topics we will develop in chapters to come. Further, the inclusion of a socially beneficial outcome clarifies the role of entrepreneurship in the economy. However, it may have detrimental effects on the long-term credibility of entrepreneurship research in political and fellow academic circles if we portrayed the micro-processes that we study as "good by definition." I suggested "pick your heroes!" once already. When the creation of new economic activity is studied in real time or the outcomes for other reasons have not been carefully assessed, it is advisable for entrepreneurship researchers to have an open attitude to the possibility of different types of outcomes on different levels. If a significant proportion of what we study in fact appear to be "redistributive ventures" (see Fig. 1.3 and surrounding text), we should see and report just that.

2.2 Previous Attempts at a Domain Delineation

Many readers may have been surprised—and more than so—that I did not include Shane and Venkataraman's (2000) entrepreneurship definition in the opening of the previous chapter. This would seem a peculiar omission as theirs has arguably been the most influential conceptual contribution to entrepreneurship research in recent years, achieving close to 8000 citations (at this point in time) on Google Scholar, receiving the AMR Decade Award for most cited paper (Shane, 2012; Venkataraman, Sarasvathy, Dew, & Forster, 2012), and having a major influence on the Entrepreneurship Division's new domain statement (cited above). The reason is not that the first version of this text was penned in 2003. I was well aware of the work back then, having heard Shane present it to us in Jönköping prior to publication in 1999 and also having discussed it with Venkat in 2000 when we were both keynote speakers at the same conference (in my current hometown Brisbane, as it were). Instead, the reason is that Shane and Venkataraman (2000)—originally Venkataraman (1997)—wisely suggested not just another attempt at defining the entrepreneurship phenomenon, but precisely the scholarly domain. So here is the more proper place to discuss their definition of the field of entrepreneurship, which reads:

> [T]he scholarly examination of how, by whom, and with what effects opportunities to create future goods and services are discovered, evaluated, and exploited (Venkataraman, 1997). Consequently the field involves the study of sources of opportunities; the processes of discovery, evaluation, and exploitation of opportunities; and the set of individuals who discover, evaluate, and exploit them. (Shane & Venkataraman, 2000: 218)

They further point out the following three sets of research questions as especially central: (1) why, when, and how opportunities for the creation of goods and services come into existence; (2) why, when, and how some people and not others discover and exploit these opportunities; and (3) why, when, and how different modes of action are used to exploit entrepreneurial opportunities. In the subsequent dialogue, they agreed with Zahra and Dess (2001) that the outcomes of the exploitation process represent a fourth important set of research questions, adding that outcomes on the level of industry and society should be considered as well (cf. Shane & Venkataraman, 2001; Venkataraman, 1997). In Davidsson (2003), I detailed the many merits I think this domain delineation has over what preceded it:

– They try to delineate the scholarly domain rather than suggesting yet another definition of the societal phenomenon. As just discussed, this distinction is important.
– Focusing on the creation of future goods and services, their delineation directs attention to the problem of emergence. This adds a distinctive feature to entrepreneurship research, an element that is missing in established theories in economics and management.
– They put the main focus on goods and services rather than including organizational change per se or creative behavior in any context. They thereby carve out a domain that has a manageable size and relatively clear boundaries and which is

consistent with Kirzner's (1973) notion that entrepreneurship is what drives the market process.

- While retaining an interest in individuals, they emphasize their actions (entrepreneurship) and fit with the specific "opportunity" rather than general characteristics of entrepreneurs. They thereby avoid the dead end of "trait research."[1]
- As to openness, their domain delineation includes two partly overlapping processes, discovery and exploitation. This refutes the view that discovery is instantaneous and that entrepreneurship consists solely of discovery (cf. Kirzner, 1973).
- No mention is made of the age, size, or ownership of the organizations in which "opportunities" are pursued. They even point out the existence of alternative modes of exploitation for given "opportunities" as an important research question. Hence, the stated domain includes corporate entrepreneurship. By implication, small business research is included only when it deals explicitly with discovery and exploitation of "opportunities" to create future goods and service.
- They do not include purposefulness in their domain delineation. They thereby avoid an overly rationalistic view and make room for the possibility of luck and serendipity in entrepreneurial processes.
- Finally, if we disregard for the moment their definition of "opportunity," their wording "…with what effects" makes the field open to different types of direct and indirect outcomes of processes of discovery and exploitation, e.g., satisfaction, learning, imitation, and retaliation in addition to financial success or failure. Importantly, the perspective suggests that an important task for entrepreneurship research is to assess outcomes not only on the microlevel but on other levels (e.g., societal wealth creation) as well (cf. Shane & Venkataraman, 2001).

That's not bad! As may be inferred from the above points, it is fair to say that it is largely in line with the entrepreneurship definition we discussed in Chap. 1. One of the few debatable points is affording general primacy to the microlevel by putting the main emphasis on the individual and the "opportunity." This does not seem to give much room for entrepreneurship research on more aggregate levels of analysis (cf. Zahra & Dess, 2001). As Shane (2012) explains, this was—in contrast to emphasis on cultural, political, economic, and industry factors at the time—to draw attention to the fact that entrepreneurship requires agency (and despite the title of Shane's, 2003 book, one couldn't or shouldn't expect one framework to be the ideal tool for all approaches to researching entrepreneurial phenomena).

[1] That's what I said in 2003. Although trait explanations will never be my favorite, I should clarify that subsequent meta-analyses (e.g., Collins, Hanges, & Locke, 2004; Rauch & Frese, 2007; Zhao & Seibert, 2006) and—ironically—Shane's own, recent work on the genetic factor in entrepreneurship (Nicolaou et al., 2011; Nicolaou et al., 2008) have to some degree reinstated stable person characteristics as explanations of entrepreneurial behavior and success.

A more important question mark is their definition of "opportunity" and, in fact, *any* use of that construct for research purposes. We'll get back to that point in Chap. 8. Right now, we have reason to consider Bill Gartner's many musings on entrepreneurship (e.g., Gartner, 1988, 1990, 1993, 2001; Gartner, Carter, & Reynolds, 2010; Gartner, Davidsson, & Zahra, 2006), which can also be regarded as attempts to delineate the field of research rather than defining or describing the phenomenon. Gartner's view—which he is careful to present as a suggestion for redirection rather than a formal "definition" (Gartner, 1988)—is that entrepreneurship is the creation (or emergence; cf. Gartner, 1993) of new organizations. This choice of focus appears to have had two origins. One was a perceived lack of treatment of organizational emergence in organization theory. Somehow, organizations were assumed to exist; theories started with existing organizations (cf. Katz & Gartner, 1988). The other was a frustration with the preoccupation that early entrepreneurship research had with personal characteristics of entrepreneurs. For these reasons, Gartner (1988) suggested that entrepreneurship research ought to focus on the *behaviors* in the process of organizational *emergence*. This focus on early-stage behavior has later been echoed—directly or indirectly and independently or as deliberate elaboration—by several other important contributors to conceptualizations of entrepreneurship (e.g., Alvarez, Barney, & Anderson, 2013; Baker & Nelson, 2005; Dimov, 2011; Foss & Klein, 2012; McMullen & Dimov, 2013; McMullen & Shepherd, 2006; Venkataraman, Sarasvathy, Dew, & Forster, 2012). So there would seem to be reason to consider what these people are saying.

In my opinion, Gartner's view certainly has a lot to commend it. For one thing, it has a clearly defined focus, addressing terrain that economics as well as management studies have treated in a stepmotherly fashion. This clear focus gives promise of providing unique contributions and avoiding overextension of the field of entrepreneurship research. Further, Gartner's view has a strong process orientation. The main problem I have with Gartner's (1988) approach is that whereas organizing is an important aspect of the exploitation process, he does not emphasize the discovery process (cf. Shane & Venkataraman's domain delineation above). Further, his approach directs no or only cursory attention to the possibility of alternative modes of exploitation for particular instances of new economic activity (Shane & Venkataraman, 2000; Wiklund & Shepherd, 2008). If interpreted as a delineation of the (entire) research domain, his take on entrepreneurship appears overly narrow in these regards. In short, I see Gartner's focus as the natural task for an organization theorist to take on *within* a somewhat broader domain.

Below, I will try to outline precisely that: a somewhat broader, yet sufficiently precise, domain delineation. What an incredibly pretentious thing to do! Well, the reason that I dare try is that I can stand on the shoulders of Gartner (1988) and Shane and Venkataraman (2000), as well as their predecessors and some other, later contributors. The little trick I will attempt below is the sewing together of their respective perspectives while ironing out the little wrinkles I think I've found, in order to arrive at a coherent domain delineation, tailor-made for entrepreneurship research.

2.3 My Suggested Domain Delineation

First, I take from Gartner (1988) the idea that entrepreneurship research should study behavior in the process of emergence. That introduces three very important components: *behavior*, *process*, and *emergence*. From Shane and Venkataraman (2000), I take the distinction between two subprocesses: *discovery* and *exploitation*. (If there were no prehistory, I would probably have chosen the labels *identification* and *implementation* instead. I include "evaluation" in the discovery process.) Further, in line with the view of entrepreneurship that we developed in Chap. 1, I agree with their notion that entrepreneurship research should not study only or primarily the emergence of new (independent) organizations, but the emergence of *new market offerings* (they say "new goods and services") through different *modes of exploitation*. Thus, when I speak of "new economic activities," I mean market-related activities and not internal reorganization, and when I say "new ventures," I do not restrict that notion to a particular mode of exploitation. From Venkataraman (1997), Shane and Venkataraman (2001), and Zahra and Dess (2001), I also adopt the idea that entrepreneurship research should study a variety of *outcomes* on different levels (see Chap. 7). The final element I take from Shane and Venkataraman (2000) is the fundamental assumption of disequilibrium: the economy *always* allows for *some* new initiatives to be successful. Further, and in part related to disequilibrium, I adopt the notion of *heterogeneity* of economic actors, which is prominent, e.g., in resource-based theory (Kraaijenbrink, Spender, & Groen, 2010).

To this I only need to add two little pieces. The first is to adopt the additional fundamental assumption that the economy is also characterized by *uncertainty* (McMullen & Shepherd, 2006). The second is that empirical entrepreneurship research need not and should not be restricted to the study of empirical cases known to qualify as "entrepreneurship" à la our definition of that phenomenon in the previous chapter. Entrepreneurship research should also study *failure* and *induced* processes of emergence. Oh, there is one more not so little thing, which has to do with uncertainty and failure among other things: I avoid that o-word, which has been so prominent in entrepreneurship research during the past decade (Hansen, Shrader, & Monllor, 2011; Karlsson, 2009; Short, Ketchen, Shook, & Ireland, 2010). Patience for now; we'll deal with it in Chap. 8.

Piecing it all together, I arrive at the following:

> Starting from assumptions of uncertainty, heterogeneity, and disequilibrium, the domain of entrepreneurship research encompasses the study of processes of (real or induced, and completed as well as terminated) emergence of new economic ventures, across organizational contexts. This entails the study of new venture ideas and their contextual fit; of actors and their behaviors in the interrelated processes of discovery and exploitation of such ideas, and of how the characteristics of ideas, actors and behaviors link to antecedents and outcomes on different levels of analysis.

Now, I can assure that there is no shortage of information hidden in those few lines, so it would be really nice if at this point the reader could stop, reflect, reread, and perhaps start counterarguing or asking follow-up questions. After playing that game for a couple of rounds, I'd be delighted if the reader imbibed my own elaborations below.

2.3.1 Uncertainty, Heterogeneity, and Disequilibrium

It could be debated whether one should really let this type of assumptions restrict a research domain. My rationale for including them is that I firmly believe that theories and research designs that do not build on such assumptions are unlikely to be useful tools for understanding the entrepreneurship phenomenon. Shane and Venkataraman (2000) have already made the disequilibrium argument quite well. As regards heterogeneity, assumptions that economic aggregates (such as an *industry* or *demand*) are made up of the sum of identical microlevel entities can hardly explain entrepreneurial action or success in meaningful, realistic ways.

Neither do I think it is illuminating for the understanding of entrepreneurship to start from a view of reality as characterized by certainty and calculable risk alone. I'd be the last to argue that all decisions for all actors are non-calculable. However, the situations in which behaviors aimed at creating new economic activity are undertaken often have this characteristic. It is also worth pointing out that we must allow theorists to build whatever fantasy worlds they like and then test the extent to which their theories have anything useful to say about the real world. This said, my belief is that to understand entrepreneurship, we need theories that admit that information collection and processing, careful planning, and calculation cannot give a conclusive and reliable answer as to whether an entrepreneurial initiative will be successful or not. Only (trial) implementation will tell. In short, entrepreneurial situations usually come with a substantial element of genuine, Knightian uncertainty (Knight, 1921). The future is not only unknown, but also unknowable (Foss & Klein, 2012; Sarasvathy, Dew, Velamuri, & Venkataraman, 2003).

On this point I disagree with the same Kirzner (1973)[2] that I leaned so heavily on in the first chapter. Very rarely are entrepreneurial situations certain in the way Kirzner portrays them. In one famous passage, Kirzner likens "entrepreneurial opportunity" with realizing that a free ten-dollar bill is resting in one's hand, ready to be grasped. If we should use the ten-dollar bill metaphor at all, I would suggest the true situation in real life is more like spotting the bill from your balcony. From that distance one would face the (calculable) risk that the bill was for anything from 1 to 100 dollars. Moreover, while you dash down the stairs, the wind might take it, or someone else might get it before you, or it turns out upon closer look that it was token money from some game or promotion rather than a real banknote. There is no way the finder can tell before she makes the decision to run down the stairs. In order to understand behaviors in such situations, it is important to start from a theoretical perspective that acknowledges or even emphasizes uncertainty.

[2] Please don't counterargue that I misinterpret Kirzner on the basis that in later works Israel Kirzner shows a greater understanding or appreciation of the dynamic and uncertain elements of the economy (Kirzner, 2009; Pollack, Vanepps, & Hayes, 2012). "Kirzner (1973)" is a theoretical argument, not a flesh-and-blood individual, and for all its merits, that argument is relatively insensitive to issues of time and uncertainty.

2.3.2 Processes of Emergence; Behaviors in the Interrelated Processes of Discovery and Exploitation

One of Gartner's (1988, 1993, 2001) great strengths is that he identified an important phenomenon—the process of emergence—on which other fields of research haven't done a very good job. Therefore, entrepreneurship research can make a real contribution if it takes on this challenge.

I agree with Shane and Venkataraman (2000) that both discovery and exploitation are required for entrepreneurship to happen and that both should be studied in entrepreneurship research. So again, I disagree with Kirzner's (1973, p. 47) claim that "Entrepreneurship does not consist of grasping a free ten-dollar bill which one has already discovered to be resting in one's hand; it consists of realizing that it is in one's hand and that it is available for the grasping." That is, he holds that entrepreneurship consists solely of discovery; exploitation presumably follows automatically or is "something else" altogether (Foss & Klein, 2012, p. 34). But returning to the balcony, nothing much happens if we just note that a ten-dollar bill seems to be lying down there, does it? How Kirzner makes restricting entrepreneurship to (instantaneous) discovery match his notion that entrepreneurship consists of the "competitive behaviors that drive the market process" beats me. There seems to be an underlying assumption in his reasoning that every actor who perceives an opportunity not only knows with certainty that it really is an opportunity but also necessarily acts upon it. Entrepreneurship researchers know that such is not the case. Many of us just have to exercise a little introspection to realize that.

Our emphasis on the interrelated processes of discovery and exploitation as new economic activities emerge implies that a very central set of research questions for entrepreneurship research concerns what individuals and other economic entities actually *do* when they initiate, refine, and realize ideas for new business ventures. This is still an area that needs much more investigation over and above the tentative steps that have been taken so far (e.g., Baker & Nelson, 2005; Bhave, 1994; Lichtenstein, Carter, Dooley, & Gartner, 2007; Sarasvathy, 2008).

The term *discovery* may be suspected to reflect an objectivist view on the entities that entrepreneurs act on when trying to get a new economic activity going. This is not a perspective I purport. Rather, like Shane and Eckhardt (2003), I use the term "discovery" to maintain consistency with prior literature, despite its potentially misleading connotations. Discovery refers to the conceptual side of venture development, from the identification of a rudimentary, initial idea to a fully developed business concept where many specific aspects of the operation are worked out in great detail, including how value is created for the customer and how the business will appropriate some of the value (Amit & Zott, 2001; G. George & Bock, 2011; Zott, Amit, & Massa, 2011). Importantly, discovery is a *process*—the venture idea is not formed as a complete and unchangeable entity at a sudden flash of divine insight (Ardichvili, Cardozo, & Ray, 2003; Dimov, 2007).

Exploitation is a negatively loaded word in some contexts and may therefore evoke negative associations. In the present context, it is a neutral term referring to action toward the realization of new economic activities. The exploitation process deals primarily with resource acquisition and coordination, as well as market

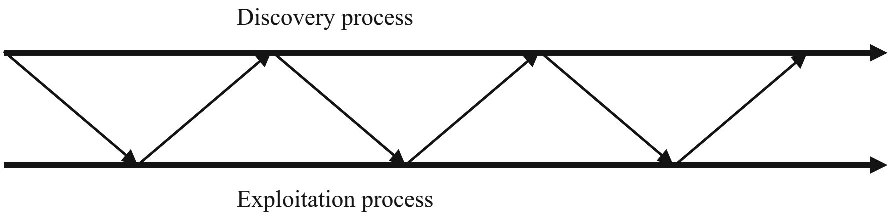


Fig. 2.1 The interrelationship between discovery and exploitation

making (see Brush, Greene, & Hart, 2001; Santos & Eisenhardt, 2009; Sarasvathy & Dew, 2005; Shane & Eckhardt, 2003). This includes all research questions pertaining to the organizing of new ventures, that is, the research agenda that Gartner (1988, 2001) emphasizes. Exploitation thus means the attempted realization of ideas. Like discovery, exploitation is a process that may or may not lead to the attainment of profit or other goals.

The emphasis on the interrelatedness of the two processes is based on empirical insights (Bhave, 1994; Sarasvathy, 2008). I think discovery and exploitation are best conceived of as overlapping processes. This is what Fig. 2.1 tries to portray.

For example, an entrepreneurial process may start with an individual perceiving what she thinks is an opportunity for a profitable business [discovery]. In the efforts to make this business happen, contacts with resource providers and prospective customers [exploitation] make it clear that the business as initially conceived will not be viable [feedback to discovery]. The individual changes the business concept accordingly [discovery] and continues her efforts to marshal and coordinate the resources needed for the realization of the revised business concept [exploitation]. Although the above process starts with an element of discovery, this is not necessarily always the case. Empirical research suggests that venture creation processes can follow almost any sequence (Carter, Gartner, & Reynolds, 1996; Gordon, 2012; Liao, Welsch, & Tan, 2005), and Bhave's (1994) study indicates that the insight (or discovery) that a problem solution one has developed for one's own needs may become a new venture idea rather late in a process that initially did not have the creation of a new venture as a goal.

Before closing this section, I should mention that "discovery" and "exploitation" do not represent the only way to conceptualize, subdivide, and label entrepreneurial processes. In my current research[3], I instead discuss aspects of "discovery" as *identification of new venture ideas* and *perception of opportunity* (the latter referring only to the individual's evaluation of a situation or an idea). Similarly, most of "exploitation" I reassign as aspects of entrepreneurial *action*, subdivided into *initiation*, *mode of exploitation*, and *process pattern*. For other conceptualizations of subprocesses, see, e.g., Bhave (1994); Delmar and Shane (2004); Gatewood, Shaver, and Gartner (1995); Katz and Gartner (1988); and Reynolds (2007).

[3] At the time of writing under review for *Journal of Prestigious Conceptual Work*, but at the time of reading possibly appearing in *Journal of Entrepreneurship & Bicycle Repair* (credit to Norris Krueger for this wonderful, generic title for journals-no-one-reads-and-which-you-don't-even-want-to-be-seen-in).

2.3.3 Real or Induced and Completed as well as Terminated

These are issues that we dealt with in the beginning of this chapter. The practicing entrepreneurs the world has seen so far do not necessarily have all the answers. That is, pure theory development and laboratory research methods may sometimes prove better avenues to arrive at normatively valid results and theories. As a case in point, one of the most interesting and influential developments in recent years, namely, Sarasvathy's theorizing about effectuation, emanates from research on induced (or hypothetical) entrepreneurial processes (Sarasvathy, 2008)[4].

Further, if we were to study successfully completed cases only, there is no telling whether terminated cases shared the same characteristics as the successful ones. This is especially important with regard to risk taking and its correlates. Risk taking should increase the span of possible outcomes. That is, the entrepreneur who takes risks should be rewarded with a greater likelihood of great success. At the same time, however, that entrepreneur incurs an increased risk of making a big belly flop. If our research design censors the terminated cases, we will systematically misinterpret the effects of risky strategies and actions (Yang & Aldrich, 2012).

2.3.4 Across Organizational Contexts

This has been thoroughly dealt with already. In Chap. 1, we parted with the "independent business" perspective on entrepreneurship. Shane and Venkataraman (2000) make a major point of this issue, emphasizing different modes of exploitation (such as internal venturing; licensing; the setting up of a new, independent firm) as a core set of research questions for entrepreneurship research. Recently, disappointment has been expressed that while entrepreneurship researchers seem to accept the importance of different modes on a conceptual level, the vast majority of empirical studies focus on independent start-ups (Foss & Klein, 2012; Shane, 2012). The emphasis on different organizational modes is in apparent conflict with Gartner's perspective. It is important to note, however, that Gartner's "creation of new organizations" should not necessarily be read as "creation of new, owner-managed firms." Gartner (1988, p. 28; cf. Gartner et al., 2010) explicitly discusses internal venturing. Although he—arguably with good reason—regards the emerging new firm as a particularly promising arena for studying it, his interest is in "organizing" in the Weickian sense (Gartner, 2001, p. 30, cf. Gartner & Carter, 2003), not necessarily the creation of *formal* and legally defined organizations.

[4]The underlying empirics were not presented in Sarasvathy (2001), presumably because in the absurd world of academic publishing, basing one's argument on armchair reasoning is sometimes more accepted than is basing it on careful and innovative empirical work that has visible warts.

"Across organizational contexts" has additional meaning beyond opening up for the study of discovery and exploitation both in emerging and existing firms, small and large, owner-managed or otherwise. This is also where we can start inviting back to the party those organizational changes in quadrant II of Fig. 1.2, which in the previous chapter were defined as not being instances of entrepreneurship. *Change* in the organizational context as *explicitly related to* the creation of new, market-related activity is clearly within the entrepreneurship research domain. Studies referred to by Ucbasaran, Westhead, and Wright (2001)(p. 64) showing that management buyouts are followed by increased development of new products are therefore examples of entrepreneurship research.

This example presumes a shift of "organizational context" within the same organizational entity. The emerging venture may also lead a life that cuts across several different organizations. What originates as an idea by an independent inventor may be acquired into an existing small firm, which is later acquired by a large organization, which decides to spin out this particular part of their business operations. This highlights the need for studies that use the emerging venture itself as the unit of analysis (Davidsson & Wiklund, 2001). Such studies would follow samples neither of individuals nor of organizations, but precisely of *new*, *emerging activities*—i.e., venture ideas and what evolves around them—from their conception and through whatever changes in human champions and organizational contexts might occur along the way. In some cases, what originated as a de novo start-up is transferred to an existing firm; in other cases, what originated within a firm may be spun out at an early stage. This is something we will also return to in later chapters.

2.3.5 New Economic Ventures; New Venture Ideas and Their Contextual Fit

"New economic ventures" include independent start-ups as well as new internal ventures and also new market offers that are so limited that the actors involved do not necessarily conceive of them as entire "ventures" (yet). However, in line with our placing entrepreneurship in a market context in the previous chapter, the suggested domain delineation is restricted to new *economic* ventures.

The reader may have noted that I have so far avoided the o-word as best I could, putting "opportunities" within quotation marks and/or only using it when citing others or occasionally for some individual's unproven, subjective *perception* of conditions as being potentially lucrative or otherwise beneficial. Instead, I have started to sneak in the concept n*ew venture idea* in its stead. This is very, very intentional, of course. In a draft version of this revised chapter, I decided this was where to slay the dragon (good luck with that, you say!), so I went on with a rant against the ills of "opportunity" over several pages. But I then decided to save you (and myself) from that grumpy-old-man detour and instead introduce a new chapter on "The Entrepreneurship Nexus" as Chap. 8.

As a sneak peek, instead of the elusive, contested, and otherwise problematic notion of "opportunity," I suggest we place the *new venture idea* (NVI) as the actor's main companion in the process. I conceive of the new venture idea as an "imagined new venture," defined as *a set of imagined combinations of product/service offerings, markets, and means of bringing these offerings into existence* (cf. Davidsson, 2015). As ideas evolve and (initially) reside only in actors' minds, this is certainly not an easy construct to work with. However, it is much more straightforward and less contentious than "entrepreneurial opportunity." A new venture idea is more than a mere dream but less than a manifest business model; it is a cognitive construct. The idea is not limited to increasing the efficiency and profitability of existing operations; it concerns introduction of new activity, although this new activity need not be innovative. New venture ideas may be easy or impossible to convert into operational ventures, and their (attempted) implementation may have good or bad consequences for the actors and for the economic system.

An NVI thus may start as a very rudimentary hunch about a technically possible product, or the perception of an unsolved problem that a market segment would be willing to pay to get solved, if one could find a solution to the problem. Over time it may be changed, honed, and elaborated to qualify as what others would called a *business concept* or a fully developed (conception of a) *business model*. The NVI is, so to speak, the focal object of the discovery and exploitation processes.

Referring back to Fig. 1.2, NVIs are ideas for new products or services or bundles thereof, introducing a new price/value relation, imitative entry, and new markets. Relating also to the heterogeneity issue, this shows that venture ideas come in different flavors. A seriously under-researched area, I would argue, concerns the characteristics of new venture ideas and how these characteristics relate to antecedents, behaviors, and outcomes. And I am as big a sinner as any; I can only point to a couple of published, empirical attempts at investigating the nature and effects of characteristics of NVIs (Davidsson, Hunter, & Klofsten, 2006; Samuelsson & Davidsson, 2009). New venture ideas have generalizable characteristics that may have generalizable effects. We need to conceptualize and operationalize such characteristics (novelty, scope, scalability; there should be many possibilities). By contrast, an abundance of studies have tried to assess the characteristics of entrepreneurs. Interestingly, this disproportionate interest in the individual is shared by diffusion research, where only about 1 % of the close to 4000 studies (now many more) have focused on the characteristics of the innovation (which may be possible to apply as NVI characteristics as well), whereas more than half of them focus on the individuals who adopted them (Rogers, 1995). An explanation for this might be the general human tendency that psychologists have dubbed "the fundamental attribution error." This is to seek explanations to events in terms of the characteristics of the individuals involved, also when structural or situational factors are the true determinants (Riggio & Garcia, 2009; Ross, 1977). Researchers beware!

Finally, as regards contextual fit, we have discussed fit between the actor and the new venture idea as the main entrepreneurship nexus. Other fit foci are certainly also possible. One obvious candidate would be the fit between NVIs and the environment. For all its qualities as an entrepreneurship hotbed, Silicon Valley might not have turned out the right environment for launching the Ice Hotel.

2.3.6 Actors (or Agents)

Having replaced one part of Shane and Venkataraman's (2000) "entrepreneurship nexus," we can now turn to the other half. They (p. 218) portray entrepreneurship as the nexus of lucrative you-know-what and enterprising individuals. I find the focus on individuals unnecessarily restrictive and somewhat misleading. For example, in the past decade we have seen a minor explosion of research on entrepreneurial *teams* (e.g., Harper, 2008; Ruef, Aldrich, & Carter, 2003; West, 2007) and the notion of *strategic entrepreneurship* (Hitt, Ireland, Sirmon, & Trahms, 2011; Ireland, Hitt, & Sirmon, 2003) where the firm or its management rather than single individuals is the most natural agent to focus on. The notion of "corporate entrepreneurship" was already well established in the year 2000 and has not died out since (Shepherd, Covin, & Kuratko, 2009).

Further, we have emphasized that different individuals and organizations can perform important roles at different stages of venture development and that it is important to study emerging ventures through changes of human champions and organizational affiliations. All in all, there are good reasons to generalize Shane and Venkataraman's individual to *actor* (*or agent*), denoting one or more individuals or firms in conjunction or in sequence, making room for team-based, corporate, and strategic forms of entrepreneurship. Notably, Shane and Venkataraman (2000) acknowledge this in the context of discussing modes of exploitation. Thus, the idea that the entrepreneurial function is performed not only by single individuals or in one particular organizational context has gained broad acceptance.

2.3.7 Antecedents and Outcomes on Different Levels of Analysis

This should be easy enough. It is standard research practice to ask questions about antecedents and outcomes. The emphasis on different levels of analysis makes our framework more inclusive. However, in order to qualify as entrepreneurship research on any given level, it needs to be *explicitly related to discovery and exploitation of new venture ideas*. Thus, we can re-invite the organizational issues in quadrant II of Fig. 1.2. The relationships between organizational characteristics and change on the one hand and discovery and exploitation of new venture ideas on the other are important questions for entrepreneurship research. However, those who think narrowly of entrepreneurship as dealing with the firm level of analysis should reflect on the fact that there are many other levels of analysis that are of equal relevance on the entrepreneurship research agenda. The potential and challenges involved in researching entrepreneurship on those *different* levels of analysis will be the central theme in chapters to come.

On the outcome side, this means that entrepreneurship research is very, very far from restricted to the question of the financial performance of new ventures or of firms. In Chap. 7, I will discuss the dependent variable in entrepreneurship research more thoroughly and argue that the relative financial performance of new ventures is not even one of the core outcomes. Entrepreneurship is about emergence; what comes after may well be of interest to entrepreneurship researchers but definitely not unique entrepreneurship terrain. At this stage of reading, migrants and visitors from strategic management should start to understand why the notion of entrepreneurship as a subfield of strategy is, should we say, a trifle incomplete. North American business schools may organize their departments any way they like; I don't think economists, sociologists, or psychologists around the world, who are interested in entrepreneurial phenomena, will care too much about that (although Baker & Pollock's, 2007, argument about hiring has some force). As we have delineated entrepreneurship research here, the core strategic management questions that are also entrepreneurship questions constitute but a corner of the totality of the entrepreneurship domain. It should be clear by now that many disciplines and subdisciplines cover different aspects of the research domain we have delineated—or *could* cover them as a natural part of their work. Hence, I do not believe in entrepreneurship as a distinct domain (Davidsson, 2003; Venkataraman, 1997) in an isolationist sense. It should be equally clear that no *one* other existing discipline or subdiscipline covers the entirety of what we here see as entrepreneurship research (Ács & Audretsch, 2010).

I would suggest that in showing a genuine interest in outcomes on different levels, and in providing a more refined and empirically informed view on "failure," entrepreneurship can distinguish itself from other fields and make strong contributions to social science at large (cf. Low, 2001; Sorenson & Stuart, 2008; Venkataraman, 1997). The question of when successful venture-level outcomes are and are not associated with successful outcomes on the societal level, and vice versa, is highly relevant but seldom asked. It is conceivable that under certain circumstances, the successful pursuit of ideas for new ventures does not benefit society (cf. Baumol, 1990). It is also possible to conceive of a situation where entrepreneurial efforts on the whole benefit society while at the same time the most likely outcome on the microlevel is a loss—and that therefore the rational decision is to refrain from entrepreneurship (cf. Olson in Sarasvathy, 2000, p. 35). Both of these situations represent important problems that entrepreneurship research can help societies to solve or avoid. The question of differential outcomes on different levels can also be asked from the perspective of the corporate manager, and this is to me an obvious question for strategic entrepreneurship: when and why does and does not new venturing—successful or not on the venture level—contribute to company performance? Again, because of potential learning and cannibalization, the answer is not a simple one to one relationship between venture- and organizational-level outcomes.

Referring back to Fig. 1.3, the issue of catalyst ventures, then, is of particular interest. Too narrow or simplistic a view on "failure" may lead to gross misrepresentation of the benefits of attempts to create new business activity, on micro- as well as aggregate levels (Levie, Don, & Leleux, 2011; Wennberg, Wiklund, DeTienne, & Cardon, 2010). What in a narrow perspective appears to be a "failure" may instead be a beneficial "catalyst" either because those directly involved in the "failure" learn for the future or because others imitate. One outcome we have already seen in deeper and more refined research into apparent "failure" is that pure failure as defined in Fig. 1.3 is far less common than previously thought. I think one of the first things entrepreneurship scholars should try to get rid of is the bias against failure. In addition to the "catalyst" potential, both theory and empirical evidence actually suggest that experimentation that may end in failure as well as the demise of less effective actors are necessary parts of a well-functioning market economy (Eliasson, 1991; Pe'er & Vertinsky, 2008; Schumpeter, 1934).

We should not forget that there are qualitatively different *types* of outcomes, too. Entrepreneurial processes do not only have financial outcomes and affect not only those directly involved in the project. Outcome assessment may also concern, e.g., satisfaction, learning, imitation, and retaliation. For researchers who have the creativity and guts to be unconventional, there are plenty of ~~opportunities~~, sorry, I should say "new research ideas," that await your discovery and exploitation.

2.4 Summary and Conclusion

In this chapter, I have argued that even though the objective of entrepreneurship research is to understand the phenomenon we call entrepreneurship, our research cannot be delimited to the study of proven empirical instances of entrepreneurship as defined in Chap. 1. Instead, I suggested the following domain delineation for entrepreneurship research, making a particular point of giving a central role to new venture ideas rather than to "opportunities":

> Starting from assumptions of uncertainty, heterogeneity, and disequilibrium, the domain of entrepreneurship research encompasses the study of processes of (real or induced, and completed as well as terminated) emergence of new economic ventures, across organizational contexts. This entails the study of new venture ideas and their contextual fit; of actors and their behaviors in the interrelated processes of discovery and exploitation of such ideas, and of how the characteristics of ideas, actors and behaviors link to antecedents and outcomes on different levels of analysis.

Building on a combination and extension of earlier contributions by Gartner and Shane and Venkataraman, the domain I suggest for entrepreneurship research is broader than either of these predecessors. This combination and extension allows the following, broadening reformulation of Shane and Venkataraman's (2000, 2001) core research questions. The focus on new venture ideas should not be interpreted as denying the existence of enabling and restricting external conditions or denying their importance. In Chap. 8, I will discuss them further under the label "external enablers."

- Why, when, where, how, and for whom do new venture ideas come into existence?
- Why, when, and how do individuals, organizations, regions, industries, cultures, and nations (or other units of analysis) differ in their propensity for identification, evaluation, development, and exploitation of new venture ideas?
- Why, when, and how are different modes of action used to identify, develop, and exploit new venture ideas?
- What are the outcomes on different levels (e.g., individual, organization, industry, society) of efforts to exploit new venture ideas?

My more inclusive attitude may lead to a less distinctive domain. If one looks closely at the above domain delineation, it is clear that most core research questions in entrepreneurship would fit in *some* existing discipline or subdiscipline. I see no problem in that; the more scholars from various disciplines invest in understanding entrepreneurship, the happier I am! Entrepreneurship as a distinctive domain, to me, is not about being exclusive but about trying to make a well-defined contribution to something bigger. I think it is also clear that entrepreneurship is not in its entirety a subdivision of any *one* established discipline or field of research. If left solely to within-discipline work (Sorenson & Stuart, 2008), there is no guarantee that a lot of research would be conducted on the most central questions of entrepreneurship, as we have here outlined that domain. Many of these questions may be peripheral to every discipline (cf. Ács & Audretsch, 2003). *A failure to collectively cover the entrepreneurship research agenda is neither a problem nor a shortcoming on the part of the existing disciplines*. Well, at least it is not a problem for any individual scholar within them. When maximizing knowledge development about the entrepreneurship phenomenon is the vantage point (Wiklund, Davidsson, Audretsch, & Karlsson, 2011; Zahra & Wright, 2011), however, this is a very real and important problem. This is the most important raison d'être for entrepreneurship research as a distinctive domain and research community. Therefore, I think we need to be a multidisciplinary community of scholars who dedicate ourselves to this phenomenon and who interact enough in order to speak roughly the same language. In line with this notion, our leading journal, the *Journal of Business Venturing*, now explicitly defines itself as a multidisciplinary journal (www.journals.elsevier.com/journal-of-business-venturing) and is organized accordingly.

Now, after this long warming up, it's about time we get to the real stuff: method-related challenges of entrepreneurship research. So that's what we'll turn to next and for the remainder of this book: empirical design and analysis issues. Oh, well, perhaps not; there was this little thing called "theory" that we have to deal with first.

References

Ács, Z. J., & Audretsch, D. B. (2003). Editor's introduction. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research*. Dordrecht, NL: Kluwer Academic.

Ács, Z. J., & Audretsch, D. B. (Eds.). (2010). *Handbook of entrepreneurship research: An interdisciplinary survey and introduction* (2nd ed.). New York, NY: Springer-Verlag.

Alvarez, S. A., Barney, J. B., & Anderson, P. (2013). Forming and exploiting opportunities: The implications of discovery and creation processes for entrepreneurial and organizational research. *Organization Science, 24*, 301–317.

Amit, R., & Zott, C. (2001). Value drivers in e-business. *Strategic Management Journal, 22*, 493–520.

Ardichvili, A., Cardozo, R., & Ray, S. (2003). A theory of entrepreneurial opportunity identification and development. *Journal of Business Venturing, 18*(1), 105–123.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329–366.

Baker, T., & Pollock, T. G. (2007). Making the marriage work: The benefits of strategy's takeover of entrepreneurship for strategic organization. *Strategic Organization, 5*(3), 297–312.

Baron, R. A. (2008). The role of affect in the entrepreneurial process. *Academy of Management Review, 33*(2), 328–340.
Baumol, W. J. (1990). Entrepreneurship: Productive, unproductive and destructive. *Journal of Political Economy, 98*(5), 893–921.
Bhave, M. P. (1994). A process model of entrepreneurial venture creation. *Journal of Business Venturing, 9*, 223–242.
Brush, C. G., Greene, P. G., & Hart, M. M. (2001). From initial idea to unique advantage: The entrepreneurial challenge of constructing a resource-base. *Academy of Management Executive, 15*(1), 64–78.
Carter, N. M., Gartner, W. B., & Reynolds, P. D. (1996). Exploring start-up event sequences. *Journal of Business Venturing, 11*, 151–166.
Collins, C. J., Hanges, P. J., & Locke, E. A. (2004). The relationship of achievement motivation to entrepreneurial behavior: A meta-analysis. *Human Performance, 17*(1), 95–117.
Crawford, G. C., & McKelvey, B. (2010). *Using simulation experiments to build and test entrepreneurship theories*. Paper presented at the BCERC conference, Lausanne, Switzerland. (A summary is available in Frontiers of Entrepreneurship Research, 30, and downloadable from http://digitalknowledge.babson.edu/fer/vol30/iss20/3).
Davidsson, P. (2002). What entrepreneurship research can do for business and policy practice. *International Journal of Entrepreneurship Education, 1*(1), 5–24.
Davidsson, P. (2003). The domain of entrepreneurship research: Some suggestions. In J. Katz & D. Shepherd (Eds.), *Advances in entrepreneurship, firm emergence and growth* (Cognitive approaches to entrepreneurship research, Vol. 6, pp. 315–372). Oxford, UK: Elsevier/JAI Press.
Davidsson, P. (2015). Entrepreneurial opportunities and the entrepreneurship nexus: A reconceptualization. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.002.
Davidsson, P., Hunter, E., & Klofsten, M. (2006). Institutional forces: The invisible hand that shapes venture ideas? *International Small Business Journal, 24*(2), 115–131.
Davidsson, P., & Wiklund, J. (2001). Levels of analysis in entrepreneurship research: Current practice and suggestions for the future. *Entrepreneurship: Theory and Practice, 25*(4), 81–99.
Delmar, F., & Shane, S. A. (2004). Legitimating first: Organizing activities and the survival of new ventures. *Journal of Business Venturing, 19*, 385–410.
Dimov, D. (2007). Beyond the single-person, single-insight attribution in understanding entrepreneurial opportunities. *Entrepreneurship: Theory and Practice, 31*(5), 713–731.
Dimov, D. (2011). Grappling with the unbearable elusiveness of entrepreneurial opportunities. *Entrepreneurship: Theory and Practice, 35*(1), 57–81.
Eliasson, G. (1991). Modeling the experimentally organized economy: Complex dynamics in an empirical micro-macro model of endogenous economic growth. *Journal of Economic Behavior and Organization, 16*, 153–182.
Fiet, J. O., & Patel, P. C. (2008). *Prescriptive entrepreneurship*. Cheltenham, UK: Edward Elgar Publishing.
Foss, N. J., & Klein, P. G. (2012). *Organizing entrepreneurial judgment: A new approach to the firm*. Cambridge, UK: Cambridge University Press.
Gartner, W. B. (1988). "Who is an Entrepreneur?" is the wrong question. *American Small Business Journal, 12*(4), 11–31.
Gartner, W. B. (1990). What are we talking about when we are talking about entrepreneurship? *Journal of Business Venturing, 5*, 15–28.
Gartner, W. B. (1993). Words lead to deeds: Towards an organizational emergence vocabulary. *Journal of Business Venturing, 8*, 231–239.
Gartner, W. B. (2001). Is there an elephant in entrepreneurship research? Blind assumptions in theory development. *Entrepreneurship: Theory and Practice, 25*(4), 27–39.
Gartner, W. B., & Carter, N. (2003). Entrepreneurial behavior and firm organizing processes. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research*. Dordrecht, NL: Kluwer Academic.

Gartner, W. B., Carter, N. M., & Reynolds, P. D. (2010). Entrepreneurial behavior: Firm organizing processes. In Z. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research* (pp. 99–127). New York, NY: Springer.

Gartner, W. B., Davidsson, P., & Zahra, S. A. (2006). Are you talking to me? The nature of community in entrepreneurship scholarship. *Entrepreneurship: Theory and Practice, 30*(3), 321–331.

Gatewood, R. D., Shaver, K. G., & Gartner, W. B. (1995). A longitudinal study of cognitive factors influencing start-up behaviors and success at new venture creation. *Journal of Business Venturing, 10*, 371–391.

George, G., & Bock, A. J. (2011). The business model in practice and its implications for entrepreneurship research. *Entrepreneurship: Theory and Practice, 35*(1), 83–111.

Gordon, S. R. (2012). *Dimensions of the venture creation process: Amount, dynamics, and sequences of action in nascent entrepreneurship*. Doctoral dissertation, Queensland University of Technology, Brisbane.

Grégoire, D. A., & Shepherd, D. A. (2012). Technology-market combinations and the identification of entrepreneurial opportunities: An investigation of the opportunity-individual nexus. *Academy of Management Journal, 55*(4), 753–785.

Hansen, D. J., Shrader, R., & Monllor, J. (2011). Defragmenting definitions of entrepreneurial opportunity. *Journal of Small Business Management, 49*(2), 283–304.

Harper, D. A. (2008). Towards a theory of entrepreneurial teams. *Journal of Business Venturing, 23*(6), 613–626.

Hitt, M. A., Ireland, R. D., Sirmon, D. G., & Trahms, C. A. (2011). Strategic entrepreneurship: Creating value for individuals, organizations and society. *Academy of Management Perspectives, 25*(2), 57–75.

Ireland, R. D., Hitt, M. A., & Sirmon, D. G. (2003). A model of strategic entrepreneurship: The construct and its dimensions. *Journal of Management, 29*(6), 963–989.

Karlsson, T. (2009). *Emergence and development of entrepreneurship research 1989–2007: Keywords and collocations. Working Paper 2009–7*. Lund, Sweden: School of Economics and Management, Lund University.

Katz, J., & Gartner, W. B. (1988). Properties of emerging organizations. *Academy of Management Review, 13*(3), 429–441.

Kirzner, I. M. (1973). *Competition and entrepreneurship*. Chicago, IL: University of Chicago Press.

Kirzner, I. M. (2009). The alert and creative entrepreneur: A clarification. *Small Business Economics, 32*(2), 145–152.

Knight, F. (1921). *Risk, uncertainty and profit*. New York, NY: Hougthon Mifflin.

Kraaijenbrink, J., Spender, J. C., & Groen, A. J. (2010). The resource-based view: A review and assessment of its critiques. *Journal of Management, 36*(1), 349–372.

Levie, J., Don, G., & Leleux, B. (2011). The new venture mortality myth. In K. Hindle & K. Klyver (Eds.), *Handbook of research on new venture creation* (pp. 194–215). Cheltenham, UK: Edward Elgar Publishing.

Liao, J., Welsch, H., & Tan, W. L. (2005). Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns. *Journal of High Technology Management Research, 16*(1), 1–22.

Lichtenstein, B. B., Carter, N. M., Dooley, K. J., & Gartner, W. B. (2007). Complexity dynamics of nascent entrepreneurship. *Journal of Business Venturing, 22*(2), 236–261.

Low, M. B. (2001). The adolescence of entrepreneurship research: specification of purpose. *Entrepreneurship: Theory and Practice, 25*, 17–25.

McMullen, J. S., & Dimov, D. (2013). Time and the entrepreneurial journey: The problems and promise of studying entrepreneurship as a process. *Journal of Management Studies, 50*(8), 1481–1512.

McMullen, J. S., & Shepherd, D. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review, 31*(1), 132–152.

Nicolaou, N., Shane, S. A., Adi, G., Mangino, M., & Harris, J. (2011). A polymorphism associated with entrepreneurship: evidence from dopamine receptor candidate genes. *Small Business Economics, 36*(2), 151–155.

Nicolaou, N., Shane, S. A., Cherkas, L., Hunkin, J., & Spector, T. D. (2008). Is the tendency to engage in entrepreneurship genetic? *Management Science, 54*(1), 167–179.

Pe'er, A., & Vertinsky, I. (2008). Firm exits as a determinant of new entry: Is there evidence of local creative destruction? *Journal of Business Venturing, 23*(3), 280–306.

Pollack, J. M., Vanepps, E. M., & Hayes, A. F. (2012). The moderating role of social ties on entrepreneurs' depressed affect and withdrawal intentions in response to economic stress. *Journal of Organizational Behavior, 33*(6), 789–810.

Rauch, A., & Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology, 16*(4), 353–385.

Reynolds, P. D. (2007). New firm creation in the United States: A PSED I overview. *Foundations and Trends in Entrepreneurship, 3*(1), 1–150.

Riggio, H. R., & Garcia, A. L. (2009). The power of situations: Jonestown and the fundamental attribution error. *Teaching of Psychology, 36*(2), 108–112.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: The Free Press.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–240). Orlando, FL: Academic Press.

Ruef, M., Aldrich, H. E., & Carter, N. M. (2003). The structure of organizational founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review, 68*(2), 195–222.

Samuelsson, M., & Davidsson, P. (2009). Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics, 33*(2), 229–255.

Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal, 52*(4), 643–671.

Sarasvathy, S. D. (2000). Seminar on research perspectives in entrepreneurship. *Journal of Business Venturing, 15*, 1–57.

Sarasvathy, S. D. (2001). Causation and effectuation: Towards a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–288.

Sarasvathy, S. D. (2008). *Effectuation: Elements of entrepreneurial expertise*. Cheltenham, UK: Edward Elgar Publishing.

Sarasvathy, S. D., & Dew, N. (2005). New market creation through transformation. *Journal of Evolutionary Economics, 15*(5), 533–565.

Sarasvathy, S. D., Dew, N., Velamuri, R., & Venkataraman, S. (2003). Three views of entrepreneurial opportunity. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research*. Dordrecht, NL: Kluwer.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, UK: MA: Harvard University Press.

Shane, S. A. (2003). *A general theory of entrepreneurship: The individual-opportunity nexus*. Cheltenham, UK: Edward Elgar Publishing.

Shane, S. A. (2012). Reflections on the 2010 AMR Decade Award: Delivering on the promise of entrepreneurship as a field of research. *Academy of Management Review, 37*(1), 10–20.

Shane, S. A., & Eckhardt, J. (2003). The individual-opportunity nexus. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research* (pp. 161–194). Dordrecht, NL: Kluwer Academic.

Shane, S. A., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review, 25*(1), 217–226.

Shane, S. A., & Venkataraman, S. (2001). Entrepreneurship as a field of research: A response to Zahra and Dess, Singh, and Erikson. *Academy of Management Review, 26*(1), 13–16.

Shepherd, D. A., Covin, J. G., & Kuratko, D. F. (2009). Project failure from corporate entrepreneurship: Managing the grief process. *Journal of Business Venturing, 24*(6), 588–600.

Short, J. C., Ketchen, D. J., Jr., Shook, C. L., & Ireland, R. D. (2010). The concept of "opportunity" in entrepreneurship research: Past accomplishments and future challenges. *Journal of Management, 36*(1), 40–65.

Sorenson, O., & Stuart, T. E. (2008). Entrepreneurship: A field of dreams? *Academy of Management Annals, 2*(1), 517–543.

Ucbasaran, D., Westhead, P., & Wright, M. (2001). The focus of entrepreneurship research: Contextual and process issues. *Entrepreneurship: Theory and Practice, 25*(4), 57–80.

Venkataraman, S. (1997). The distinctive domain of entrepreneurship research: An editor's perspective. In J. Katz & J. Brockhaus (Eds.), *Advances in entrepreneurship, firm emergence, and growth* (Vol. 3, pp. 119–138). Greenwich, CT: JAI Press.

Venkataraman, S., Sarasvathy, S. D., Dew, N., & Forster, W. R. (2012). Reflections on the 2010 AMR Decade Award: Whither the promise? Moving forward with entrepreneurship as a science of the artificial. *Academy of Management Review, 37*(1), 21–33.

Wennberg, K., Wiklund, J., DeTienne, D. R., & Cardon, M. S. (2010). Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing, 25*(4), 361–375.

West, G. P. (2007). Collective cognition: When entrepreneurial teams, not individuals, make decisions. *Entrepreneurship: Theory and Practice, 31*(1), 77–102.

Wiklund, J., Davidsson, P., Audretsch, D. B., & Karlsson, C. (2011). The future of entrepreneurship research. *Entrepreneurship: Theory and Practice, 35*(1), 1–9.

Wiklund, J., & Shepherd, D. A. (2008). Portfolio entrepreneurship: Habitual and novice founders, new entry, and mode of organizing. *Entrepreneurship: Theory and Practice, 32*(4), 701–725.

Yang, T., & Aldrich, H. E. (2012). Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing, 27*(4), 477–492.

Zahra, S. A., & Dess, G. G. (2001). Entrepreneurship as a field of research: Encouraging dialogue and debate. *Academy of Management Review, 26*(1), 8–10.

Zahra, S. A., & Wright, M. (2011). Entrepreneurship's next act. *Academy of Management Perspectives, 25*(4), 67–83.

Zhao, H., & Seibert, S. E. (2006). The big five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology, 91*(2), 259.

Zott, C., Amit, R., & Massa, L. (2011). The business model: Recent developments and future research. *Journal of Management, 37*(4), 1019–1042.

This Thing Called "Theory"

3

Abstract

How and why can theory help us understand entrepreneurial phenomena? The contemplative nature of theory may seem antithetical to the bold action associated with entrepreneurship. Theory is important in research because it is the abstracted and reflected sensemaking of theory that makes empirical observations meaningful. However, an exaggerated focus on "theoretical contributions" can also hamper the development of a scholarly field. This chapter discusses what theory is and is not: its various roles in the research process and the pros and cons of focusing on theory. The specific requirements on theoretical tools suitable for the study of entrepreneurship are also considered.

3.1 Confessions of a Sinner

I confess! I am a sinner! I haven't always practiced what I preach as far as theory goes. Some of the projects I have been involved in, and where I have enjoyed access to excellent empirical data, haven't been as theory driven or theory developing as they should. Pressed for time and in the face of intriguing empirical relationships, I have sometimes neglected the conceptual side of research. But that is really my loss. No matter how intriguing an empirical result may seem here and now, it is the sensemaking of theory that makes it travel through space and stand the test of time. It is theoretical interpretations that uncover the implications of empirical results, so that they can properly guide practitioner behavior and the design of continued research efforts. In short, theory is crucially important. In the absence of theory, empirical research will be poorly designed, and the results will have little meaning. Besides, I'm not the worst of sinners. As will be boasted below, one of the main contributions of my very first attempts in the field of entrepreneurship research—i.e., my dissertation project—was to increase the level of abstraction (Davidsson, 1991) Similarly, when we set up an Australian counterpart project to the Panel

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_3

Study of Entrepreneurial Dynamics II (Reynolds & Curtin, 2008)—the CAUSEE project—one distinct feature was to make it more explicitly theory driven and theory testing than earlier panel studies of nascent ventures and—entrepreneurs (Davidsson, 2006; Davidsson & Gordon, 2012).

Broadly speaking, there are two fundamentally different types of research. In (prototypical) *basic research*, we want to understand, or find meaningful ways to discuss and communicate, how some aspect of the world works "in general," where general can be broader (e.g., assuming validity for "all firms") or narrower (assuming validity only for, e.g., young firms, or young firms in mature industries, or young firms in mature industries in transition economies). For basic research, the value of theory should be pretty obvious; the goal is not to list the facts pertaining to specific firms in a particular place at a particular time, but to develop a conceptual toolbox that allows us to make probabilistically true statements about the behavior, problems, or success of many different firms (if that is what the theory is about) in many different places, in the past, present, and future. *Applied research* uses scholarly tools (theories, methods) to investigate an issue with the aim to help solving a particular problem for a particular actor. It is thus *not* the case that theory is not useful or needed in applied research. In (prototypical) applied research, we use theory and methods to select what facts to look for, to get those facts right, and to understand what they imply for particular cases at a particular time. If we don't believe much in objective "facts" or our ability to use them to shape our fate, theory is still useful for unearthing otherwise hidden issues that are somehow important "here and now" and provide a language for us to discuss them. That is, theory helps in designing applied research and in interpreting their results. Moreover, as I argue below, there is no realistic alternative to letting theory guide one's research. Research studies are not undertaken with or without theoretical input; what varies are how articulated that theory is and how useful it is for solving or illuminating the issue at hand.

So despite sometimes having sinned, and knowing that there are colleagues who would raise an eyebrow at my writing this particular chapter, I am a great fan of theory. Below, I will try to explain why. I will start with an attempt to demystify theory—something which seems to be needed at least when addressing students. I will then dwell on the advantages of abstractions for some time, before turning to the different roles of theory in the research process. After that, we should be ready to discuss whether entrepreneurship needs its own theory development. Perhaps theories from existing disciplines and fields of research suffice?

But first, a little update in response to developments over the past decade.

3.2 Theory Is No Mystery

Back to base camp! Let's start here: theory is not the opposite of reality and not the opposite of practice. These are the first things I tell students about theory. Theory is not some mystical, unworldly exercise of ivory-tower academics. On the contrary, we all use theory all the time. There's no escape! If I turn on my computer, wait, enter my credentials, wait, open a web browser, enter the URL www.dn.se (the leading daily newspaper in Sweden), and hit "Enter," I do so because I have an experience-based "lay theory" which says that if I follow these steps, the newspaper with all its wisdom will be revealed to me. But the outcome of this sequence of actions is not a *fact* until I have really undertaken them and been rewarded with a successful result. There are many reasons why the theory might not work a particular day, i.e., not give an accurate prediction. The paper's website may be down or hijacked by hackers, or my connection may be faulty, or an electrical outage might cut my session short, or an outburst of strong radiation from unusual solar activity might distort the data traffic so badly that I would have to give up, or an asteroid might hit and wipe us all out before I even got to hit "Enter." So my actions are guided by theory, not certainty. In fact, when viewed this way, all of our goal-directed behavior is governed by some kind of theory about the workings of the world. If no form of theory should guide action, random behavior would be the only alternative. For example, out of all the possible random actions available, I could have verbally and physically abused the object in front of me[1]—which happened to be a computer—in the hope of getting to read my cherished newspaper.

So the issue in design and interpretation of research is not *whether* it is guided by theory, but how articulated and suitable for the purpose that theory is. As I see it, theories are best regarded as *tools*. "Science is tooled knowledge," as Schumpeter (1954, p. 7) puts it. Scholarly theories have usually been developed by clever, hard-working people trying to do their very best. This is a reason to be respectful toward such theories and not disregard them until one is certain one has a better alternative. However, the toolmakers cannot possibly know exactly what tools are needed to solve your particular problem. It is not the toolmaker's fault if you make a mess trying to open a can with a hammer. The researcher is well advised to rummage the toolbox a little more or have a look in another toolbox (another discipline or field of research) before trying such questionable solutions. However, sometimes the available toolboxes really do not contain any suitable theory. This is when a sound amount of creative disrespect may be needed vis-à-vis theory. Theories are not untouchables; one might need to adapt and combine them in order to get the tools needed to solve the problem. Using theory is not a reason to stop thinking for oneself! Sometimes, entirely new tools may need to be developed. However, if you ask me, it is more often researchers' ignorance or creative itches rather than real lack of available tools that make them try to develop genuinely new theory. Mind you, there are—and should be—more skilled tool users than toolmakers (cf. Sutton & Staw, 1995).

Theories are not the opposite of reality, but neither are they perfect images of reality. In order to be useful, they must be simplified representations of reality. As Bacharach (1989) relates, theory prevents the observer from being dazzled by the complexity of the social realities we study. Some people get annoyed when some aspects of a theory do not hold true with respect to a specific case or that the

[1] I normally apply this latter behavioral sequence only to printers, not computers themselves. After multiple replications, I have come to the conclusion that with respect to achieving goal attainment it is not a valid theory, or at least printers lie outside its boundaries.

theory does not account for the full richness of empirical cases. This could be a sign of deficiency on the part of the theory, but more likely, it shows that those people simply have not understood what theory is and what it is good for (and regrettably this problem is not confined to students). When a theory fits perfectly with a specific case, it is in all likelihood no longer a theory but an idiosyncratic description of little use for understanding *other* cases. A useful theory must abstract and generalize and thus neglect many of the fine details. Theorizing involves *abstraction*. This is what we will turn to in the next subsection.

3.3 The Need for Abstraction and Understanding

In all honesty, the newspaper reading theory above wasn't much of a theory, after all. The example made the point that the behavior was guided by experience-based guesses, not facts. It may be best described as a *behavioral script* or an *empirical generalization*, which is a forerunner to—or a very primitive form of—theory. There are different definitions and descriptions of what is needed for something to qualify as a "theory." Bacharach (1989, p. 496) suggests that "Theory is a statement of relations among concepts within a set of boundary assumptions and constraints." Locke (2007, p. 889) refers to the lexical definition "Systematically organized knowledge applicable in a relatively wide variety of circumstances, especially a system of assumptions [based on facts?], accepted principles, and rules of procedure devised to analyze, predict, or otherwise explain the nature of behavior of a specified set of phenomena." Most definitions of theory require the following two elements:

1. A set of well-defined, abstracted concepts
2. A set of well-specified relationships among those concepts

This is something that can in part be expressed as a formula or graphically in the form of a boxes-and-arrows diagram. In other words, it can be expressed as a *model*. Most scholars would say that a theory is more than that (see Sutton & Staw, 1995). For example:

3. Some demarcation of within which context or under what conditions the theory is supposed to hold (boundary conditions)
4. A deeper *understanding* of why the relationships exist and what they imply

Concepts. Concepts are incredibly useful abstractions that allow us to exchange ideas about things outside of our immediate, shared perceptions and sensations. Even when the relationships of our theories turn out to have weak or unknown generalizability, it is still of great value to be able to talk about "profit," "costs," "job satisfaction," and "sustainable competitive advantage" instead of having to point at a pile of money or a set of numbers in a profit statement or listen to repeated, elaborate descriptions of how this and that person feels about their job or the specific circumstances that give particular firms a privileged position in the marketplace. "Concept" and "construct" are used differently by different authors (cf. Bacharach, 1989; Locke, 2012). I use "concept" as the broader of the two. There are many different, tangible entities that are all captured by the concept "dog," which is an abstraction. "Entrepreneurial bricolage" (Baker & Nelson, 2005) is an example of a concept where the empirical manifestations are less tangible and less agreed upon. It is an example of a construct, which can only be measured indirectly (Senyard, Baker, Steffens & Davidsson, 2014).

To see the power of concepts, consider the following example. Some 20 years ago, Husqvarna AB, which happens to be located in what was then my hometown, introduced a radically new type of lawn mower. This machine, called the Solar Mower, was like a modern sheep, walking the garden at random and on its own, cutting the grass little by little. Solar cells generated the power needed, and random walk software in the small internal computer made sure the device cut the entire lawn, within magnetic cord demarcations. For safety and security, the device was equipped with an anti-theft alarm and cut the grass with a piece of plastic string rather than a metal knife. Got it? Seen one? The product now exists in a number of modified forms under a variety of brand names and has finally taken off big time, at least in Sweden. There are many reasons why it was not a big seller early on or in all markets. One particular problem in northern Sweden at the time was that this lawn mower was not "macho" enough. Up there, tractor-like machines—preferably bigger than your neighbors'—were the name of the game. In the UK, market acceptance was slow for seemingly an entirely different reason: it didn't make stripes! I could not believe the seriousness of this objection when the product manager first told me. Shortly thereafter, I visited the UK and told a colleague—Peter Rosa—about this innovation. Whereupon his wife, who overheard the conversation, made that exact remark: "But then it doesn't make stripes!?" Apparently, UK homeowners have a very strong preference for stripes on their lawns, just like on the soccer fields you may have seen on TV. Presumably, these stripes should preferably be straighter than your neighbo*u*rs'.

What do we learn from this? Don't introduce sissy lawn mowers up north in Sweden? If you are a lawn mower manufacturer considering the UK market, don't forget the stripes? As yet, all we have is a couple of cute little marketing anecdotes. We could try to incorporate these two events in a "lawn mower launch theory," but that theory would just be a (long) list of historical particularities that have somehow, somewhere hampered or facilitated the market acceptance of a new lawn mower (and lists are another of Sutton & Staw's, 1995, examples of what theory is not). However, if instead from instances like these we are able to distill the abstracted concept *compatibility*, we can see that the "macho" and "stripes" issues are in a sense aspects of the *same* type of problem: lacking compatibility with prevailing norms. See what has happened here? We just made a giant leap from perhaps funny but largely useless anecdotes about particular instances to having a concept that is useful not only with respect to historical cases or to lawn mowers but one that applies to the diffusion of any innovation in any society at any time. *Compatibility* is something worth considering with respect to every innovation there ever was and ever will be and probably differentially so for different potential customers or user groups. Does the innovation fit with technical systems and behavioral patterns they are already applying? Does using it fit with prevailing norms? Practitioners and researchers involved in other innovations in the future now have a meaningful issue to consider, along with the issues of *relative advantage*, *complexity*, *trialability*, and *observability*—the other generic attributes of innovations that theorists have abstracted from empirical instances that first might seem unique and unrelated (Rogers, 1995). That's the power of theory. The specific manifestations of those concepts will differ, but these *types* of problem will always remain a potential threat to innovators' success.

Concepts (or constructs) are "the foundation of theory" (Suddaby, 2010, p. 346) and the "What?" in Whetten's (1989) discussion of theoretical contributions in terms of *What*?, *How*?, *Why*?, *Who*?, *When*?, *and Where*? No building can be strong without a good foundation and neither can a theory. It is therefore a sad reality we share with management research that many of our core concepts are ill-conceived or poorly defined, leading to internal inconsistencies and confused conversations (Bacharach, 1989; McKinley, 2007; Locke, 2012; Suddaby, 2010). This pertains even to very central concepts—try "entrepreneurship," "resources," "dynamic capabilities," "performance," and my very own favorite pet peeve "entrepreneurial opportunity" (Arend, 2006; Arend & Bromiley, 2009; S. Carter, 2011; Hansen et al., 2011; Miller, Washburn, & Glick, 2013; Priem & Butler, 2001; Zahra, Sapienza, & Davidsson, 2006). I can only hope the next generation of scholars in entrepreneurship will take concept clarity more seriously[2]. More about this in Chaps. 4, 6, 7, and 8. Taking Suddaby (2010) to heart is a good start.

As another example of the need for abstraction and understanding, consider the following example. When I first entered the field of small business research, it was with an interest in the growth (and non-growth) of small firms. I soon got to view growth as an instance of (continued) entrepreneurship—a view I have subsequently revised and refined (see Davidsson et al., 2002 and Chap. 1, above). By reading a large number of empirically based studies on entrepreneurship and small firm growth, the picture I got of their determinants was something like the horrors of Fig. 3.1.

Now, how does one deal with this? One approach would be to cover everything and estimate (or conceptually try to tease out) all relationships. This is not feasible; we would soon lose track of the important overriding structure and arrive at the result that "the world is complex." This is something we probably knew from the very beginning. So this is the consequence of lack of theory: even if we measure all relevant variables and estimate all relevant relationships—which is highly unlikely to happen in the first place if we do not have theoretical insights—we will not really understand much. Using the concepts in Fig. 3.1, one would at best arrive at empirical generalization, not theoretical understanding. Assume, for example, that we find a reliable negative relationship between firm age and entrepreneurship. What does that *mean*? The relationship, no matter how strong and statistically significant, is

[2] I leave it to you to decide whether to fix this problem before or after fixing global warming.

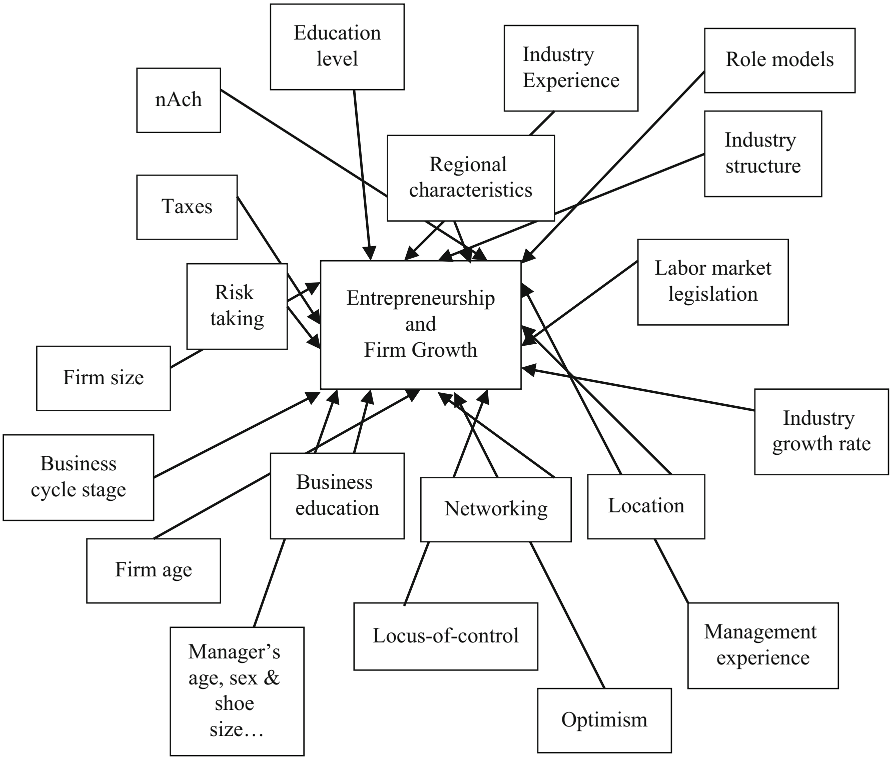


Fig. 3.1 An unsorted array of factors possibly affecting entrepreneurship (and growth)

empty and pretty meaningless without interpretation. What is it about older firms that make them less entrepreneurial? Is this a problem for older firms? If so, what can they do about it? Perhaps older firms are happily non-entrepreneurial, but is this a problem for society? If so, what can policy-makers do about that? A theoretical understanding of the relationship would answer at least some of those questions; the empirical generalization itself just leaves us wondering.

Arguably a better way of attacking the problem of the complexity in Fig. 3.1 is to move up the ladder to a higher level of abstraction. This means including a lot of the specifics but to view them as aspects of more general concepts. A good question to ask oneself in order to raise the level of abstraction is: *This is a special case of what*? (and the answer may well be found outside or "above" entrepreneurship). While reading and rereading a large number of empirical studies when I was working on my doctoral dissertation, I asked myself that question a number of times. After several attempts at summarizing the findings of previous studies in terms of more abstracted concepts, I came up with three rather simple ones. All of the specifics in the "messy" picture could be regarded as aspects of *ability*, *need*, or *opportunity*. More specifically, I let the model depicted in Fig. 3.2 guide the analyses.

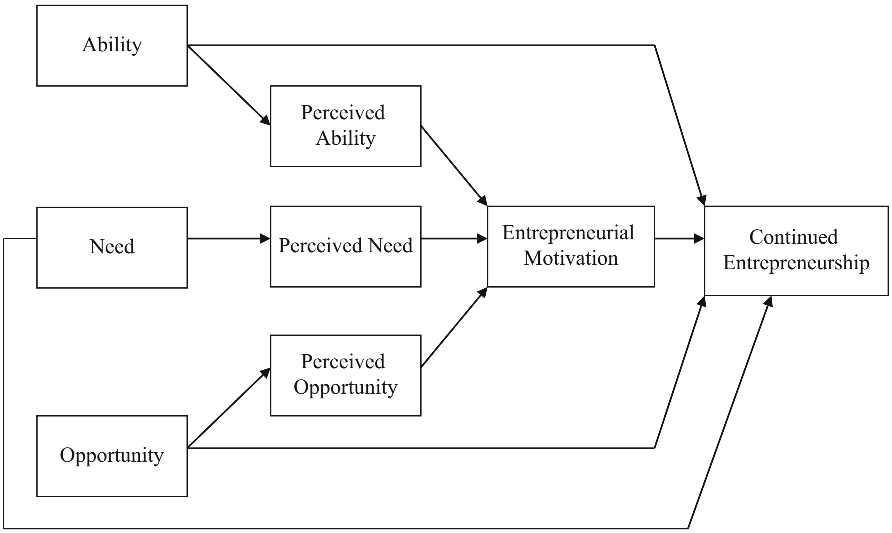


Fig. 3.2 A somewhat more sorted array of factors possibly affecting entrepreneurship

Relationships and their understanding. Relationships are Whetten's (1989) *How*? and the explanation of their existence is his *Why*? Altering or bettering the *Why* is what he sees as "probably the most fruitful, but also the most difficult avenue of theory development" (p. 493). Returning to Fig. 3.2, in addition to relying on a smaller set of more general concepts, the theorizing underlying this graphical model represents an improvement over Fig. 3.1 because it suggests meaningful explanations for empirical observations, some of which had already appeared repeatedly in the literature. Why would older and larger firms grow less? Because they have less *need* to grow! The older and larger you are, the more likely it is that you have already attained the minimum efficient size (firm size and firm age were used as empirical indicators of *need*). Why would affiliation with a growing industry and location in a major city be associated with higher (entrepreneurial) growth? Because there is more *opportunity*[3] in such environments! Do we really believe that education *causes* entrepreneurship? Do we believe that experience does? I would say no, but we may believe that some *ability* is needed for entrepreneurship to come about and that measures of experience and education are two out of many possible indicators of *ability*. Faced with a specific situation, those with more education and experience may regard themselves more able to exploit that situation (ability → perceived ability) and therefore become more motivated to act on it (perceived ability → entrepreneurial motivation). We may also be justified in suggesting that those with

[3] Alert readers may raise an eyebrow at my embracing of the o-word here. Well, for starters, this is 25–30 years ago. Second, I here use "opportunity" as an uncountable (cf. Davidsson, 2003) to denote generally favorable circumstances and not to denote specific, preexisting entities ready to be picked and converted into successful new economic activities. My empirical items were indicators of a favorable resource situation as well as of a munificent environment.

more education and experience are better at making plans materialize (ability → continued entrepreneurship).

Secondly, applying this level of abstraction makes it possible to actually put a model that considers many specific sources of influence to statistical analysis without getting lost. The use of partial least squares analysis—a LISREL-like structural equation modeling technique (Fornell & Larcker, 1981; Vinzi, Chin, Henseler, & Wang, 2010)—allowed me to analyze versions of this model with up to 72 manifest (low-level) variables in the same analysis (Davidsson, 1991). Doing so was certainly no walk in the park back then with the programs and documentation (un) available and would not be easy today, either. However, analyzing the pairwise covariation among such a large number of variables would lead nowhere but to bewilderment. Abstraction is a blessing, and theoretical understanding is what makes research more fun than producing a largely meaningless list of empirical generalizations.

The relationships in a theory may be expressed with varying degrees of precision and sophistication. In the most rudimentary case, the theory would say "A affects B." To increase precision, the theory may suggest the *direction* (sign) of the relationship: "A has a positive effect on B." Most published entrepreneurship research today would theorize the direction of the influence. However, we have not come very far at all in precision as regards the expected *magnitude* of the relationship (Edwards & Berry, 2010), i.e., "A has a positive effect on B of at least magnitude x." Hypotheses concerning relative magnitude are starting to appear more frequently, though, i.e., "The effect of A_1 on B is significantly larger that the effect of A_2 on B" or "the effect of A on B is significantly larger than the effect of A on C" (e.g., Naldi & Davidsson, 2014).

The predicted *form* of the relationship is often unstated and usually tested as linear in empirical analysis, although hypotheses suggesting, e.g., "A has an inverted U-shape effect on B" are becoming more common. I have to admit that my first nonlinear hypothesis ever saw the light of day only recently—and wasn't supported by the data (Senyard et al., 2014). Sadly, the most common instances of nonlinear specification may be those where a variable transformation or the application of a particular analysis technique actually makes the relationship nonlinear without the analyst reflecting on this fact.

Much more has happened with regard to theorizing *contingent* relationships, i.e., that the influence of A on B is contingent on its influence on an intermediate variable W (mediation; see Fig. 3.2) or on its interaction with another explanatory variable *Z* (moderation) or both at the same time (Baron & Kenny, 1986; Edwards & Lambert, 2007; Hayes, 2009). Theorizing contingent relationships has a longer history in management and organizational theories in the form of contingency or configuration meta-theories (see Dess, Lumpkin, & Covin, 1997; Wiklund & Shepherd, 2005, for entrepreneurship applications). However, whether under these specific labels or not, they seem to have gained considerable popularity in entrepreneurship during the last decade. I had reason to check a few years ago, and it turned out that almost every issue of the 2006–2008 volumes of the leading niche journals (*Journal of Business Venturing* and *Entrepreneurship Theory & Practice*) included one or more articles applying some form of analysis of moderated relationships. What remains to be determined is how many of these moderation hypotheses are theoretically and practically meaningful, and replicable, as moderated effects have been shown to be even harder to reproduce (by a wide margin) than are main effects (Aarts et al., 2015).

Boundary conditions. With this concept, I refer to the conditions under which theories should and should not be expected to apply. Whetten (1989) somewhat discounts improved explication of a theory's boundary conditions by suggesting that (as a theoretical contribution) it is insufficient to point out limitations in current conceptions of a theory's range of application. There are several signs that this type of contribution is more appreciated today. The ubiquitous moderation hypotheses are actually an example of this; the magnitude or even sign of a relationship is theorized to vary due to other factors not considered in the original formulation of the theory. This actually amounts to drawing boundary conditions for the "original" or (previously) presumed "universal" theory: under other conditions, the relationships are different.

Not least the experience of conflicting results leading to a frustrating lack of cumulative, "certain" evidence has led to an increased emphasis on context. Johns (2006) provides an excellent overview which can serve as inspiration both for applied theorizing and for research design. In entrepreneurship, several scholars have recently called for more theoretical attention to context (Welter, 2011; Zahra, 2007; Zahra & Wright, 2011) which can also be interpreted as a call for specification of boundary conditions. Boundary conditions can concern dimensions other than spatial (or industrial) context. For example, reviewers of research on (small) firm growth have called into question any attempts to explain firms' amount of *total* growth, instead calling for theory-driven research on specific forms of growth (cf. Davidsson, Achtenhagen, & Naldi, 2010; McKelvie & Wiklund, 2010; Shepherd & Wiklund, 2009; Chandler, McKelvie, & Davidsson, 2009; Lockett, Wiklund, Davidsson, & Girma, 2011; Naldi & Davidsson, 2014).

Above, we have discussed *concepts*, *relationships* and their interpretation/explanation, and *boundary conditions*. In my darker moments, I fear that the boundary conditions for social theories are actually so narrow that we cannot transfer much at all of *relationships* from one context to the next. That is, depending on a myriad of cultural, institutional, and macroeconomic factors as well as historical particularities, it may be the case that the influence of A on B turns out weaker or stronger or even reversed. If this is the case, it means that if we really want to know what's true for a particular place at a particular time, the relationship has to be retested in the particular context where an important decision currently is to be made. *Concepts*—if well conceived and well defined—arguably have much higher generality than relationships have. Concepts allow us to contemplate and discuss what the pertinent issues may be, so that we can design the empirical test that may be necessary if we really want to know the effects in a particular context. If this is realistically the best we can do most of the time in the social sciences, all the more reason to take the development and definitions of concept seriously.

3.4 The Role(s) of Theory in the Research Process

Theories are commonly described as *deductive* or *inductive*. In the extreme form of the former case, the theory is an abstract, an internally consistent fantasy world that the theorist has come up with. The closest real example in the social sciences is perhaps (basic) microeconomics, with its highly stylized firms (essentially a cost function) and strong behavioral assumptions. Once in place, we can test whether the theory has anything to say about the real world by *deducing* from the relationships among concepts in the theory how empirical variables in the real world ought to behave, if the theory is valid. If we fail to falsify the theory, we find it useful and may want to act on its implications. Locke (2007) instead argues for inductive theory, which generalizes from repeated observations of empirical regularities. According to him, it takes many observations over a long time in order to develop useful theory. Arguably, induction should result in more realistic theory. However, if based on observational data, the risk is that the observed reality is too complex for any sufficiently clear and general concepts and relationships to emerge. If instead the theory is to be developed through a series of experiments, well, then at least some rudimentary theory must already have been developed from which one can deduce when the first experiment is designed (at least "A affects B") otherwise there would be no reason to manipulate A and measure B in the experiment.

Thus, deductive theorists must get their inspiration from *some* kind of input (which introduces an element of induction), and inductive theorists must let some a priori ideas guide their attention (which introduces an element of deduction). So real theorizing happens somewhere between these theoretical extremes, and this in-between is sometimes portrayed as a useful strategy and given a name of its own: *abduction*[4] (Alvesson & Sköldberg, 2009). Anyway, there are two major roles for theory in the research process, namely to *guide the design and/or analysis of empirical studies* and to *interpret the results of empirical research or other empirical observations*. Although this distinction has some clear overlap with inductive-deductive, I will now let those latter terms rest for a while. The theory or theories that guide the design should logically also be used for interpretation of that same research. The converse does not necessarily hold true, although it often does. It is conceivable that theory that was not considered at the design stage may still be useful for interpreting and understanding the results. Within those two roles—design and interpretation—there are a variety of cases, some of which will be discussed below.

3.4.1 Theory as Guide to Research Design and Analysis Mark I: The Theory Test

The most obvious case that comes to mind is the pure theory test. This type of research starts either from an interest in the theory and its applicability or from an interest in the entrepreneurship phenomenon to be investigated. Either way, the

[4] I sometimes wonder what supernatural beliefs or sense of humor drives my fellow Swedes when they coin theoretical terms like "abduction" or "psychic distance" (Johanson & Vahlne, 2009).

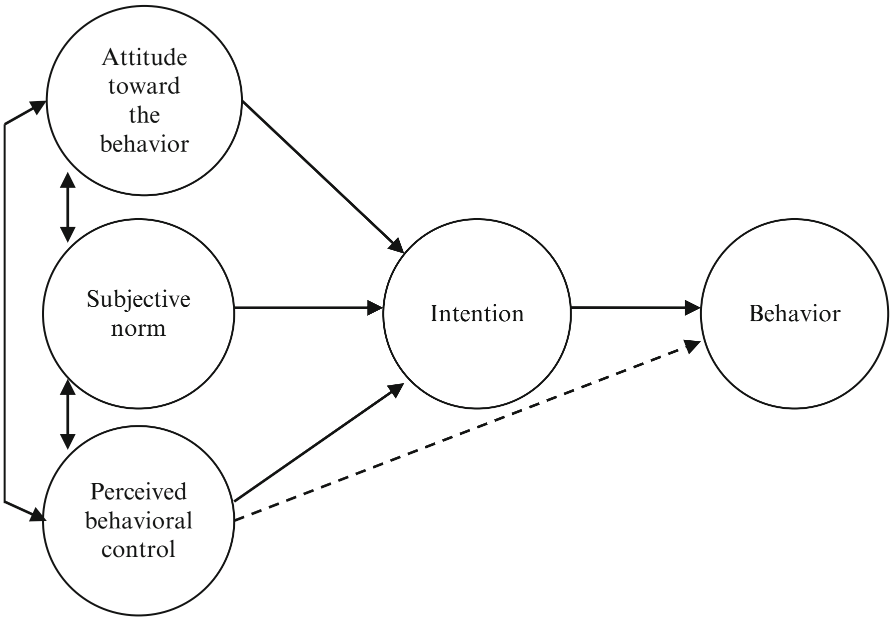


Fig. 3.3 A graphical representation of the theory of planned behavior

researcher realizes that an existing theory has implications for an entrepreneurship phenomenon. The researcher therefore designs an empirical test of the propositions made by—or hypotheses derived from—the theory. A well-specified theory is likely to guide the researcher on issues like:

- What kind of units/level of analysis do I need data on (e.g., individuals or firms or networks; general population or business founders or habitual entrepreneurs)?
- Will cross-sectional data suffice, or do I need more than one wave of data collection?
- What core concepts do I need to operationalize?
- What relationships should be tested and with what analysis technique?

As an example, consider a test of Ajzen's (1991) theory of planned behavior (TPB)—assuming it has not previously been applied in an entrepreneurship context (which is not true; see, e.g., Carr & Sequeira, 2007; Leroy, Manigart, Meuleman, & Collewaert, 2015; Shook & Bratianu, 2010). The core concepts and relationships in this theory are depicted in Fig. 3.3.

Behavior would in the current context mean starting a business venture or taking concrete actions toward realizing a business start-up. Intentions would reflect willingness or a plan to do so. Attitude reflects the extent to which the focal individual regards starting a venture as a good or bad thing to do, whereas subjective norm is his/her assessment of what valued others (e.g., family, friends) think about it.

Perceived behavioral control would here be high for individuals who feel they have the knowledge, contacts, and means needed to get a business going and lower for those who feel they lack one or more of those requirements.

A test of this theory is a piece of research that could be initiated by a researcher specializing in TPB and looking for other arenas for testing the theory's applicability after already having established its relevance for behaviors like weight loss, use of contraceptives, and quitting smoking. Alternatively, we may have to do with a researcher interested specifically in why people go into business for themselves and who in TPB has found a theory that seems promising for (partially) explaining this phenomenon. Either way, the theory would be helpful in answering the above questions. What subjects? Clearly, psychological theory gears one to the individual level of analysis, but not to practicing entrepreneurs in this case. The general population seems to be a prime suspect, possibly deducting those who already run their own businesses (unless you want to include plans to start another venture) and people permanently out of the workforce. Could we use students? Yes, as a second best solution and at the cost of making the theory test much more restricted. Preferably, one would then use MBAs or last year undergraduates rather than freshmen, who are for the moment rather far from making real career choices. Cross-sectional data? You would possibly get away with it for a test of the left part of the model, but not for the important intention-behavior link (and you still wonder why you have seen more papers on the rather unimportant question of predicting intentions than those taking on the important task of explaining actual entrepreneurial behavior?).

What concepts need to be operationalized? This is obvious from the model; just don't think that any measure is a valid measure. Psychological concepts like these are likely to need multiple indicators whose internal consistency needs to be tested in factor and/or reliability analyses (see Chap. 6). What relationships should be investigated? This is also obvious from the model, but don't forget that the non-included arrows should be empirically ruled out and not just assumed not to exist; otherwise, we haven't really tested alternatives to the theory. As to techniques, the co-occurrence of indirectly measured constructs and direct as well as indirect relationships points toward some kind of structural equation modeling (SEM) technique (Goodhue, Lewis, & Thompson, 2012; Marcoulides, Chin, & Saunders, 2009), although if the ultimate dependent variable is dichotomous, their applicability may be restricted. Interactions and nonlinear relationships should also be tested (possibly in separate analyses), as finding those to be nonexistent strengthens the support for the theory. Remember also that if we do not have temporal division between the three sets of variables in the model, we have not ruled out the possibility of reverse causality. That is, if everything is measured at the same time, we cannot rule out that the arrows (partly) run in the other direction.

The design suggested by the theory, then, would be a longitudinal study of a representative sample from the working age population, using validated operationalizations of the core concepts and a set of analyses, performed with adequate techniques, of the relationships predicted by the theory as well as other possible alternatives. The analysis should also control for other factors (age, sex, education, etc.) in order to rule out that we have incorrectly ascribed their influence to our theoretical variables. If supportive, a well-designed test of this kind would lead us to conclude that TPB is valid for entrepreneurial behavior. From the perspective of entrepreneurship research, a supportive test would make our beliefs about why and how business start-ups come to be somewhat less speculative. If the outcome were negative, we would conclude that the explanations offered by TPB do not apply to (this type of) entrepreneurial behavior.

This would constitute valuable contributions. However, there are some limitations as well. With respect to understanding the phenomenon, the problem with a single-theory test is that one might only obtain a very partial and fragmented understanding. Firstly, what about entrepreneurial behavior that is not particularly planned, especially not far in advance (Bhave, 1994; Sarasvathy, 2008)? And what about spousal (Ruef et al., 2003) and other team start-ups (Steffens, Terjesen, & Davidsson, 2012)—how much of that dynamic can be captured in this research? Further, like many psychological theories, the explanations are rather proximal psychological constructs. What more fundamental variables (e.g., those that policy-makers can affect) influence relevant attitudes, norms, and perceptions of behavioral control?

From a more (career-)tactical point of view, Whetten's (1989) criteria for theoretical contributions and Colquitt and Zapata-Phelan's (2007) findings concerning what actually gets published in top journals give reasons for concern. Although valuable, a straight theory test might get turned down on the basis of not making a theoretical contribution. If so, you might need to add something, to offer a theoretical extension. My dear friend, the phenomenal Dean Shepherd has a favorite trick, namely to identify a gap in a theory or literature and then help fill that gap not in some haphazard way, but by using some *other* theory or literature to fill the gap. In the case of TPB, this could concern moderators of the intention → behavior or perceived behavioral control → intention relationships. That is, can we find and test a theoretical explanation, particularly relevant in the entrepreneurship domain, why some people do and others don't act on their capacity or realize their entrepreneurial plans? Such a contribution, i.e., one that actually follows the cases and relates intentions to behaviors, and which also can explain theoretically why that relationship is sometimes strong and sometimes weak, would probably have a decent chance to get published in a really good outlet. In the first edition of this book, I suggested a *triggering event* (such as unemployment, divorce, milestone birthday, inheritance, or other windfall gain) as moderator between intention and entrepreneurial potential on the one hand and behavior on the other, referring to Shapero and Sokol (1982). Today, you would probably have to find another reference and have to pretend we learn more if you make the same argument by referring to some *big name* in a *big journal* who said something about half as insightful under the rubric of *theory*.

It may seem self-evident that one should let theory guide the design. Sometimes, it is not as easy as it might first seem, however, to find a satisfactory solution to this matching need. One common mismatch in entrepreneurship is the use of a more Schumpeter-like conceptualization of entrepreneurship and its effects on the economy (Schumpeter, 1934) and then use "self-employment" as the dependent variable—simply because no measure of what the theory actually speaks about is available (S. Carter, 2011; Sanandaji & Leeson, 2013). The "little" issue with this, of course, is that most self-employed never provide significant innovations and never expand their operations. Following my argument in Chap. 1, I acknowledge they fulfill the entrepreneurial function to some degree when they first enter the market, but in most cases, the degree of entrepreneurship is very limited.

The mismatch may be much subtler. Because there are different conceptualizations of "the firm," there may be a mismatch between the theoretical and empirical notion of "firm" even though the *level* of analysis seems to be right (Davidsson & Wiklund, 2000). We will elaborate on this issue in Chap. 5. It may also be the case that no data can be obtained that perfectly matches the theoretical level of analysis. For example, ecological and evolutionary theories deal with "species" and "populations" of organizations that share certain competence and/or other characteristics. Empirically, however, membership of a population is often equated with having the same industry classification code, and "populations" then become equal to what industrial economists call "industries" (Aldrich, 1999, p. 224). Obtaining samples that more closely mirror the theoretical definition of organizational "species" would require costly and cumbersome collection of primary data (Gratzer, 1999). There are interesting role models in which to seek inspiration, though (e.g., Usher & Evans, 1996).

Sometimes, data do not exist for the chosen level of analysis but can be aggregated from lower levels. For example, in *Culture and Entrepreneurship* (Davidsson, 1995a; Davidsson & Wiklund, 1997), I needed data on cultural (values and beliefs) differences across regions. As no regional data on such variables were available, I had to collect them from representative samples of individuals in different regions. The averages for individuals in regions were then used as regional level scores (cf. Obschonka et al., 2015; Stuetzer et al., 2015). As a bonus, I got data that were excellent also for analyses on the individual level (Davidsson, 1995b).

A special case of the theory-testing approach is when theory is not used to design the data collection, but only to guide the analysis of data that are already available. I will give some examples of this in the "Replicating Others" section of Chap. 9.

3.4.2 Theory as Guide to Research Design Mark II: Understanding the Phenomenon Through an Eclectic Framework

A main problem with mindless application of the theory-testing formula in a maturing field is that it may gear research toward filling lesser and lesser gaps concerning increasingly uninteresting or unimportant details without really contributing much to understanding the main forces that give shape to the phenomena at hand. When research starts from a *true interest in the phenomenon*, one will soon detect that a single theory can only offer very limited and partial insights into phenomena that are truly complex. Theory testing as described above deserves the position as sovereign ruler only in situations when all the theories are part of the same puzzle, so that the insights gained from one theory test fit neatly alongside previous knowledge development. I hold that these conditions are currently not met in entrepreneurship research. Therefore, if sinning means deviating from the ideal of straight theory testing (even with minor extensions), then one should allow oneself a little sinful escapade at times. Actually, I think we put it quite nicely in the introduction to a recent book (Davidsson & Wiklund, 2013, pp. 2–3), so why not reuse that piece of prose to underline this point? The work commented on is Wiklund, Patzelt, and Shepherd (2009).

> It is decidedly phenomenon-driven but not atheoretical; instead it takes on the task of integrating several theoretical perspectives. It uses a large sample representing a select set of different industries and size classes, thus allowing for broad generalisation while avoiding some of the heterogeneity and micro-firm dominance of simple random samples. It employs a longitudinal design, thus separating the actual growth in time from its theorised antecedents. It lets both theory and data speak by allowing revision of the theoretical model. It assesses growth with multiple indicators rather than a single measure that may be differentially valid across a broad sample. Finally, it explains a sizeable share of the variance of "total" growth—not just a statistically significant but possibly practically irrelevant share of a variance that through the researchers' design choices has already been made much smaller than that occurring in the real economy. We agree that the type of carefully designed single-theory research on homogeneous samples of firms that is the current fashion in leading management journals can reach further and deeper as regards both measurement and estimation of the effects of the factors highlighted by that particular theory. However, we hold that it benefits the collective sobriety of academic research on firm growth to once in a while get an idea of where these factors sit in the bigger scheme of things.

There are some variations to the theme of "understanding the phenomenon through an eclectic framework." One, which is the closest to the theory-testing approach just described, is to design the study to test more than one theory in parallel (but separately) in order to be able to determine which of them best explains the phenomenon (cf. Combs & Ketchen, 2003; Krueger, Reilly, & Carsrud, 2000). That is not very sinful at all and can be highly useful. The more representative case is when the researcher shows a little more creative disrespect for the original theories. Theories are tools, remember? Your wish to modify the tool and combine it with other tools does not imply a shortcoming on the part of the toolmaker—she/he could not possibly know about your specific needs.

Some examples of applying an eclectic framework by building on elements from several theories come close to theory testing with theoretical extension, as discussed above. The distinction depends on how many different theories are brought in and how the research is presented. If presented as one main theory with one or two additions from other origins—preferably articulated theories rather than "just" empirical observations—to plug clearly identified gaps, the research is usually well received in the current research culture (you should realize by now that I'm not 100 % sympathetic to this culture). In the context of TPB, to make the model more complete in terms of providing understanding of the phenomenon, such extensions could be to suggest additional (theory-backed) predictors associated with a higher probability of exerting entrepreneurial behavior, whether planned or not. This could be, for example, sex, the presence of close role models, or prior entrepreneurial experience. It could also be the addition of more tangible (theory-backed) antecedents of the psychological variables at the rear end of the model.

The model in Fig. 3.2 (above), which I developed in my dissertation study, can be regarded as a purer example of an eclectic framework approach, as it builds on inputs originating from a range of theoretical and empirical sources, which were processed through my own sensemaking. I have argued that such an approach can lead to a more complete understanding of the studied phenomenon. However, every framework has to be incomplete—we cannot include the whole world in our models. And—as every experienced empirical researcher knows—even with the most comprehensive model, we are unlikely to explain more than half of the variance. There is just too much idiosyncratic variation and unavoidable measurement error, so it is actually sound practice to be suspicious about research that purports to reach farther than that. Moreover, using a comprehensive, eclectic framework for designing one's study makes for a more demanding research task. Therefore, if you are an American doctoral student and approach your committee with a research proposal of this kind, you can be sure that the question "Will they reject it?" belongs in the "Is the Pope Catholic?" category. I was as blessed by European, 1980s freedom when conducting my thesis work as have many others been cursed by the exact same freedom to take on research tasks of an enormous magnitude (Davidsson, 2014). Generally speaking, "narrow down!" *is* good advice for doctoral candidates.

In more general terms, there are two main drawbacks of the eclectic framework approach. First, compiling elements from different sources may lead to a theoretical mishmash where arguments pertaining to different levels of analysis are mixed, where consideration of time and process varies across different parts of the model, where there is conceptual overlap between constructs, and where logical inconsistencies abound. Second, to achieve evidence-based understanding of the phenomenon in its full complexity may truly be a *mission impossible*. It may well be beyond our cognitive abilities and the capacity of our methods to correctly model and interpret what is going on. With some effort, we may be able to map out the terrain insofar as describing the empirical world in terms of useful concepts, but to also get the relationships right in a complex world may be beyond us and our tools in a very fundamental way. We then have to simplify to understand anything at all, a topic we will dig deeper into in the following chapters.

This said, I maintain that more comprehensive research built on eclectic theoretical frameworks is also needed. Entrepreneurship *is* a complex phenomenon, and that complexity does not go away just because we make our designs more manageable. Piling up the many small pieces from separate theory tests is important. However, that alone will not lay the puzzle for us without more comprehensive studies that help us see what pieces do and do not belong in our puzzle and which makes the fitting pieces fall into their proper places.

3.4.3 Theory as Tool for Interpretation: The Theory Test

I have pointed out already that the theory that is used for designing the study should also be used for interpreting its results. Does this really need to be pointed out? I think it does, not because researchers frequently make a sudden turn to other theories when they interpret their results, but because there is this kind of really dull, quantitative research where there is *no* interpretation at all. Early entrepreneurship research was full of this type of studies. That is, the concluding section of the paper merely restated that these hypotheses were supported whereas those were rejected. Period. Yawn. For goodness' sake—what does it *mean*? Our vantage point was our curiosity about an interesting theory or a really important societal phenomenon, right? Then maybe we should return to discussing this in the tail end of the manuscript? Seems a good idea. In my opinion, it is often the quality of the concluding discussion that distinguishes an excellent and highly interesting piece of research from merely publishable, standard research (in combination with an excellent introduction, because without that the excellent discussion might not get read at all). So please don't run out of steam. Finish the job and devote considerable time and energy to telling us what the results (might) imply for practice as well as for future research and theory development.

As regards further theory development, the interpretation of a pure theory test is pretty straightforward. The possible outcomes can be sketched as four cases. The first is that the theory holds up, all its proposed relationships are found in the data, and the theory provides strong explanation of the phenomenon under study. This shows that the theory is applicable to the investigated domain and strengthens the general validity of the theory. The second case is limited support. That is, the hypothesized relationships hold up, but they are weak and much of the variance remains unexplained. This is likely to be the case in our TPB example. One interpretation would be that the theory is valid (also) in this domain, but that we need to add more variables and relationships to it in order to get a more complete explanation of the studied phenomenon. Alternatively, an even more useful route might be to see if the support is stronger and weaker for different subgroups in the sample. This may lead to a better specification of the theory's boundary conditions; a theoretical contribution of the *Who*? *When*? *Where*? kind in Whetten's (1989) terminology. The third case is partial support, i.e., that some of the relationships proposed by the theory hold up empirically whereas others do not. This points to a need to modify the theory rather than—or in addition to—supplementing it. This could reflect a shortcoming of the theory that has also emerged in other empirical contexts. If so, the theory would likely benefit from a general revision, whereas a domain-specific deviation from the theory's predictions only calls for domain-specific modifications. The fourth case is straightforward. This is when the results of the empirical test do not at all support the theory. Assuming that the empirical test is performed through a well-designed study, we would then reject the theory, at least with respect to the specific phenomenon under study, and start looking for other alternatives. Importantly, rejection of a theory is not a failure. From a Popperian perspective, falsification is the main route to knowledge development (Popper, 1992). I don't completely buy that and we probably shouldn't (Klayman & Ha, 1987; Locke, 2007). However, given the ubiquity of confirmation bias—the tendency to only look for supportive information (Davidsson & Wahlund, 1992)—researchers would be well served by showing some more appreciation of the informational value of nonsupport.

With respect to understanding the phenomenon and the findings' implications for practice, the interpretation work is firstly a matter of going back to the original theory and the detailed arguments as to *why* the proposed—and now confirmed—relationships should manifest themselves. If we forget this important step, we are back to relatively empty empirical generalizations. The theoretical mechanisms are what give direct hints at what practitioners can and cannot do in order to increase or decrease the dependent variable in question.

3.4.4 Theory as Tool for Interpretation: The Eclectic Framework Approach

Most of what has been said above is equally applicable to the eclectic framework situation. This is especially true if the various theoretical elements are combined a priori to a well-specified model or framework with precise predictions as to the direction, sign, and form of relationships among the included constructs. When this is the case, the only difference to the theory test as described above is that the results should perhaps be regarded as more tentative, as we are now dealing with a new theory rather than one that has been proven valid before in other domains.

In most real cases, the eclectic framework is not that well specified before data collection and/or analysis starts. For sure, the sampling and data collection have been guided by theories, but rather than making our minds up as regards TPB or some other theory, we may have included items that could serve as operationalizations of either, and we may have included measures of concepts from various theories without working out beforehand how we expect these to relate to each another. We may also have included in our eclectic framework variables that eventually do not show up in our reported results, because on closer thought they were not logically compatible with other concepts in the framework or they turned out to be unimportant.

The framework in Fig. 3.2 is an example of this. When the point of departure is a genuine interest in the phenomenon and the current state of theory-based knowledge about this phenomenon is short of what it could be, I see nothing wrong in such an approach. Good research is often a matter of abductive wrestling between theory and data. However, the more inductive the process has been, the more tentative are the results and the greater reason there is to give room for caution and alternative explanations in the interpretation. Importantly, when the research process has been a matter of theory-data wrestling, it should be portrayed as such. Regrettably, it seems likely that it is not uncommon that what in published research is portrayed as hypotheses were little more than hazy ideas before the analysis work began. Presenting exploration as theory testing is not good scholarly practice, no matter how firmly institutionalized the practice is. We should not fool ourselves—and others. The proper names for such practices are deceit, cheating, and fraud. I think those who believe that packaging (partly) exploratory research as purely deductive is what it takes to get published in a good, scholarly journal should stop believing so. That is, stop believing that such a journal truly is a good and scholarly one.

3.4.5 Theory as Tool for Interpretation: Post Hoc Theorizing

We have now approached the exploratory situation, when theory that was not used for design is used or created in the interpretation phase. This does not mean that *no* theory was used for designing the study. In line with my earlier reasoning, theory-free data collection is hardly possible. What we really mean is that the theory guiding the research design was vague, unarticulated, rudimentary—or just different—relative to the theory used or created for interpretation.

I think there is reason to be very, very wary about anything looking like truth claims from exploratory research. This skeptical attitude is in part formed by explicit persuasion during my research training, e.g., making me read Armstrong (1970) and his compelling example of how prone we are to find post hoc rationales for any empirical relationships—even when mistakenly based on an analysis of random numbers. More important, however, is my training in cognitive psychology. Learning about the selective nature of perception, attention and information search, the constructive nature of memory, and how easy it is to manipulate our perception, recollection and sensemaking of events has a humbling effect on a researcher's belief in his/her ability to distill any form of generalizable "truth" (or intersubjectively meaningful knowledge) from unsorted and complex data (Anderson, 1990; Goldman, 1986). There is a very real risk that the theories that actually guide the analysis are the prejudice and preconceptions of which I am not consciously aware.

However, there are many situations in which the post hoc use or development of theory can be justified. The first is when we get support for the theory that was used for designing the study. In trying to understand the phenomenon, and teasing out the implications for practice, we may also want to use *other* theories for our interpretation. This may seem strange, but is in fact not strange at all. To illustrate this, let us return to the TPB example. We noted above that in order to get a more complete picture, we might want to expand the design to include elements from other theories. However, even if we did not do that in the design phase, we can add understanding in the interpretation phase by doing so.

For example, a possible outcome of a test of TPB in an entrepreneurship context is that *perceived behavioral control* comes out as the strongest predictor of entrepreneurial intention and that this variable also has a strong direct influence on behavior. We now want to know where these perceptions of behavioral control come from. Assume that TPB—like its graphical representation in Fig. 3.3—does not provide sufficient explanation of this. There is then no reason for us to refrain from looking around and asking, "Does some other theory, or established empirical generalization, give us a clue?" Such a search would likely lead us to *social learning theory* (Bandura, 1982, 1986) and to noticing that the concept of *self-efficacy* has a large overlap with Ajzen's concept of perceived behavioral control (Eagly & Chaiken, 1993). Bandura (1982) elaborates on the sources of self-efficacy and holds that individuals develop and strengthen such beliefs in four ways: (1) mastery experience, (2) modeling or observational learning, (3) social persuasion, and (4) judgment of their own physiological state. Thus, by adding elements from this theory to our interpretation, we can reach much farther in our understanding of the results and their practical implications. We may also have data available to run exploratory analyses, which may confirm that different indicators of direct and indirect experience are positively associated with perceived behavioral control. So far, so good; this can be material for a theoretical extension of TPB. Regrettably, many scholars in this situation—sometimes encouraged by reviewers and editors—would present this exploration as if it were part of the a priori theory. This is fraudulent or at least deceptive, because given the results you can very often find *some* theory that would have predicted them.

The second case also deals with the situation when we have received support for our tested theory. Even without a need for additional theory-based interpretation, it is good scholarly practice in this situation to admit that even though our theory was supported, there may be alternative explanations for the results. I cannot see any reason why theories not used in the design should be banned from the discussion of alternative explanations. Given the problem with post hoc theorizing being presented as theory testing, we should actually demand more of this: do the results also accord fully with other theories, or do they lend *stronger* support to the focal theory than to any other explanation?

The third situation is when we get an unexpected result that runs counter to our hypothesis or when a control variable[5] turns out to have an effect of unexpected form or strength. Yes, for sure, we can stop at just noting that the hypothesis was not supported or noting the effect. However, I see little reason why we should not share our after-the-fact speculations about why those unexpected effects turn up. In particular, I see little reason why it would be worse to base such speculation on previously unused theory rather than armchair reasoning or reference to possible method artifacts. The history of science is full of cases where surprises and mistakes have led to breakthroughs. It is probably even fuller of silent cases of chances foregone, where researchers have simply discarded their unexpected results in disappointment instead of trying to understand what they mean. This said, one should clearly distinguish such tentative, exploratory findings from support for the theory that was used for study design. Again, this is because we can always find *some* explanation after the fact (even for correlations among random numbers, remember?). Further, although one should not refrain from speculation about why the unexpected results occurred, this speculation may not always be best placed in the discussion section of an article with other front-end theory. Instead, it can serve as inspiration for the next study, where the effect can be studied through a proper theory test.

The fourth case is when it is not the researchers' ignorance or creative itches that make them want to go explorative, but a real lack of (obviously) relevant theory, when the knowledge of a phenomenon is at a nascent stage (Edmondson & McManus, 2007). Exploratory empirical work may be justified, and it may only be the combination of such data and the researcher's creative ability that reveals that the studied phenomenon can be seen as a special case of something seemingly unrelated and that using theory originally developed in the context of that other phenomenon may be illuminating. I see no problem with that.

[5] A control variable is an explanatory variable that is included not because we have a theoretical interest in its effect, but because omitting it may lead to incorrect estimation of the effects of those variables we do have a theoretical interest in (Kish, 1987).

3.4.6 Is It the Theory or the Data That Is Supported or Should Be Rejected?

Thoughtful researchers are damned, aren't they? I mean, shouldn't it suffice that we get statistically significant support for our theory and say, "Good; theory proven true!" or fail to get this heavenly authorization and therefore have to say, "Tough luck—theory proven wrong"? No, I don't think that suffices. We have noted already that a given set of empirical results may be consistent with several theoretical interpretations. In addition, an inescapable problem inherent in any theory test is that the outcome we get can either be ascribed to qualities of the theory or to qualities of the data (or method). We don't really know (Locke, 2007, 2012).

For example, in the TPB example, we may get support for the attitude → intention relationship. What is this? Probably a correlation between two paper and pencil behaviors, conducted only a few minutes apart. Is this evidence that if we can affect people's attitudes to entrepreneurial behavior, they will as a consequence develop more entrepreneurial intentions? This may be the reason why we obtained the result, but it could also be due to a personality-, mood-, or response style-based method artifact (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Conversely, we may fail to get support for this relationship. Does this show that the theory is wrong? Possibly, but it could also be the case that poor operationalizations led to such grave measurement error that the true relationship does not emerge from the data.

We have to face it: we don't know for sure. Perhaps the theory should be rejected; perhaps we should instead conclude our data are not up to scratch. Perhaps we are justified in strengthening our trust in a theory, but then again we may have been misled by some peculiarity of the method. We don't *know*—and probably never will.

This means that such a horrible, subjective thing as *judgment* must have a big role in the research process. It also means that this even more horrible thing called *rhetoric* will have a profound role in knowledge dissemination. As consolation, I offer this: perfect democracy and perfect justice are not possible to achieve. This does not mean they are not worth striving for. Similarly, there is nothing wrong with judgment when it is good, and this is easier to achieve when the judgment is based on good evidence, i.e., clear support or rejection of the theory based on good data. And rhetoric is not "just" rhetoric—its likelihood of success is related not only to the form but also to the quality of its contents. But the allure of rhetoric again points to the need for *replication*—an issue we will return to in Chap. 9. When results are replicated in several studies using similar but slightly different samples and operationalizations, our belief or disbelief in theories will be much less contingent on "mere" skillful rhetoric.

3.5 Do We Need Specific Entrepreneurship Theory?

In Chap. 2, I offered the following domain delineation for entrepreneurship research:

> Starting from assumptions of uncertainty, heterogeneity, and disequilibrium, the domain of entrepreneurship research encompasses the study of processes of (real or induced, and completed as well as terminated) emergence of new economic ventures, across organizational

> contexts. This entails the study of new venture ideas and their contextual fit; of actors and their behaviors in the interrelated processes of discovery and exploitation of such ideas, and of how the characteristics of ideas, actors and behaviors link to antecedents and outcomes on different levels of analysis.

If we examine this statement, we have to conclude that there are few contingencies of interest to entrepreneurship scholars that could not be the topic of theory in at least some discipline in the social sciences (Ács & Audretsch, 2010). Not making full use of the tools available within the disciplines would be a wasteful practice. It is not so easy, however, that all the theory entrepreneurship researchers need already exists in the disciplines or in other branches of management and organizational studies. Even if it is true that there are few contingencies of interest to entrepreneurship scholars that are not the topic of theory in at least some discipline in the social sciences, it is equally true that theorizing about entrepreneurship is not the *main* responsibility of any discipline. I have stressed already that many of "our" questions may be peripheral to every discipline. Therefore, although the disciplines have developed many sophisticated tools, these tools may not always be adequate for the task at hand (cf. Davidsson & Wiklund, 2000; Sarasvathy & Venkataraman, 2011). In relation to the above domain delineation, some of the questions one should ask before applying existing theory "as is" are the following:

1. Does the theory acknowledge uncertainty, heterogeneity, and disequilibrium?
2. Can it be applied to the problem of emergence, or does it presuppose the existence of markets, products, organizations, or resource bundles in a way that clashes with the research questions?
3. Does the theory allow a process perspective?
4. Does it apply to the preferred level of analysis (e.g., "venture idea" or "emerging venture" rather than "firm" or "individual")?
5. Is it compatible with an interest in the types of outcomes that are most relevant from an entrepreneurship point of view (cf. Chap. 7)?

Theories exist, and whenever possible, entrepreneurship research should deductively test theory from other fields. However, as a scrutiny of some existing theories in relation to the five questions above would show, they are not always optimal for research questions addressing the processes and analysis levels of most relevance to entrepreneurship research. For example, when Sorenson and Stuart (2008) propagate the promise of a disciplinary approach, their examples can be read as support for just the opposite stance. This is because they tend to focus on concepts that can explain why emergence is difficult or impossible (uncertainty, legitimacy, familiarity) or processes that come after initial emergence (diffusion) rather than theoretical notions that can explain how emergence is possible against the odds of social inertia being stacked against it. As a reviewer-editor-supervisor, I have just recently found reason to reflect on the suitability of conceptualizations or at least established operationalizations of "dynamic capabilities," "ambidexterity," and even "environmental dynamism."

Arend (2014) recently bemoaned what he sees as the continued atheoretical state of entrepreneurship research. His solution is not to borrow more but rather to develop theory in our domain, starting from radically different assumptions than those used in most extant theories. His examples include process vs. variance, disequilibrium vs. equilibrium (see above), dynamic vs. static, complex vs. parsimonious, nonlinear vs. linear, ambiguous vs. risky, and multilevel vs. single. These are thoughts worth contemplating when adapting theories as well as asking new questions through inductive, theory-building approaches. However, I for one do not believe theory is the answer to everything. A risk with an overly strong focus on theory in combination with a tendency to *only* use extant theory (which is understandable; it is much easier to use theory than to develop it) is that researchers may avoid many questions which are at the core of the phenomenon (Wiklund et al., 2011) simply because extant theories aren't there to help us. But it is work on these questions that is needed the most!

One additional reason why entrepreneurship research needs to build its own theory is that the various theoretical fragments developed within the disciplines (of direct relevance to entrepreneurial phenomena) are not likely to form a coherent whole. Again, this failure to collectively cover the entrepreneurship agenda is neither a problem nor a shortcoming on the part of the disciplines. They never had that goal or obligation. However, this state of affairs is a problem and a shortcoming from the perspective of entrepreneurship research, and our field therefore has additional needs to develop its own theories. The emerging theories of effectuation and entrepreneurial bricolage (Baker & Nelson, 2005; Sarasvathy, 2008) are interesting developments in this direction, but we have a long way to go before entrepreneurship has a body of theoretical tools that have been developed within our field.

3.6 A Defense Speech by a Proud Nonbeliever[6]

My introductory confession was written at a time when entrepreneurship research was still under theorized. In my estimation, it was not until around year 2000 that entrepreneurship research started in earnest to have a section explicitly called "theory" or something to that effect, describe its variables as operationalizations of theoretical constructs, present hypotheses (in quantitative work), and make its way into leading journals in various disciplines as well as in mainstream management. Today, we may instead have a problem with overly strong and universal emphasis on theory. My scholarship has always had an element of rebellion—possibly my interest in entrepreneurship even started as psychological reactance to the fact that I grew up in a steel town totally dominated by one, large, and more than 100 years old employer—so in the current climate of a rather *singular* focus on "theoretical

[6] If you are an early-stage research student, you may find this section hard to follow before having read at least the remainder of this chapter and the next one. Being immersed in a research-oriented academic environment for a little longer will also help comprehension, so hopefully the section makes more sense next time you read it.

contributions" in top-tier publishing, I have already vented some skepticism against the sovereign rule of theory. Accordingly, I think there is reason to lend some space to the case *against* exaggerating the importance of always providing theoretical contributions. As I say elsewhere in this book, I do not subscribe to any religious-like *faith* when it comes to making progress in knowledge production. As scholars, we cannot capitulate and let "scripture" or "priesthood"—including this book and its author—dictate what we believe in or how we should conduct and report our research. Although we would be idiots not to listen to our elders, we need to always have our critical minds switched on, and that is not compatible with religious-like devotion to or belief in anything.

One main problem with the idea that *every* paper worthy of publication in the top tier should make a "theoretical contribution" is that it overemphasizes one step in the knowledge creation process—formulating new theoretical ideas—at the expense of other, equally important steps. These include providing *interesting and seemingly important empirical observations*—be they correlations or intriguing observations from case studies—that our current theories cannot explain and which may even call established theories in question[7]. To demand that those who discover them also provide a theoretical explanation is actually *not* taking theory very seriously, because developing good theory takes time (Locke, 2007). And to demand that they do so in the form of front-end theory leading to hypotheses *when this was in fact not the way it happened* is an invitation to academic dishonesty. If leading journals are really convinced that what they publish should be produced with deduction and theory *drivenness*, they should perhaps demand that authors submit their theory, hypotheses, and data collection plans *before* they collect the data and accept papers on that basis (on condition that the data be of the promised quality), *not* on results confirming the theory. Further, they should publish these a priori hypotheses online, and then wait for the data to come in. This would counteract two unsound practices: (a) researchers crafting hypotheses (and choosing theories) based on already known results and (b) reviewers/editors accepting papers not on the basis on how well they are designed or argued, but based on whether or not enough support is found for the proposed theoretical relationships. For some reason, I have never seen anything like this—publication of hypotheses before data collection—realized, or even suggested in our field[8].

The other important knowledge creation step that I keep arguing is suppressed is *replication* (cf. Chap. 9). Given how complex social reality is and considering the fact that most reported findings are not stunningly strong, it is exceedingly naïve to believe that we know much about the truth of a theoretical explanation after one test

---

[7] You believe providing descriptive statistics is about as low as you can get on the scale of scholarly contributions? You think such descriptions are necessarily dull, unimportant, and unlikely to have any impact in the absence of theory? Maybe think again: www.youtube.com/watch?v=ezVk1ahRF78; www.youtube.com/watch?v=WU0kYxhzQvo; www.youtube.com/watch?v=hVimVzgtD6w. Don't forget to compare the number of views with your favorite theorist's citation stats.

[8] It does exist elsewhere: http://editorsupdate.elsevier.com/short-communications/journal-cortex-launches-registered-reports/. Good on them!

in one top-tier journal. Other disciplines, who seem to take their own knowledge creation endeavor more seriously, do not act in this way—and we should be damned happy that medical doctors don't start to experiment on us with new treatments on the basis of as weak evidence as most of our theories have. Hambrick (2007) expresses these problems so eloquently that I should not tire you with more of my own prose:

> Many nice things can be said about theory. Theories help us organize our thoughts, generate coherent explanations, and improve our predictions. In short, theories help us achieve understanding. But theories are not ends in themselves and members of the academic field of management should keep in mind that a blanket insistence on theory, or the requirement of an articulation of theory in everything we write, actually retards our ability to achieve our end: understanding. Our field's theory fetish, for instance, prevents the reporting of rich detail about interesting phenomena for which no theory yet exists. And it bans the reporting of facts—no matter how important—that lack explanation, but that, once reported, might stimulate the search for an explanation (…) I am not aware of any other field in which theory is viewed with such religious fervor. (p. 1346) (…) All other academic fields I am aware of—especially those that have professional constituencies that rely on a formal body of knowledge—attach significant value to straightforward tests of previously proposed theories, ideas, and operating mechanisms. We in management, however, are so riveted on new and revised theories, and so dismissive of simple generation of facts and evidence, that our revealed ethos is that we care much more about what's fresh and novel than about what's right. (p. 1350)

Well said. As regards replication, academic journals should, of course, be allowed to have whatever editorial policies they like (that adhere to accepted scholarly principles according to some academic "tribe"). However, when journals that have published the original theoretical ideas (and perhaps a first test of them) explicitly refuse to assume responsibility for publishing replications which can confirm or refute those ideas, it really makes us look like a field that doesn't take our science very seriously. I can think of a few reasons why we have ended up in this situation:

1. Influential people in our field really do "care much more about what's fresh and novel than about what's right." I don't really believe that.
2. Influential people in our field believe a single study can provide stronger and more conclusive evidence than it really can. I believe that to some extent, but experienced empiricists should definitely know better.
3. Influential people in our field mix up two things: (a) studies that are not anchored in previous research and therefore approach a research question as if it were done for the first time (and which also possibly apply a home brew of new operationalizations) and (b) serious efforts at direct replication, with or without specified extension and/or testing in a different context representing some theoretically significant deviation from the original study. I suspect the aversion against replication is historically grounded in a disappointment with studies of type (a) (of which we used to see a great deal) and not a thought through devaluation of (b) of which we for obvious reasons do not see much at all (cf. Chap. 9).
4. Influential people in our field (i.e., editors) are concerned that replication studies will not attract as many citations as studies presenting novel theoretical ideas, thus hurting the journal's holy "impact factor." Well, we could also argue that it may not be possible to win the Tour de France without doping, but beating others

with the help of forbidden, performance-enhancing drugs remains legally and morally wrong all the same, *n'est-ce pas*? Integrity should rule. If giving room for replications is the right thing to do, then we should give room to replications, no matter how mad the world is. Besides, I'm not sure this "citation theory" is correct. Studies retesting theoretical ideas that have taken off—and these are the most important to replicate—should get quite a few citations from their association with the original paper.

But there is hope. Entrepreneurship research is published outside of management, and there you find highly ranked journals with a more balanced appreciation for different types of intellectual contributions. In their analysis of articles published in the *Academy of Management Journal* (AMJ) over time, Colquitt and Zapata-Phelan (2007) find reason for optimism as they interpret their data as saying that *both* theory building and theory testing have increased over time (mainly at the cost of mere '"reporting" of atheoretical, empirical observations). Looking at the same data, what I see is that theory testing has leveled off (at best) since year 2000 (p. 1290) and that pure testing of previously proposed theoretical ideas peaked in the early 1990s with 40 % of the manuscripts and had shrunk to something like a tenth of that proportion by 2007—today that category is probably nonexistent. Anyway, there is hope for AMJ; see for example the very promising editorial policy change recently announced by Gerry George (2014)[9]. Some influential colleagues actually do care about the totality of knowledge creation in management and entrepreneurship, and the launch of the new *Academy of Management Discoveries* and *Journal of Business Venturing Insights* is another sign of this. These journals are designed to help remedy the current imbalance by making room for unexplained discoveries (front end) and verification or disproof through replication (rear end). If they can achieve sufficient prestige, this may alleviate the pressure to try to publish important, exploratory work under false pretense of theory drivenness and give replications their proper place in our building of scholarly evidence.

There are additional reasons to be wary about excessive emphasis on theory and theoretical contribution. One we have already touched upon is the fantasy that we can develop useful theory quickly and in the context of a single empirical study. Developing useful theory takes time (Locke, 2007), and a system that rushes it may produce only short-lived theory of little value. Emphasis on theoretical contribution also stimulates the launching of new constructs to capture essentially the same phenomenon, leading to construct proliferation, construct redundancy, and a more confused conversation instead of theory facilitating it (Colquitt & Zapata-Phelan, 2007; Suddaby, 2010). Further, although we may need to specialize as a discipline matures, having a class of pure theorists who never get any empirical dirt under their nails—in combination with having most ideas proposed in the leading theory outlet never being put to the test by others (Hambrick, 2007, p. 1350)—may lead to overvaluing rhetorical skill. Seemingly clear and useful ideas may turn out to be quite muddy

[9] Based on this sample of one, I happily proclaim that our research culture will rapidly approach sanity the more editors we get who have a background in entrepreneurship research!

and inoperable when confronted with a complex, empirical reality. Theorists who are deeply familiar with their relevant empirical domain may actually be better equipped to develop useful constructs and suggest realistic relationships, as well as being better at specifying their boundary conditions.

Finally, not all theory is good, neither logically nor morally. Sutton and Staw (1995) worry that we demand too much in terms of empirical evidence for theories. Rhetorically, they ask "whether the evidence provided by people like Freud, Marx, or Darwin would meet the empirical standards of the top journals" (p. 383). Well, don't the former two provide *excellent* examples of cases where we *should* have asked for stronger evidence, not before allowing the publication of their theories, but before implementing the ideas on a large scale, arguably doing quite a bit of damage? Further, even when they contain a grain of truth, theories may exaggerate that grain out of proportion, thereby legitimizing and increasing the instance of morally questionable or otherwise harmful behaviors (Ghoshal, 2005; Marwell & Ames, 1981). Even some Nobel Laureates have a great deal to answer for.

To sum up, theory is an indispensable tool in our quest for understanding. This said, not everything about theory is worthy of our admiration, and theory should never be a reason to turn your brain off. Ghoshal, Hambrick, and Locke are no lightweights and no dummies—I strongly recommend reading them in full in the original.

3.7 Summary and Conclusion

Congratulations! You have just made it through my random walk through personal confessions, lashes at a not-as-sound-as-it-could-be publication culture, and some actual reasoning about what theory is; its inescapability; what's good and not so good about it; and its places in research. I have argued in this chapter that theories are tools that give our research more relevance and a longer life. I described two main roles for theory: *guiding the design* and *interpreting the results*. Two theory-based designs were discussed: the *theory test* and combining elements from several theories into an *eclectic framework*. The former represents a more straightforward and manageable type of study, whereas the latter—when successful—can lead to more complete understanding of the studied phenomenon. With either approach, the theory used for design should logically also be used for analysis. However, there are situations when one is justified in introducing additional theoretical tools at the interpretation stage. These include: (a) when so doing helps *deepening the interpretation* of positive results and (b) when discussing the *alternative explanations* that should be admitted when the researcher has been lucky enough to get support for her theory. Yet another case is (c) when we want to *speculate about the reasons for an unexpected result*, such as the opposite of a hypothesized relationship or a very strong effect of a control variable. As long as the interpretation is admittedly speculative, basing one's speculations on previously overlooked theory can be no great sin.

Exploration-based generation of theory can also be justified when it truly is the case that no relevant theory exists. I have argued that such situations are not unlikely in a field like entrepreneurship, which is young and at the periphery of established disciplines. However, although one should ask questions about their applicability, the main rule for entrepreneurship researchers should be to use the theoretical tools already developed within psychology, sociology, economics, and various branches of management and organizational research. At the very least, before we decide not to do so, we should have made an effort to more fully understand what we are rejecting.

Finally, we have noted that questions of theory and method are intertwined. We want to accept or reject theory on the basis of its relevance for understanding real-world phenomena. However, in empirical testing, we run the risk of accepting theories because of method artifacts or rejecting them because of poor sampling or measurement error. So we do not really know whether the theory or the data should be rejected. In order to better justify our judgment and rhetoric about entrepreneurship theories, it is therefore critically important that the greatest care be taken in the design and execution of empirical studies. This is the topic for the next few chapters.

References

Aarts, A. A. et al. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251) (28 August 2015). Doi: 10.1126/science.aac4716. (270 co-authors under B. Nosek's leadership).

Ács, Z. J., & Audretsch, D. B. (Eds.). (2010). *Handbook of entrepreneurship research: An interdisciplinary survey and introduction* (2nd ed.). New York, NY: Springer-Verlag.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179–211.

Aldrich, H. E. (1999). *Organizations evolving*. Newbury Park, CA: Sage.

Alvesson, M., & Sköldberg, K. (2009). *Reflexive methodology: New vistas for qualitative research*. Newbury Park, CA: Sage.

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York: W.H. Freeman and Co.

Arend, R. J. (2006). Tests of the resource-based view: Do the empirics have any clothes? *Strategic Organization, 4*(4), 409–422.

Arend, R. J. (2014). Promises, premises…An alternative view on the effects of the Shane and Venkataraman 2000 AMR note. *Journal of Management Inquiry, 23*(1), 38–50.

Arend, R. J., & Bromiley, P. (2009). Assessing the dynamic capabilities view: Spare change, everyone? *Strategic Organization, 7*(1), 75.

Armstrong, J. S. (1970). How to avoid exploratory research. *Journal of Advertising Research, 10*(4), 27–30.

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of Management Review, 14*(4), 496–515.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329–366.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*, 122–147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs: Prentice-Hall Inc.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173.

Bhave, M. P. (1994). A process model of entrepreneurial venture creation. *Journal of Business Venturing, 9*, 223–242.

Carr, J. C., & Sequeira, J. M. (2007). Prior family business exposure as intergenerational influence and entrepreneurial intent: A theory of planned behavior approach. *Journal of Business Research, 60*(10), 1090–1098.

Carter, S. (2011). The rewards of entrepreneurship: Exploring the incomes, wealth, and economic well-being of entrepreneurial households. *Entrepreneurship: Theory and Practice, 35*(1), 39–55.

Chandler, G. N., McKelvie, A., & Davidsson, P. (2009). Asset specificity and behavioral uncertainty as moderators of the sales growth – Employment growth relationship in emerging ventures. *Journal of Business Venturing, 24*(4), 373–387.

Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal, 50*(6), 1281–1303.

Combs, J. G., & Ketchen, D. J. (2003). Why do firms use franchising as an entrepreneurial strategy? A meta-analysis. *Journal of Management, 29*(3), 443–465.

Davidsson, P. (1991). Continued entrepreneurship: Ability, need, and opportunity as determinants of small firm growth. *Journal of Business Venturing, 6*(6), 405–429.

Davidsson, P. (1995a). Culture, structure and regional levels of entrepreneurship. *Entrepreneurship & Regional Development, 7*, 41–62.

Davidsson, P. (1995b). *Determinants of entrepreneurial intentions working paper 1995:1*. Jönköping: Jönköping International Business School. http://eprints.qut.edu.au/2076/1/RENT_IX.pdf.

Davidsson, P. (2003). The domain of entrepreneurship research: Some suggestions. In J. Katz & D. Shepherd (Eds.), *Advances in entrepreneurship, firm emergence and growth* (Cognitive approaches to entrepreneurship research, Vol. 6, pp. 315–372). Oxford, UK: Elsevier/JAI Press.

Davidsson, P. (2006). Nascent entrepreneurship: Empirical studies and developments. *Foundations and Trends in Entrepreneurship, 2*(1), 1–76.

Davidsson, P. (2014). Getting published—and cited—in entrepreneurship: Reflections on ten papers. In A. Fayolle & M. Wright (Eds.), *How to get published in the best entrepreneurship journals. A guide to steer your academic career*. Cheltenham, UK: Edward Elgar Publishing.

Davidsson, P., Achtenhagen, L., & Naldi, L. (2010). Small firm growth. *Foundations and Trends in Entrepreneurship, 6*(2), 69–166.

Davidsson, P., Delmar, F., & Wiklund, J. (2002). Entrepreneurship as growth; growth as entrepreneurship. In M. A. Hitt, R. D. Ireland, S. M. Camp, & D. L. Sexton (Eds.), *Strategic entrepreneurship: Creating a new mindset* (pp. 328–342). Oxford, UK: Basil Blackwell & Mott, Ltd.

Davidsson, P., & Gordon, S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics, 39*(4), 853–876.

Davidsson, P., & Wahlund, R. (1992). A note on the failure to use negative information. *Journal of Economic Psychology, 13*, 343–353.

Davidsson, P., & Wiklund, J. (1997). Values, beliefs and regional variations in new firm formation rates. *Journal of Economic Psychology, 18*, 179–199.

Davidsson, P., & Wiklund, J. (2000). Conceptual and empirical challenges in the study of firm growth. In D. Sexton & H. Landström (Eds.), *The Blackwell handbook of entrepreneurship* (pp. 26–44). Oxford, MA: Blackwell Business.

Davidsson, P., & Wiklund, J. (2013). *New perspectives on firm growth*. Cheltenham, UK: Edward Elgar Publishing.

Dess, G. G., Lumpkin, G. T., & Covin, J. G. (1997). Entrepreneurial strategy making and firm performance: Tests of contingency and configurational models. *Strategic Management Journal, 18*(9), 677–695.

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jonanovich, Inc.

Edmondson, A. S., & McManus, S. (2007). Methodological fit in management field research. *Academy of Management Review, 32*(4), 1155–1179.

Edwards, J. R., & Berry, J. W. (2010). The presence of something or the absence of nothing: Increasing theoretical precision in management research. *Organizational Research Methods, 13*(4), 668.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12*(1), 1.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

George, G. (2014). Rethinking management scholarship. *Academy of Management Journal, 57*(1), 1–6.

Ghoshal, S. (2005). Bad management theories are destroying good management practices. *Academy of Management Learning and Education, 4*(1), 75–91.

Goldman, A. I. (1986). *Epistemology and cognition*. Cambridge, MA.: Harvard University Press.

Goodhue, D. L., Lewis, W., & Thompson, R. (2012). Comparing PLS to Regression and LISREL: A response to Marcoulides, Chin, and Saunders. *MIS Quarterly Management Information Systems, 36*(3), 703.

Gratzer, K. (1999). The making of a new industry – the introduction of fast food in Sweden. In B. Johannisson & H. Landström (Eds.), *Images of Entrepreneurship Research -- Emergent Swedish Contributions to Academic Research* (pp. 82–114). Lund, Sweden: Studentlitteratur.

Hambrick, D. C. (2007). The field of management's devotion to theory: Too much of a good thing? *Academy of Management Journal, 50*(6), 1346–1352.

Hansen, D. J., Shrader, R., & Monllor, J. (2011). Defragmenting definitions of entrepreneurial opportunity. *Journal of Small Business Management, 49*(2), 283–304.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408–420.

Johanson, J., & Vahlne, J. E. (2009). The Uppsala internationalization process model revisited: From liability of foreignness to liability of outsidership. *Journal of International Business Studies, 40*(9), 1411–1431.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386–408.

Kish, L. (1987). *Statistical design for research*. New York, NY: John Wiley & Sons, Inc.

Klayman, J., & Ha, Y. W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*(2), 211.

Krueger, N. F., Reilly, M. D., & Carsrud, A. L. (2000). Competing models of entrepreneurial intentions. *Journal of Business Venturing, 15*(5/6), 411–432.

Leroy, H., Manigart, S., Meuleman, M., & Collewaert, V. (2015). Understanding the continuation of firm activities when entrepreneurs exit their firms using theory of planned behavior. *Journal of Small Business Management, 53*(2), 400–415.

Locke, E. A. (2007). The case for inductive theory building. *Journal of Management, 33*(6), 867–890.

Locke, E. A. (2012). Construct validity vs. concept validity. *Human Resource Management Review, 22*, 146–148.

Lockett, A., Wiklund, J., Davidsson, P., & Girma, S. (2011). Organic and acquisitive growth: Re-examining, testing and extending Penrose's growth theory. *Journal of Management Studies, 48*(1), 48–74.

Marcoulides, G. A., Chin, W. W., & Saunders, C. (2009). A critical look at partial least squares modeling. *MIS Quarterly, 33*(1), 171–175.

Marwell, G., & Ames, R. E. (1981). Economists free ride, does anyone else? Experiments on the provision of public goods. *Journal of Public Economics, 15*(3), 295–310.

McKelvie, A., & Wiklund, J. (2010). Advancing firm growth research: A focus on growth mode instead of growth rate. *Entrepreneurship: Theory and Practice, 34*(2), 261–288.

McKinley, W. (2007). Managing knowledge in organization studies through instrumentation. *Organization, 14*(1), 123–146.

Miller, C. C., Washburn, N. T., & Glick, W. H. (2013). The myth of firm performance. *Organization Science, 24*(3), 948–964.

Naldi, L., & Davidsson, P. (2014). Entrepreneurial growth: The role of international knowledge acquisition as moderated by firm age. *Journal of Business Venturing, 29*(5), 697–703.

Obschonka, M., Stuetzer, M., Gosling, S. D., Rentfrow, P. J., Lamb, M. E., Potter, J., et al. (2015). Entrepreneurial regions: do macro-psychological cultural characteristics of regions help solve the "knowledge paradox" of economics? *PloS One, 10*(6), e0129332.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903.

Popper, K. (1992). *The logic of scientific discovery* (4th ed.). London, UK: Routledge Peterson & Co.

Priem, R. M., & Butler, J. E. (2001). Is the resource-based "view" a useful perspective for strategic management research? *Academy of Management Review, 26*(1), 22–40.

Reynolds, P. D., & Curtin, R. T. (2008). Business creation in the United States: Panel Study of Entrepreneurial Dynamics II initial assessment. *Foundations and Trends in Entrepreneurship, 4*(3).

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: The Free Press.

Ruef, M., Aldrich, H. E., & Carter, N. M. (2003). The structure of organizational founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review, 68*(2), 195–222.

Sanandaji, T., & Leeson, P. T. (2013). Billionaires. *Industrial and Corporate Change, 22*(1), 313–337.

Sarasvathy, S. D. (2008). *Effectuation: Elements of entrepreneurial expertise*. Cheltenham, UK: Edward Elgar Publishing.

Sarasvathy, S. D., & Venkataraman, S. (2011). Entrepreneurship as method: Open questions for an entrepreneurial future. *Entrepreneurship: Theory and Practice, 35*(1), 113–135.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, UK: MA: Harvard University Press.

Schumpeter, J. A. (1954). *History of economic analysis*. London, UK: Allen & Unwin.

Senyard, J., Baker, T., Steffens, P., & Davidsson, P. (2014). Bricolage as a path to innovativeness for resource-constrained new firms. *Journal of Product Innovation Management, 31*(2), 211–230.

Shapero, A., & Sokol, L. (1982). The social dimension of entrepreneurship. In C. A. Kent, D. L. Sexton, & K. H. Vesper (Eds.), *The encyclopedia of entrepreneurship* (pp. 72–90). Englewood Cliffs, NJ: Prentice-Hall.

Shepherd, D. A., & Wiklund, J. (2009). Are we comparing apples with apples or apples with oranges? Appropriateness of knowledge accumulation across growth studies. *Entrepreneurship: Theory and Practice, 33*(1), 105–123.

Shook, C. L., & Bratianu, C. (2010). Entrepreneurial intent in a transitional economy: An application of the theory of planned behavior to Romanian students. *International Entrepreneurship and Management Journal, 6*(3), 231–247.

Sorenson, O., & Stuart, T. E. (2008). Entrepreneurship: A field of dreams? *Academy of Management Annals, 2*(1), 517–543.

Steffens, P. R., Terjesen, S., & Davidsson, P. (2012). Birds of a feather get lost together: New venture team composition and performance. *Small Business Economics, 39*(3), 727–743.

Stuetzer, M. Obschonka, M., Audretsch, D.B., Wyrwich, M., Rentfrow, P.J., Coombes, M., Shaw-Taylor, L & Satchell, M. (2015). Industry structure, entrepreneurship, and culture: An empirical analysis using historical coalfields. *European Economic Review* (forthcoming).

Suddaby, R. (2010). Editor's comments: Construct clarity in theories of management and organization. *Academy of Management Review, 35*(3), 346–357.

Sutton, R., & Staw, B. (1995). What theory is not. *Administrative Science Quarterly, 40*, 371–384.

Usher, J. M., & Evans, M. G. (1996). Life and death along gasoline alley: Darwinian and Lamarckian processes in a differentiating population. *Academy of Management Journal, 39*(5), 1428–1466.

Vinzi, E. V., Chin, W. W., Henseler, J., & Wang, H. (Eds.). (2010). *Handbook of partial least squares: Concepts, methods and applications*. New York, NY: Springer.

Welter, F. (2011). Contextualizing entrepreneurship—conceptual challenges and ways forward. *Entrepreneurship: Theory and Practice, 35*(1), 165–184.

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review, 14*(4), 490–495.

Wiklund, J., Davidsson, P., Audretsch, D. B., & Karlsson, C. (2011). The future of entrepreneurship research. *Entrepreneurship: Theory and Practice, 35*(1), 1–9.

Wiklund, J., Patzelt, H., & Shepherd, D. A. (2009). Building an integrative model of small business growth. *Small Business Economics, 32*(4), 351–374.

Wiklund, J., & Shepherd, D. A. (2005). Entrepreneurial orientation and small business performance: A configurational approach. *Journal of Business Venturing, 20*(1), 71–91.
Zahra, S. A. (2007). Contextualizing theory building in entrepreneurship research. *Journal of Business Venturing, 22*(3), 443–452.
Zahra, S. A., Sapienza, H., & Davidsson, P. (2006). Entrepreneurship and dynamic capabilities: A review, model and research agenda. *Journal of Management Studies, 25*(4), 917–955.
Zahra, S. A., & Wright, M. (2011). Entrepreneurship's next act. *Academy of Management Perspectives, 25*(4), 67–83.

General Design Issues

4

Abstract

What is it we are really trying to do, when doing "research"? This chapter aims to set a sound foundation for more detailed discussions of methods issues. This also aims to organize familiar research vocabulary (e.g., sample, inference, validity, model specification, boundary conditions, causality) within a holistic framework. Further, the chapter discusses pros and cons of different research approaches ("qualitative" vs. "quantitative") and different types of data (archival, survey, laboratory, case studies) in an entrepreneurship research context. Discussions of entrepreneurship-specific challenges relating to process, analysis level, and heterogeneity form an important part of the chapter.

4.1 Getting Started at Last

After that three-chapter warm-up, it's about time we get to the real contents of this book: method issues. Well, the previous chapters have in fact more than touched on general design issues already. For example, Chap. 3 emphasized the need for theory-driven designs (although allowing for exploration when needed) and mentioned at least briefly the matching of theory and level of analysis. In Chap. 2, the detailed explanation of keywords in my proposed domain delineation—e.g., *heterogeneity*, *process*, *emergence*, *discovery/exploitation*, *contextual fit*, and *antecedents and outcomes on different levels*—contained many implicit and some explicit design suggestions. Hence, parts of this chapter will be a recapitulation and elaboration of previously introduced themes. But first of all, let's have a bit of a think about what—more generally—it is that we are trying to do when we are conducting "research." This discussion, which follows immediately below, introduces or (hopefully) reiterates a number of issues we care about in research design, as well as providing them with a structure to explain how they fit together. I will then discuss "qualitative" and "quantitative" approaches to empirical research. This is followed by sections honing in on design implications of the process- and multilevel nature of entrepreneurship, respectively. After that, I elaborate on dealing with the

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_4

heterogeneity that often plagues observational data, making comparisons difficult and conclusions uncertain. Finally, I discuss design pros and cons of different types of data: primary survey data, archival data, laboratory research, and qualitative process studies.

4.2 What We Are Trying to Do When Doing "Research"

4.2.1 The World About Which We Wish to Know and Tell

To guide our discussion, please consider the sketchy representation in Fig. 4.1. No, please don't leave yet, you structuration, process, or complexity theorists; social constructionists; and people with preferences for in-depth studies and "qualitative" data. This book *is* biased toward the types of research with which I happen to be most familiar, but the symbols in this figure should be given a generous interpretation; it does not suggest a delineation to quantitative, cause-effect "variance studies" (Van de Ven, 2007), and most of the issues discussed below apply to a broad set of research approaches.

The ellipse in the top panel denotes "the world about which we wish to know." This *may* be restricted to a particular, empirical population. In the Global Entrepreneurship Monitor (GEM), each country team tries to find out the level of entrepreneurial activity in their country during a particular year (Amorós, Bosma, & Levie, 2013). However, as I will discuss further in Chap. 5, the restrictions of a particular, existing, and accessible population are *not* usually what delimits the

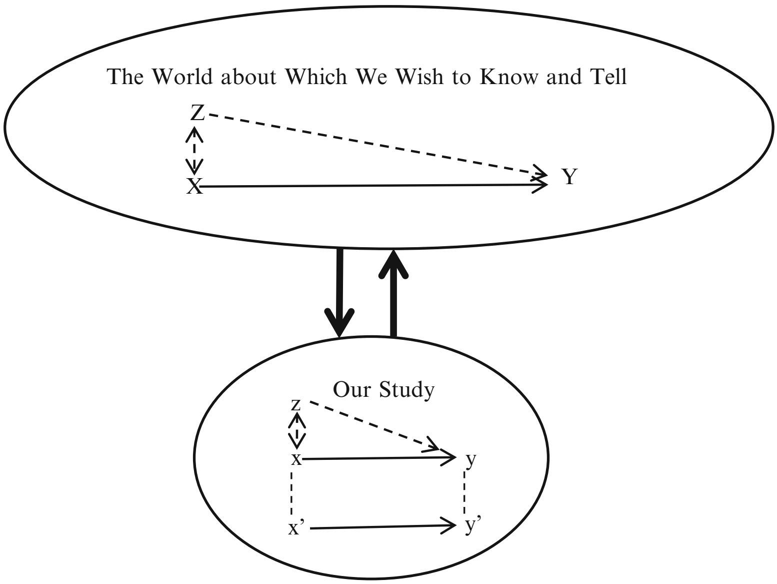


Fig. 4.1 Knowledge interest vs. empirical design

social scientist's interest. Alternatively, we may think of "the World" in question as *the abstracted phenomenon* that we are interested in. For example, "new ventures," "industrial districts/entrepreneurial clusters," or "venture capital investments."[1]

A first observation from the figure is that "the World" has an envelope, an outer border. In the context of "new ventures," this may refer to whether or not "new social ventures," "new self-employment," or "new corporate ventures" reside inside or outside of what our theory and results are assumed valid for. It can also delimit our interest in *time* (do we want to understand new venturing activity through all times, during some specific historic era, the present, or the future?), *level* (are we interested in all aspects of this phenomenon at all levels of granularity, or do we frame our interest as pertaining exclusively to the global, national, regional, firm, venture, team, or individual level?), and *space* (e.g., are new ventures in other-than-our [type of] country included, urban vs. rural, growing vs. disadvantaged regions, etc.).

In "the World" we also find Y. This denotes the specific phenomenon or phenomena that we would like to ultimately see explained. No, I'm not suggesting that each and every study should have an explanatory goal and a design that allows its achievement, but I am suggesting that eventually, we would like to see our collective journey of knowledge development succeed in explaining to some meaningful degree the interesting phenomena that we observe. *Y* can be represented by one or more theoretical constructs and one or more empirical variables. It can be broadly (or vaguely?) defined as "the dynamics of industrial districts" or more specifically as "the duration of the creation processes for innovative ventures."

The next element is *X*, the constructs or variables which may explain (some aspect or part of) *Y* and in which we have a theoretical (and/or practical) interest. No, I am again not suggesting that every study is or should be driven by "independent and dependent variables" logic. However, a lot of research more or less explicitly follows that logic and our collective knowledge-development processes as a whole do so to an even greater extent. Regardless of how precisely we actually can predict or explain a particular phenomenon *Y*—and regardless of how realistic it is that we will ever be able to do so—it is commonly the case that we have an interest in knowing what factors are worthy of attention (e.g., *relative advantage*, *compatibility*, *complexity*, *observability*, *and trialability* of innovations, as discussed in the previous chapter). If *Y* is new venture creation outcomes on the microlevel, *X* in the broadest terms would reflect some or all of the following: the characteristics of the individuals involved, the new venture idea, the resources put in, the environment, and the actions undertaken in the process. And perhaps something else that I'm not smart enough to think about—but hopefully you are….

[1] I'm almost tempted here to refer to ancient Greek philosophy for (I believe) the first time in my academic career and to the possible a- or be-musement of some colleagues. Plato held that the world of "ideas" or "forms" was more real than the empirical world "out there" (Ross, 1951). For example, the abstract notion of "cat" is the "real" cat; no particular, imperfect, flesh-and-blood cats are worthy of that honor. In a sense, it is truths about the "world of ideas" that we ultimately want to uncover through research. However, Plato would not give a rat's (or cat's) posterior for our attempts to get to such truths through empirical studies; therefore, I officially refrain from referring to ancient Greek philosophy…

Then there is the horizontal arrow, which denotes the question of how *X* and *Y* are related. Just as *X* and *Y* are shorthand for a group of constructs and/or variables, the arrows in Fig. 4.1 denote not one, unidirectional, causal, and linear relationship but any form of relationships among *X* and *Y*, where both *X* and *Y* may refer to several constructs or variables. The arrow between *X* and *Y* represents the actual relationships in "the World" we are interested in. The task of explanatory-causal research is to find out what those relationships are and to help us understand why they are what they are. To further broaden the perspective, we may also take the arrow not to mean "influence" but "time" or "sequence" and apply the logic of "process theory" (van de Ven, 2007).

Finally, we have the nasty one, *Z*, and its associated, dashed arrow. These denote the influence of factors other than those we are interested in (X) which may nevertheless have an impact on *Y*. If the influence of *Z* factors were unrelated to the effects of *X*, we could safely let them be, as the only consequence would be that our explanation of *Y* turned out less complete than it could have been. When *X* and *Z* are interrelated, we have a problem—a very important problem—because if we do not consider *Z*, our picture of the role played by *X* will be distorted. This is a major issue for research design. Why do *Z* factors exist? Because of our ignorance (internal attribution) or the complexity of the world (external attribution; see Weiner, 1985)! We may simply not know that *Z* exists. Or it may be the case that *Z* are constants in the context (place, time period) we are used to, and we don't realize that elsewhere *Z* varies and has an influence (Lieberson, 1985). Or we may neglect *Z* because it represents factors we (policy-makers, managers, business founders) cannot influence anyway or because we do not have any data on *Z*. In yet other cases, it may simply be that *Z* doesn't interest us—they're boring! But the thing is that if *Y* and *X* do interest us, we need to consider *Z*, or we will get something wrong—sometimes very wrong.

OK, we have now gone through the basic elements of "The World About Which We Wish to Know and Tell." From this, we can deduce that a research study can have one or more of the following foci:

1. What is *Y*? Identifying and describing *Y*. Revealing interesting and important phenomena for further study. For example, [types of] "high-growth firms" (Delmar, Davidsson, & Gartner, 2003) or "[venture] idea sets" (Hill & Birkinshaw, 2010).
2. What is *X*? Identifying and describing *X* and the mechanisms by which it may be related to *Y*. Given our interest in *Y*, the hunt for possible explanations begins. For example, the introduction of *effectuation* (Sarasvathy, 2001) and *bricolage* (Baker & Nelson, 2005) into entrepreneurship research can be portrayed in this way. The same goes for *celebrity capital* in our domain (Hunter, Burgers, & Davidsson, 2009).
3. How can *X, Y*, and their relationship be studied? Develop ways to sample, assess, or measure *Y* or *X* or both. Methods development that is necessary for valid studies of how *X* and *Y* are related. For example, developing measures of "opportunity recognition beliefs" (Grégoire, Shepherd, & Schurer Lambert, 2010) or "market newness" (Dahlqvist & Wiklund, 2012) or introducing new approaches to studying entrepreneurial processes (Uy, Foo, & Aguinis, 2010).

4. How are *X* and *Y* related? Assessing the direction, sign, magnitude, structure, and form of *X*–*Y* relationships. The examples are endless… except for the much more restricted number undertaking the important task of reconfirming or challenging the veracity of these relationships as portrayed in original publications (see Chap. 9).
5. What is *Z* and how can we effectively solve the problem of its distorting influence on perceived or estimated *X*–*Y* relationships (e.g., Davidsson, 2008)?
6. Where is the border of the relevant "World"? Defining the outline, finding the boundary conditions, and determining outside what sphere the focal phenomenon or population cannot meaningfully be said to "be the same" or "work in the same way." The Global Entrepreneurship Monitor encountered a problem of this kind early on when it became increasingly clear that the phenomenon of "nascent entrepreneurship" did not necessarily have quite the same meaning in countries at different levels of economic development (see Chap. 6).

These categories are worth reflecting on for a while, both for PhD student newbies and journal editors singularly focused on getting "theoretical [mechanism] contributions" from each and every submission. Clearly, although several types of studies are needed in order to complete the total knowledge-development task, there seems to be more journal demand for some types of study over others. For our current purposes, the most important takeaway here is that *since research can have so diverse objectives*, *it can hardly be the case that any one general approach or method can be superior for all purposes*. We will get back to knowledge interest and research design later in the chapter.

4.2.2 Our Study

Unfortunately, we cannot study "the whole World." If our "World" is a particular empirical population and that population is big, we can probably not afford to study all the members of the population. If our "World" pertains to an abstracted phenomenon which exists also in places where we cannot go (like the past, the future, Narnia, or North Korea), it is not just lack of resources that makes it impossible to study "The World About Which We Wish to Know and Tell" in its entirety. It is simply not accessible for empirical assessment. Therefore, our empirical study will not be able to study (all of) the relevant "World." This is what the lower panel in Fig. 4.1 is getting at. I should admit up front that when discussing "Our Study" I will mostly assume an explanatory-causal type of (ultimate) knowledge interest.

This graph illustrates a few things. First, "Our Study" is visibly different from the relevant "World." The way it is drawn it is not even clear whether "Our Study" forms part of the relevant "World." Further, "Our Study" is smaller than the relevant "World." In reality, "Our Study" is usually a tiny corner of the latter. These observations highlight issues of *representativeness* and *sampling*. Does what we include in our study adequately represent "The World About Which We Wish to Know and Tell"? If the "World" is a well-defined and accessible population, we can help the situation by applying *random* (or rather, *probabilistic*) *sampling* and techniques to avoid and assess *non-response bias*. This is all standard methods textbook stuff, which I will not repeat in any detail. It's worth reminding, though, that we can never know or guarantee that a sample is representative; we can just apply procedures that reduce the risk that it is not.

If the "World" is a world of past, present, future, and other-place phenomena, then the issue of probabilistic sampling may also enter the picture but only as a secondary concern. The more important question is whether the slice of empirical reality we choose to study adequately represents the phenomenon we wish to study. That is, whether the studied context and/or sample is *theoretically relevant*. Participants in an experiment are rarely randomly sampled, and few people have a problem with that—but does the experimental setting represent the corresponding world life situation, and do the participants represent a good testing ground for the theoretical question you are asking? This question explains why it is important to triangulate results using different methods (Cialdini, 1980) and why it may not be so smart to use undergraduate students as experimental participants if you wish to study the emotions or decision-making of "entrepreneurs." Further, should you let modest/mundane start-ups dominate your sample just because they are more frequent in empirical populations (Davidsson & Gordon, 2012)? We will discuss sampling and representativeness at length in Chap. 5.

The figure also hints at other design issues. We will not even try to study all of X and Y; due to limits to access and affordability, we instead only get at the subset X and y. But it doesn't end there: we do not even get to these directly but only through our *operationalizations* x' and y', which may be subject to *measurement error* of two types: random and systematic. Thus, there are issues to consider regarding the *validity* and *reliability* of our operationalizations. This is represented by the dotted lines between X and x' and between Y and y', respectively. If the measures (operationalizations) are sufficiently free from error, we have established *measurement validity*. If we have not measured X and y with adequate precision, we cannot correctly assess how they are related, either. This is also standard methods stuff that is covered in many other books. This said, we will devote Chap. 6 to operationalization in the context of entrepreneurship.

The next task is to get the $x' \rightarrow y'$ arrow right. This is about *model specification* (deciding on the structure and form of the relationship) and *statistical estimation* (assessing the sign and strength of the relationship assuming that the model is correctly specified). We should note that statistical estimation can to an extent also be used for the dashed x–x' and y–y' links (e.g., Cronbach's alpha, inter-rater reliability assessment, measurement model [as opposed to structural model] results in structural equations modeling)[2].

The first questions pertaining to "the arrow" are whether X and Y are related at all. Is there a *correlation* of meaningful magnitude?[3] If there is, we want to know

[2] If this is Greek to you, don't worry! You just have to learn Greek.

[3] Note that the correlation may not be linear and may be obscured by multivariate patterns in the data, so the zero-order, Pearson correlation (and similar) only gives a rough indication of whether any type of meaningful association exists between the variables.

whether the relationship is causal or merely associative. The notion of *causality* is in itself tricky indeed and the subject of many learned works (e.g., Pearl, 2000; Russell, 1912). I will in this book presume that it is meaningful to talk about and investigate relationships as if they were causal in nature. However, it makes a lot of sense for those who research complex, social phenomena such as entrepreneurship to study notions like "probabilistic causation" (Menzies, 1989), "deterministic chaos" (Derbyshire & Garnsey, 2014), and "INUS conditions" (Horsten & Weber, 2005)—and to reflect from time to time on what we are really doing when we are suggesting causal relationships.

I have drawn the arrow from X to Y, thereby bypassing the next (or rather, parallel) important question, which is of *direction*: does X precede and/or influence Y, or could it be the other way around? If our theory says the direction is from X to Y, this implies our design should strive for *time separation* such that our measure x' predates our measure y'. Other important aspects of the relationships—the arrow—pertain to the *sign*, *magnitude*, *structure*, and *form* of the relationship. Is the effect positive or negative? How strong is it (Cohen, 1990)? Is it direct, mediated, or moderated (Baron & Kenny, 1986; Edwards & Lambert, 2007)? Linear or nonlinear? Specifically, what nonlinear shape? Our analysis will give us *some* results no matter how we specify our model; however, it will only yield *valid* ("true") results if our model captures the true direction, structure, and form of relationships in "the World."

This includes *ruling out alternative explanations* which entails dealing with nasty Z. How can we eliminate biasing influence from factors we are not interested in? We will deal with this later in this chapter and return to the issue as needed in upcoming chapters. However, we should note here that as we move from "the World" to "Our Study," we may introduce additional Z factors in the form of *method artifacts*. That is, our estimated relationships may be unintentionally inflated (or deflated) by elements of our design. For example, our x' and y' may correlate highly in part because they were measured in similar ways, e.g., Likert-scale responses given by the same person on the same occasion (*common-method bias*, Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). In an experimental study, members of my own sex may underperform if the experiment is led by a stunning woman because they get so eager to "impress the Sheila"[4] that, of course, they fail (sorry, can't remember the reference for this beautiful result, but yes, we're soooo pathetic!).

Finally, taking an explicit process perspective, we might focus on the "content" of the arrow and ask what characterizes the journey from nonexistence to existence of new ventures (Gordon, 2012; Lichtenstein, Carter, Dooley, & Gartner, 2007). What are the necessary steps and their observed/optimal sequence (Shane & Delmar, 2004)?

If in "Our Study" we have managed to establish that (a) there is a relationship which (b) given our study design, it is reasonable to interpret as causal, (c) we have also made a good case that our measures are valid and (d) that our model correctly represents the true structure and form of relationships, and (e) we have properly

[4] In Australia, Sheila is a slightly derogatory—and disappearing—label for the human female.

eliminated other explanations and biasing influences of nasty Z—then we have established *internal* (*conclusion*) *validity*. That is, the results are valid within the limits of our study; we have found out what is true for those cases actually studied.

There is reason to stop and reflect here, because this collection of requirements clearly shows that many things can go wrong. Doing social science is very, very challenging, and that fact should humble us to firstly really make our best effort to get our research design right and secondly not to get too loud and cocksure about our findings. We should not be too sure about the findings of other studies, either, when evaluated in isolation. This is why replications—sadly underappreciated as they are (Evanschitzky, Baumgarth, Hubbard, & Armstrong, 2007)—are crucially important if we are serious about building knowledge and not just about getting published. Despite all of this, the impossibility of doing perfect research should not dissuade us from trying or from holding up our best available evidence against loudly expressed opinions lacking every ounce of systematic backing.

Getting to a point where you can be confident of the internal validity of your study's findings—and able to convince others about it—is quite an achievement. And yet, internal validity is but one step on the way. The next crucial step is to find reason to argue that what we found in "Our Study" actually speaks to "The World About Which We Wish to Know and Tell." Can we trust that our internally valid results apply elsewhere? This is the question about *generalizability* and/or *external validity*. When we have the limited aim to generalize on statistical grounds to a specific, empirical population, we can use *statistical inference theory* and its specific tools such as *statistical significance tests* and *confidence intervals*. Note that statistical inference is different from statistical estimation. The latter is about finding the best representation of relationships in the data at hand and is, if you ask me, quite an impressive and very useful set of tools. Statistical inference is about using "Our Study" to make statements about "the World." It is the "$p<0.01$" and "$\alpha=0.05$" and the asterisks (*, **, ***) that sprinkle research articles. It does not speak directly about the strength of effects (that's estimation) but about the statistical uncertainty of our estimated results.

Statistical inference theory (significance testing) is a useful tool for some purposes, e.g., in the context of randomly assigned experimentation and—if we are a bit generous—studies like GEM. In most research applications that you will encounter it is more misleading than illuminating, not least because the conditions for its valid use are usually not fulfilled (by a country mile or two) but also because it is a much weaker truth criterion than the users believe it is (or want to wish it were). We shall have more to say about this in Chaps. 5 and 9. When "the World" we are interested in is bigger and more vaguely delineated than a particular, empirical population, the help we can get from statistical testing is even more limited. Since this normally *is* the case for researchers' worlds of interest, it is inescapable that generalization from an individual study will have to rely on the quality of the design of the empirical study and the logical argument that links it to the abstracted phenomenon we are interested in. If we do well, we can convince readers that our study represents one little relevant corner of the relevant world; however, no one has reason to believe that our single study, no matter how well executed, is the final say about "the World" in its entirety.

In a qualitative study, we may have extracted a seemingly useful new construct. We can then sit at our desk and analytically think through how this construct may apply and manifest itself in a range of other empirical contexts. This is useful, but at some point there will be a need to test the veracity of our hunches. As regards theory-testing research—the arrow in Fig. 4.1—the relationships need to be retested and extended in replications. There are so many threats to internal and external validity that relying on a single study—no matter how well executed and in what prestigious outlet it appeared—is just ridiculous if we are interested in "the World" rather than just "Our Study."

In this section, I tried to create a road map that organizes a number of important notions pertaining to research design. Figure 4.2 reiterates Fig. 4.1 with many of these notions inserted in their proper places. In the rest of this chapter as well as in chapters to follow, I will occasionally refer back to these figures and the design terms they contain.

Is there anything entrepreneurship-specific about what's depicted in Fig. 4.2? Well, although entrepreneurship research is maturing, it is still a comparatively young line of research, and some of its branches are relative newbies. This implies

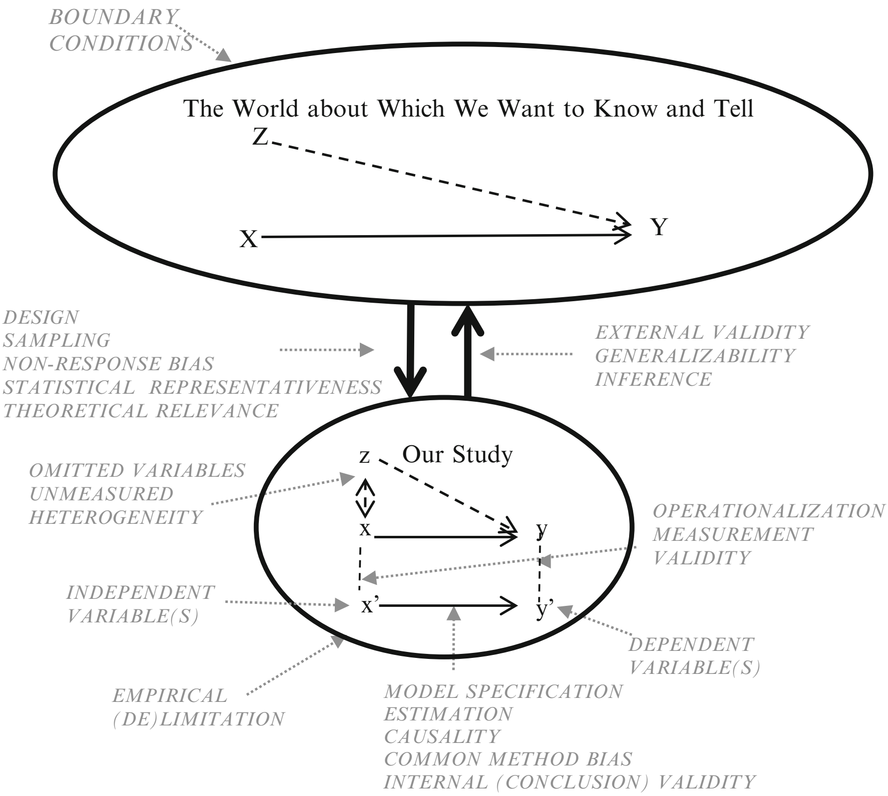


Fig. 4.2 Where design terms belong in the research landscape

that research with a main emphasis on identifying, describing, and operationalizing some X or Y may be more justified and needed than is the case in a fully mature discipline. Sometimes the evidence on X–Y relationships has to wait until we have the basics sorted (Dimov, 2011). Let us also revisit some of the descriptors from my domain delineation. *Heterogeneity*: arguably, early-stage activities that have not been subjected to the full forces of market selection and pressures to conform will be even more heterogeneous than are more mature business activities. This gives reason to be particularly wary about Z variable distortions. The emphasis on *contextual fit* highlights a similar issue. Heterogeneity applies not only to what X variables influence Y but also the strength and form of their influence. This highlights the notion of contingent relationships, e.g., that the arrows in the figure include the possibility of moderations and mediation. The notions of *emergence* and *process* point to even more intricate challenges that may be particularly pronounced in an entrepreneurship context. Some X or Y variables may not meaningfully exist at very early stages, or they may radically change during the entrepreneurial process. This may be the case for the founder's motivation and goals or the strategy that drives the process. Similarly, the $X \rightarrow Y$ relationships may be stage dependent, as may relevant subjective and objective success criteria (Y). Lots to think about. Many ideas for how we can do things a little bit better than the previous study. Didn't I say somewhere early in this book that entrepreneurship research is challenging and therefore great intellectual fun?

4.3 "Qualitative" and "Quantitative" Studies

4.3.1 "Quantitative" vs. "Qualitative": A Confused Distinction

Discussions of "qualitative" vs. "quantitative" research are often confused for a number of reasons. Firstly, people often unsystematically bundle three aspects of "quantitative": (a) the use of many cases (population studies or surveys), (b) the use of formal measurement, and (c) the use of computerized (mathematical or statistical) tools for data analysis. These may be related but do not always go together. For example, conjoint studies can produce quantitative estimates for the entire sample as well as for small subgroups or even single individuals (Lohrke, Holloway, & Woolley, 2010); case studies can be highly quantified in the measurement sense, and qualitative research increasingly relies on computerized tools for analysis of verbal data, similar to statistical packages. Second, people often and non-justifiably equate the nature of the data with issues of philosophy of science, rigor, and depth. There are no such one-on-one relationships. Data themselves do not know how the researchers are going to use them, and shallow, qualitative work is as possible as is sloppy (i.e., non-rigorous) quantitative work. Third, people often implicitly or explicitly portray the one as generally superior to the other. That's stupid. Quality has to do with using the right tool for the purpose at hand and how skilled you are at operating and creatively combining the appropriate tools.

My take on these things largely overlaps those expressed by Edmondson and McManus (2007). It is an article well worth studying, preferably in conjunction with Cialdini's (1980) essay on "full cycle" research. In essence, Edmonson and McManus associate qualitative data with *exploratory* research of phenomena that are not yet well understood. In terms of Fig. 4.1, this means "What is *Y*?", "What is *X*?, and "By what mechanism may they be related?". However, quantitative data can also be used for exploration. Several quantitative techniques, such as data mining algorithms, cluster analysis, and (exploratory) factor analysis, can unveil unknown connections among variables. They can group related variables or cases, which can lead to identification of new theoretical categories or constructs. This said, quantitative data are typically more associated with *explanatory* or *theory-testing* research, e.g., testing theory-driven hypotheses about the sign and magnitude of direct and moderated/mediated effects of *X* on *Y*. Some research questions or tasks are *inherently* quantitative in nature. For example, *statistical* generalization to a larger population only makes sense when the number of cases studied is relatively large and probabilistically sampled from a target population. Estimating the *strength* of a relationship across cases or over time inherently requires quantitative data at least in the measurement sense. If you talk about group differences or relationship strength without measurement, you're speculating.

Edmondson and McManus (2007) propose there is a progression of knowledge states over time, where research moves from *nascent* over *intermediate* to the *mature*. This progression is paralleled by qualitative-exploratory to mixed methods to quantitative, theory-testing approaches. Such patterns are identifiable in entrepreneurship research. For example, the early work on *entrepreneurial bricolage* was typically case based (Baker & Nelson, 2005; Garud & Karnøe, 2003) and has just recently moved on to measurement and theory testing (Rönkkö, Peltonen, & Arenius, 2013; Senyard, Baker, Steffens, & Davidsson, 2014).

This book deals mainly with so-called "quantitative" research. Again, this is not because of an alleged general superiority of such approaches but a simple consequence of my lack of expertise in qualitative methods. Although I embrace the ideology that our knowledge development processes are incomplete without theory testing, I firmly believe that different types of research are helpful for gaining insight into entrepreneurship. In fact, many of the studies that I cite often or otherwise admire and hold as exemplars to my research students and build on small samples and/or exploratory approaches (e.g., Amit & Zott, 2001; Baker & Nelson, 2005; Bhave, 1994; Fauchart & Gruber, 2011; Sarasvathy, 2001; Shane, 2000; Van de Ven, Polley, Garud, & Venkataraman, 1999). There are multiple routes to insight. Therefore, researchers who say or think "I cannot see any meaningful knowledge coming out of that research approach" should realize that this may reflect a deficiency of "I" and not necessarily an inherent shortcoming of "that approach" (cf. van Burg & Romme, 2014).

There are some characteristics of the entrepreneurship research domain as I have portrayed it, which point at a need for exploratory—including qualitative—research. One is the relative *youth* of the field, although it has matured by a decade since the first edition of this book. At least for some aspects of entrepreneurship, we may simply not have had time enough yet to familiarize ourselves with all facets of this empirical phenomenon or to exploratively develop all the theory we need (identifying the *X*s and *Y*s in Fig. 4.1 and pondering the possible mechanisms that link them). Another indicator of need for qualitative-exploratory research is the *heterogeneity* of the phenomenon, which I pointed out in Chap. 2. If we only did research at arm's-length distance, there are the risks that because the relationships are different for different parts of the heterogeneous population we would either come out with only weak results or results that are "true" on average but not for most individual cases (this is an aspect of "nasty Z" and also of the structure and form of "the arrow"). Close-up information may be needed in order to learn about the heterogeneity, so as to assess what abstractions and generalizations we probably can and cannot justifiably make (making it an aspect of the border of the relevant "World" as well). Further, at least when we think about more spectacular forms of innovative entrepreneurship, we are dealing with events that are *infrequent*, *unanticipated*, and/or *extraordinary*. Phenomena of this kind may be difficult to capture with conventional, quantitative approaches. It is worth pondering that at the extreme of conventionalism, the most spectacular instances of entrepreneurship would invariably end up as disturbing and possibly deleted outliers in regression analyses (cf. Crawford, Aguinis, Lichtenstein, Davidsson, & McKelvey, 2015; and Chap. 10).

Another aspect that I have highlighted in my entrepreneurship research domain delineation is the *process* character of entrepreneurship. This may also call for qualitative, or close-up, approaches (cf. McMullen & Dimov, 2013; Van de Ven et al., 1999). An early insight I had as a researcher was, in fact, the difficulty of capturing processes in survey research. One of the cases in the pilot study for my dissertation made this particularly clear. This case was about a small manufacturing firm in a shambles. The rational thing to do would have been to file for bankruptcy, but the founder-manager just couldn't stand the thought of it. The firm was heavily in debt to suppliers and tax authorities alike. Old, uncomfortable facilities led to high personnel turnover and difficulties with recruiting. Insufficient profit margins made it impossible to catch up. At this point, a series of events led to a turnaround. The son assumed a serious management position and started by checking the profitability of different customers, an exercise that led to the conclusion that many long-term relationships were in fact unprofitable and should be reconsidered. By fortuitous coincidence, the firm secured two new customers in a growing industry, which led to a reorientation that made it easier to attract—and develop special products or services for—additional, profitable, and growing customers in that industry. Around the same time, the firm reached a deal with the tax authorities for a realistic plan for catching up with tax payments. Backed by this deal and the new customers, the firm was offered new facilities on favorable terms in the municipality's modern industry park, which made the firm a much more attractive place of work. The reduced personnel turnover and increased job satisfaction in turn led to higher productivity and profits and so on (Davidsson, 1986).

Some of these events are causally related (although the sequence could equally well have been a different one) whereas others just happened to coincide. Clearly, virtuous cycles of this kind would be very hard to capture in quantitative work and entirely impossible with a cross-sectional survey design. For such reasons, early attempts to make sense of start-up sequences were not particularly successful at finding meaningful patterns (e.g., Liao & Welsch, 2008; Liao, Welsch, & Tan, 2005). I still believe it possible to capture important aspects of entrepreneurial processes in longitudinal, quantitative studies, and the evidence suggests that with better theoretical abstractions and improved empirical techniques, we are getting there (Gordon, 2012; Lichtenstein et al., 2007). I am equally convinced, however, that close-up insights from cases like the one above are an indispensable input to the theorizing that is necessary for good, quantitative work on entrepreneurial processes.

4.3.2 Bad Research Practices: Addressing "Quantitative" Questions with "Qualitative" Research and Dressing Up Your "Quantitative" Research in Stolen Outfits

I thus argue that entrepreneurship research needs both exploratory and explanatory approaches, and qualitative as well as quantitative data. However, there has to be a proper match between the research question and the chosen approach. The main problem I have with qualitative research is when researchers using such approaches make claims about issues their approach is fundamentally inadequate for addressing. Let me share a little anecdote to illustrate this. Some years ago, I was at a presentation of a qualitative study of business founders. The cases were chosen because the founders were female and the start-ups were in a particular, recently deregulated industry. The data were collected through retrospective interviews. Several of the interviewees reported they had difficulties obtaining the bank loans they needed, and when prompted, some of them ascribed this to the fact that they were women. Because of this, the researcher publicly claimed that women entrepreneurs were discriminated against by the banks[5]. I protested.

Saying that banks systematically discriminate against a particular group is a very serious accusation, and because people have a high degree of faith in what researchers say, I get pretty upset when researchers make strong claims like this on the basis of very shaky—or in this case, no—evidence. For heaven's sake, if we want to establish that women business founders have difficulty obtaining bank loans because they are women, then for a minimum, we need to investigate a group of subjects that is *representative* for the category "women business founders." That is, we need to establish correspondence between "Our Study" and "The World About Which We Wish to Know and Tell" (Fig. 4.1). These women came from a relevant population (actually a well-selected one). However, due to the small, nonrandom sample and the absence of men in the study, we cannot know that the credit difficulty experiences of the investigated women represent the experience of women in general in

[5] You may note that there is no evidence of solid grounding in some philosophical stream in this research or much evidence of depth of analysis. Such virtues do not follow automatically from the choice of data of verbal or "small n" nature, as some seem to believe. By the way, those who think my politically incorrect choice of example proves I'm an MCP are referred to the fact that I was a *proud* co-supervisor of Helene Ahl's doctoral dissertation (cf. Ahl 2002, 2006), which in my view is an excellent piece of feminist research. What the choice of example proves is probably that I have not encountered too many blatant examples like this, which is somewhat encouraging.

the empirical population from which they were sampled, or that they represent the theoretical category "women [who start businesses in newly deregulated industries]" in this regard, or that their credit difficulties had anything to do with their womanhood. In addition to a less questionable sample, we would have to *measure* the frequency of loan refusals and *compare* the results with another group relative to whom women are said to be discriminated against (i.e., probably male business founders). There can be absolutely no escape from these requirements. In addition, we should preferably also be able to rule out other substantive explanations (such as industry, experience, education, venture size, size of loan application relative to own funds, methods artifacts such as prompting the interviewee with the possibility of sex discrimination, and so on; the nasty *Z*s). Further, if we can make significance testing remotely applicable, we would like to rule out the possibility that the group difference we have established could easily be due to stochastic variation. In this case, we had none of this. None. All we had was a few women from a judgment (or convenience) sample saying they had problems getting loans and—when prompted—reporting that this might have something to do with the fact that they were women (i.e., an attribution, cf. Weiner, 1985).

I am quite convinced there is solid research evidence elsewhere that women are discriminated against in many societies. For example, although I do not remember the specific reference, I have come across rather convincing research showing that as the proportion of women in a profession increases, the relative salary level goes down (at least back in the 1990s). Certainly an uncomfortable truth. Knowledge of discrimination in other domains in conjunction with previous experiences by these women may have made their suspicion of discrimination a reasonable hypothesis. The research described above is not, however, the quality of evidence needed for researchers to make strong claims about sex discrimination on the part of a specific group of actors. To make matters worse, the accusation was in all likelihood false. Comprehensive and systematic research on precisely that matter (in the same country) was published at about the same time, arriving at the conclusion that women entrepreneurs were *not* discriminated against by banks (Björnsson, 2001), and a review of the then available international entrepreneurship research literature suggested that the evidence in support of sex discrimination by banks was very limited (Ahl, 2002).

More generally, the simple fact is that questions about quantitative differences (more, better, stronger, more often, etc.) between groups, or about such within-group changes over time, require a research design that matches the question. My design choice in this case would have been an experiment where identical loan applications for identical businesses were submitted to actual loan officers, only manipulating the sex of the applicant. This is not to suggest that "small n" research and/or research using nonnumerical data could in no way inform the question of whether women entrepreneurs are discriminated against by the banks. It is just that retrospective interviewing of a handful of women entrepreneurs about their experience and attributions of problems in getting loans is probably the worst conceivable design in this case (in all honesty, this was not the central question for the researcher in question, either). A piece of useful research that would be classified as "qualitative" by conventional criteria and which would get at issues that are unlikely to be within reach for a survey approach would be a participant observation study, where the loan officers' way of talking to and about male and female loan applicants were studied. If there were discrimination, such a study would not only give strong indications of this fact but also offer insights into the mechanisms behind it, which may elude "my" experimental study. In order to impress a researcher of my own ilk, however, the presented evidence should not just be a number of illustrative quotes that support the researcher's hypothesis but convincing evidence that the loan officers' treatment of women applicants was *systematically* different and that this was to their disadvantage.

The quantitative equivalent of the above case would be a study that rushes out and tries to measure variables and estimate relationships before enough qualitative knowledge and conceptual development have been established—and which then delivers strong policy advice with implications for people's livelihood on that basis.

More commonly, when quantitative research is disappointingly bad, it is bad in other ways: asking a meaningless research question, applying poor sampling or poor design of experimental manipulations or questionnaire contents, or using incorrect modeling, analysis techniques, or interpretation of the results. We have seen it all and all too often (however, we should remember that doing really good research is darn difficult). A particularly disappointing type of bad quantitative research is the silently sanctioned (even by rather prestigious journals and scholarly associations) "normal practice" of telling a streamlined, hypothesis-testing story where most of the relationships come out as theoretically predicted, while we never get to learn about many of the dozens of tweaking choices that it took to get there, after having examined the (preliminary) empirical relationships in the data (O'Boyle, Banks, & Gonzalez-Mulé, 2014; Simmons, Nelson, & Simonsohn, 2011). This includes trimming or expansion of the sample, selection of the dependent variable that "works best" when several indicators are available (e.g., of "performance"), selection of independent and control variables to include, adjustment of the structural relationships in the model[6], etc. Many of these adjustments *may* be sound in terms of finding out what relationships in the data we should believe in, but if so they reflect shortcomings in the design and data collection stages. It remains an inescapable fact that the procedures make the research rather exploratory and leads to gross exaggeration of the confidence we should have in the findings on statistical (inference) grounds. Few active researchers are completely free of guilt in this mat-

[6] I recall with a mix of horror and amusement taking a course in the then dominant structural equation modeling (SEM) package LISREL offered by its creators, Jöreskog and Sörbom, in the late 1980s, and seeing Sörbom proudly introduce their latest invention—the automatic modification index—which allowed you to obtain maximum fit between model and data at the push of a button. Now, exploration isn't necessarily wrong if the revised model makes sense and the research process is transparent in the published report (see, e.g., Davidsson, 1989a, Chaps. 4 and 5). Further, the exploration can be combined with testing if half the sample (tryout) is used for finding the model and the second half (holdout) is used for testing it on untouched data. This is recommended in some methods textbooks. The problems are that researchers don't heed the call—and journals don't sufficiently demand or reward this procedure.

ter (I'm not), and it is a tall order to single-handedly try to change a system (I've tried… and I keep trying). This said, I think we have an obligation to be honest to ourselves and call this phenomenon exactly what it is: a form of academic misconduct. If that's what we wish to be involved in, we should now go on with business as usual. If not, we should do our best to change our ways.

4.3.3 The Best of Both Worlds

As I see it, the most fruitful way forward for entrepreneurship research would be integrated research programs that include several types of research addressing different aspects of the same issues. This would make for real cross-fertilization between different approaches, rather than having different camps of researchers develop separate discourses that are ignored by the other camps. I have very positive experiences from a research environment that combined a focus on the phenomenon of entrepreneurship with an openness regarding paradigms and techniques for studying that phenomenon. The entrepreneurship-related doctoral dissertations from the Jönköping International Business School[7] have been a healthy mix of longitudinal survey studies (Dahlqvist, 2007; Jenkins, 2012; McKelvie, 2007; Naldi, 2008; Samuelsson, 2004; Wiklund, 1998) and a broad variety of qualitative approaches to data collection and analysis (Brundin, 2002; Garvi, 2007; Hang, 2007; Karlsson, 2005; Lövstål, 2008; Markowska, 2011) including an ethnography (Wigren, 2003) and a Foucauldian discourse analysis using text as data (Ahl, 2002)—to be contrasted with experimental and other "laboratory" approaches (Bruns, 2004; Gustafsson, 2004; Hunter, 2009); work based on advanced, customized, longitudinal data sets (Hellerstedt, 2009) and those using mixed, qualitative-quantitative data (van Weezel, 2009). For anyone who doesn't embrace the totally unacademic notion that "all the good/smart guys are/do like us," this is a wonderful type of research environment to be in, and I feel confident that some pressure to address quality standards of other paradigms, i.e., to have to deal with the blind spots of one's own paradigm, had a positive influence on all of these works, contributing to several of the authors winning national and international awards and continuing to successful journal publication.

4.4 Entrepreneurship Research as the Study of Processes of Emergence of New Ventures

What are the method consequences of the research focus implied by this heading and by my previously presented domain delineation? The keyword *new ventures* suggests that in order to belong in the entrepreneurship domain, the research has to meet the requirement of explicit consideration of new venturing within or

[7] The comments on the Jönköping environment are copied close to verbatim from Davidsson (2013).

associated with the studied type of entity. As long as this requirement is fulfilled, the research can be conducted on any level of analysis—individual, firm, industry, region, nation, or something else (cf. Davidsson & Wiklund, 2001). That is, the research design should at least include the middle box in Fig. 4.3. Preferably, the research should pay attention to antecedents and/or outcomes as well, but this is not indispensable in the same way. On the individual level, we thus cannot confine the research to, for example, owner-managers' personal characteristics on dimensions assumed to be entrepreneurial, as related to the size (an outcome) of their businesses. In order to qualify as entrepreneurship research, there should be assessment of the middle box: new venturing activities by these individuals. On the region level, studying the relationship between structural characteristics of regions and their economic growth (or well-being) does not become entrepreneurship research until the quantity and/or quality of regional business venturing is introduced as the mechanism of such a relationship. We will return to sampling and operationalization issues on different levels of analysis in the coming two chapters.

The emphasis on *processes* implies that we need longitudinal research, which has traditionally been short in supply in entrepreneurship research (Aldrich & Baker, 1997; Chandler & Lyon, 2001) although their prevalence has increased in recent years (Crook, Shook, Morris, & Madden, 2010). What do we need longitudinal research for? First, in order to establish causality, we need to establish for a minimum that the alleged cause precedes the ensuing effect. To take an entrepreneurship example of this problem, consider the hypothesis in early entrepreneurship research that entrepreneurs were characterized by a more internal *locus-of-control* (Rauch & Frese, 2007). Having a more internal locus-of-control roughly means that you believe in your own ability to control your destiny, as opposed to it being governed by fate or powerful others. Some cross-sectional studies have supported the idea that entrepreneurs (here meaning business founders and/or owner-managers) have a more internal orientation than others. But would not such an orientation be a likely *outcome* of being a business owner-manager, as opposed to being bossed around by superiors within a hierarchy? Hence, a positive correlation is not enough. In the absence of longitudinal research showing that internal locus-of-control precedes business founding the hypothesis that an internal orientation *causes* individuals' choices of an entrepreneurial career remains just that: a hypothesis.

The study of processes involves more, however, than static comparison of a beginning state and an end state. Quite a number of things happen between the initiation of a venture start-up process and its completion/termination (Bhave, 1994; Carter, Gartner, & Reynolds, 1996; Davidsson & Honig, 2003; Davidsson &

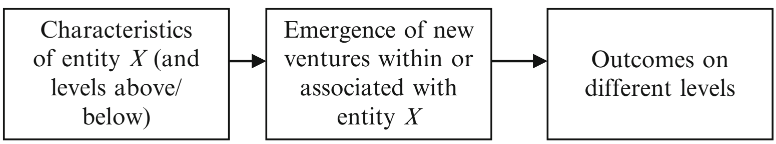


Fig. 4.3 Entrepreneurship research design possibilities

Klofsten, 2003; Delmar & Shane, 2004; Gartner & Carter, 2003; Gordon, 2012; Katz & Gartner, 1988; McMullen & Dimov, 2013; Lichtenstein et al., 2007; Sarasvathy, 2001; Van de Ven et al., 1999). Therefore, we need longitudinal designs with repeated assessment of the ventures' development over time in order to adequately capture those processes.

The emphasis on *emergence* suggests that we should catch new ventures early in the process (cf. Davidsson, Gordon, & Bergmann, 2011; Gartner & Shaver, 2012; Gartner, Shaver, Carter, & Reynolds, 2004). How can we study emergence? All existing business activities are eligible for retrospective studies, but such studies would be subject to severe selection, retrospection, and hindsight biases. For several reasons, it is preferable to study the processes *as they happen* or as close to that ideal as possible. Regarding retrospection and hindsight biases, it is well known in cognitive psychology that memory is constructive in nature (Anderson, 1990). This means that no matter how honest and careful a respondent is, he or she will in retrospect distort the image of what happened during the start-up process. Dead ends will likely be forgotten, and certain actions will be ascribed a rationale that only fell into place afterward. Such problems can to some extent be remedied through triangulation (second informant, written documentation), but serious distortions are likely to remain regardless of such efforts. Selection bias concerns the need to study also "unsuccessful" or preoperationally terminated processes (cf. Chaps. 2 and 8). For one thing, this is needed in order to acknowledge the fundamental *uncertainty* that we highlighted in the domain delineation. If we study only completed start-up processes, we may forget that completion is by no means a certain outcome for the newly initiated project.

The problem of selection bias is potentially even more serious than retrospection. In order to illustrate this, consider the following example. Imagine that for some peculiar reason, we wanted to study "factors that lead to success at gambling." We design the study so that we include only those gamblers who actually won and thus left the day at the races (or whatever the venue might be) with a net gain. This is akin to studying only those founders who actually got their venture up and running. Analyzing our data, we would arrive at the following conclusions:

- To gamble is profitable.
- The more you bet, the more you will win.
- The more unlikely (higher odds) winners you bet on, the more you will win.

While true for *winners*, these conclusions are, of course, blatantly false inferences for the *entire population of gamblers*[8]. On average, gamblers do not win; the organizer of the gambling does. Likewise, ceteris paribus the expected loss increases linearly with the size of the bet and not the other way around. And, of course, the

[8] Remember this definition: *Gambling=A tax on people who are bad at math*. Sadly, politicians in my current home country have really taken this definition to heart and use gambling as a means of taxation quite unscrupulously. Gambling typically generates double-digit percentages of total state-level tax revenue in Australia. It is a national disgrace, if you ask me.

proportion of gamblers who lose is larger among those who bet on long shots. But since we study only winners, the above are the results we will get. The scary fact is that by studying only those start-up processes that led to a successful start-up, we make ourselves guilty of the same kind of error and open up for the potential of arriving at equally biased results. We will continue the discussion of early catch when dealing with sampling issues in the next chapter (Chap. 5).

4.5 A Few Words About Levels of Analysis

My domain delineation (Chap. 2) emphasizes "antecedents and outcomes on different *levels of analysis*." This is about where in a hierarchy of aggregation the *X* and *Y* in Fig. 4.1 reside. If both are on the same level (e.g., the business and human demographics of regions explaining differences in regional start-up rates), we have a *single-level design*. If *X* and *Y* are on different levels (e.g., founder characteristics explaining venture outcomes), we have a *cross-level* design. If more than one level is explicitly considered on either the *X* or the *Y* side, the design is *multilevel*.

My interest in levels grew out of problems encountered in my own research and led to the musings in Davidsson and Wiklund (2001). I noted levels ambiguity when we were designing the *Panel Study of Entrepreneurial Dynamics* (PSED; Gartner, et al., 2004; Reynolds, 2007). Was it founders or emerging ventures we were following over time? When the venture is led by a team of changing composition and/or the idea behind the venture changes radically over time, the specification of level makes a big difference. Likewise, my research on "high-growth firms" (Davidsson & Delmar, 2006; Delmar et al., 2003) highlighted the difference between job creation on the level of the firm and the level of the entire economy, not least because old and large "high-growth firms" grow mainly through acquisitions (which means moving existing jobs from one organization to another). Further, success for the entrepreneur and/or their venture may not translate to good societal outcomes, a distinction I personally find extremely important. If entrepreneurship were solely about how individuals create fortunes for themselves, I would have very little interest in it (cf. Chap. 1—and the sad state of my bank statements…).

Of course, levels problems had been observed and contemplated before in other branches of organizational studies, a fact my good friend Michael Frese bluntly pointed out (thereby saving me from embarrassment in the psychology camp) when I was drafting Davidsson (2006). So readers should not rely solely on me but also consult the classics on this issue (Rousseau, 1985; K. Klein & Kozlowski, 2000) as well as recent, "non-entrepreneurship" sources (e.g., Kozlowski, Chao, Grand, Braun, & Kuljanin, 2013).

At any rate, levels are important, and studies should be clear about their level of analysis. An industry-level study may show that innovative industries have higher levels of profit. This does not necessarily mean that it is the innovative firms in such industries that enjoy the higher profitability; the profits could equally well be captured by smart imitators who have lower development costs and who are perhaps more business savvy. Conversely, if firm-level studies show a positive relationship between growth and profitability, this *may* actually be an industry-level effect. That is, firms in growing industries may enjoy above-average growth and above-average profitability because of the growth of the industry and not because firm growth makes firms more profitable[9]. Likewise, a regional or national-level study may show that on that level, the prevalence of certain values or attitudes is associated with higher rates of new firm formation. This does not necessarily mean that it is individuals who hold such values or attitudes that are particularly prone to start new businesses. Alternatively, it may be that being *surrounded* by such values or attitudes make anyone more likely to strike out on their own, regardless of their own psychological makeup (cf. Davidsson, 1995a).

One major project I am currently involved in is the Comprehensive Australian Study of Entrepreneurial Emergence, CAUSEE (Davidsson & Steffens, 2011; Davidsson, Steffens, & Gordon, 2011). This is a PSED-like panel study that follows the development of nascent (and young) ventures over time. When designing this project, I tried to put some of what I preach about levels to practice. First, we made a decision that the main level of analysis was that of the emerging *venture* itself (cf. Chap. 2). Although we only had one respondent per venture, we regarded this respondent as a spokesperson for the venture and for the entire start-up team (when applicable). This had effects both on what questions we asked and how we asked them[10]. One insight was that we would to a considerable extent have to give up on psychological characteristics, because the respondent could hardly be expected to correctly report the *neuroticism* (and what have you) of other team members. Second, we built some hard-earned lessons about cross-levels issues into the design. Figure 4.4 exemplifies one such instance.

Quite a lot of early entrepreneurship research implicitly employed a "one individual = one venture" design assumption. As depicted in Fig. 4.4 (and supported by research on team entrepreneurship and habitual entrepreneurs; Ruef, Aldrich, & Carter, 2003; Ucbasaran, Westhead, & Wright, 2006), this is very often not the case. If one applies the "one individual = one venture" assumption in the design, the influence of other team members and/or the distraction (or supplementary resources) provided by the founders' other business activities become *Z* factors (in Fig. 4.1) which distort the results[11]. This design flaw may explain, for example, why the estimated average effect of "human capital" on outcomes in entrepreneurship research

[9] This is exactly what the review by Capon, Farley, and Hoenig (1990) suggests, if you study their results carefully. See also Davidsson, Steffens, and Fitzsimmons (2009). In terms of Fig. 4.1, the researchers may think they have found an $X \rightarrow Y$ relationship on the firm level. However, the real driver of firm-level (growth and) profitability is characteristics of the industries. This calls for a cross-level design, or at least that the researchers control for industry-level effects, which otherwise end up being biasing *Z* factors.

[10] For one thing, once a firm was identified as a team start-up, all applicable skip patterns and question wording reflected this knowledge. Since we also adapted wording to what type of venture (e.g., product vs. service based) we were dealing with, it made for some interesting complexity in programming the interviewing, and to questionnaires printing out to some 60–100 pages per interview wave. Really nerdy readers can study this in detail at http://eprints.qut.edu.au/49327/.

[11] Note that the arrows in Fig. 4.4 form a "Z"—how clever is that????

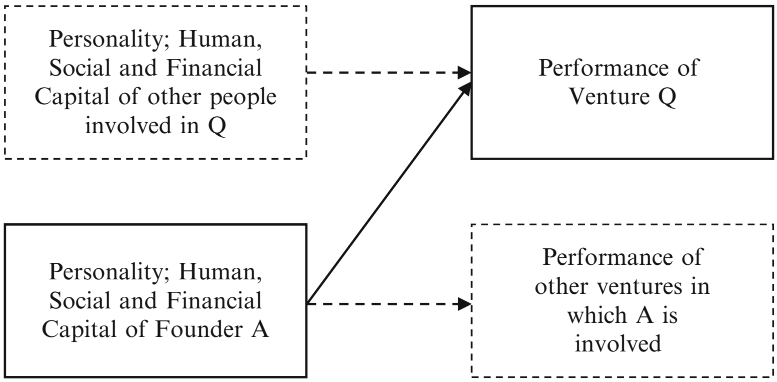


Fig. 4.4 A design fallacy in entrepreneurship research

is so surprisingly weak (Unger, Rauch, Frese, & Rosenbusch, 2009). In some studies, the effect is likely to be under estimated because the human capital provided by other team members was not considered (cf. Dimov, 2010, who chose to include only solo founders for this very reason). In other studies the possible effect does not materialize because portfolio entrepreneurs "invest" most of their human capital in other ventures than the sampled one.

The questionable but recurring "design favorite" to use (only) individual level *X* and venture/firm level *Y* variables is a peculiar type of "upward cross-level" design that does not even form part of the multilevel design experts' standard repertoire (Rousseau, 1985; Klein & Kozlowski, 2000). It is probably only researchers in entrepreneurship and leadership who entertain such a heroic view of single individuals (well, I would give Steve Jobs a nod, but otherwise…). If you reflect on it, it is rather optimistic to think that a few psychological and socio-demographic characteristics of the founder would explain much of the action patterns and outcomes of venture creation processes when there are so many factors pertaining to the venture itself as well as industry, regional, and macroeconomic environments (not to mention a lot of idiosyncratic, situational happenstance) that may also be influential. This brings us to our next topic: how to deal with all those factors (*Z*) that we aren't really interested in but which nevertheless influence the phenomenon under study.

4.6 Dealing with Heterogeneity in Design[12]

In Chap. 1, I highlighted heterogeneity as a fundamental characteristic of entrepreneurship and something to be embraced (or at least acknowledged). However, we cannot deal effectively with all possible heterogeneity at once. The heterogeneity

[12] This section borrows from a deceased paper from a few years back (a development of Davidsson, 2008) which I coauthored with Frédéric Delmar, and it may contain specific words that are his. The paper was rejected in round 2 in part because the special issue action editor did not like us

that is not in our theoretical focus instead becomes a method problem and a threat to the validity of our findings. Below I discuss three types of heterogeneity problems:

1. *The problem of unobserved heterogeneity* (e.g., Shugan, 2006). Also discussed under labels such as *omitted variable bias* (L. Lee, 1982) or *confounding variables* (Kish, 1987), this is the central heterogeneity problem that if Z (see Fig. 4.1) has substantial correlations with both x' and y', excluding Z will lead to serious bias in the coefficients.
2. *The problem of causal heterogeneity* (e.g., Western, 1998). This concept denotes the problem that the sign, magnitude, or form of the effect of X on Y may not be uniform across elements or subgroups of the studied population. In terms of Fig. 4.1, the arrow is a bit unstable or fuzzy. For example, the effect of some personality characteristic on the propensity to engage in new ventures may be different by biological sex of the founder(s), and the effect of some actions (e.g., business planning) on venture creation outcomes may be different depending on the type of entrepreneur (novice/expert) or the type of venture (imitative/innovative; however, see Garonne, 2014).
3. *The problem of uneven validity*. This is a special aspect of the x–x' and y–y' relationships in Fig. 4.1. Different aspects of this problem are highlighted particularly in cross-cultural research under labels such as *construct equivalence*, *instrument equivalence*, *measurement equivalence*, and *measurement invariance* (e.g., Byrne & Watkins, 2003; Schaffer & Riordan, 2003). This problem means that the chosen operationalizations (our topic for Chap. 6) are not equally suitable for all subgroups of the population. As a result, $X \rightarrow Y$ relationships will be misestimated. For example, if respondents at different level of educational attainment interpret the questions or response alternatives differently, we may get the result that the effect of X on Y is stronger for highly educated people when in actual fact what the results show is that the quality of the x' measure is better for this group.

So what can we do, design-wise, to eliminate or at least reduce the influence of these problems? Here are some standard suggestions:

Identify the Z factors and include them in the design. For example, in the CAUSEE study, we devote the entire first section of the interview to classifying the cases according to industry, team vs. solo, service vs. product, high tech vs. low tech, male vs. female, experienced vs. novice founder, etc. This strategy is sound but can only take us so far, because if the number of Z variables is large, modeling

to confuse the field with the new and unknown term "heterogeneity" ("unmeasured heterogeneity" = 7,770/1,870 Google/Google Scholar hits; "causal heterogeneity" = 8,400/2,010 Google/Google Scholar hits).

their influence in the analysis may be overwhelming if they are also a source of causal heterogeneity. For example, simply including a control variable for sex in the analysis does not solve the problem if the *effect* of *X* is different by sex. We will elaborate on a systematic procedure for identifying *Z* variables further below.

Use a narrowly defined sample. For example, in the *Culture and Entrepreneurship* study (e.g., Davidsson, 1995a), I needed to obtain primary regional data on "cultural values" from samples of individuals. In order for observed region differences not to be driven by random differences in the age and sex of the respondents, I sampled narrow age cohorts (18–19-year olds and 35–36-year olds) with equal male and female representation. Turning to greater authorities, you may note that studies published in top journals often study a narrow empirical context: one industry, a cohort of firms of similar age, and solo start-ups only. They do not try to obtain a statistically representative sample of every available empirical context that they may be interested in or for which they suspect their theory is valid. Instead they study one (or a few) context(s) they deem theoretically relevant and try to do that really well. This has several advantages. First, it keeps some *Z* constant and reduces the variability in others (Problem 1 above). Second, it also reduces the risk of severe causal heterogeneity (Problem 2). For example, if a particular *X* has a negative effect among women and an equally strong positive effect among men, a female only sample will reveal the negative effect whereas a mixed sample could conceal both effects completely. Third, narrow sampling reduces the risk of uneven validity (Problem 3), and as a bonus it allows customized (e.g., industry-specific) operationalizations which may have higher measurement validity in the specific context than would a more generic measure (cf. the measure of organizational innovation used by Cliff, Jennings, & Greenwood, 2006).

Hold (*many*) *Z variables constant*. Narrow sampling achieves this to an extent. In experimental and simulation studies that test relationships in an artificial, researcher-controlled situation, this strategy can be taken much further. For example, in real-life decision-making, any number of criteria may be used and used differently by different individuals. By contrast, in conjoint experiments (Lohrke et al., 2010) it is the experimenter who decides which factors differ across alternatives and by how much. Everything else about the choice alternatives is assumed to be equal.

Randomize the influence of Z. This is the other trick employed in experimentation. Participants in experiments are randomly assigned to the various experimental conditions. This is to try to make sure that variance in the experimental variable is not highly correlated with some person factor (e.g., age, sex, education, looks, mental stability, experience; preferred James Bond actor…) that may also influence the experiment's dependent variable. This is almost certain to work in the long run. However, if there are many *Z* variables and the number of participants in each treatment group is small, we *may* of course still get most of the religious fanatics in one group and all the closet superheroes in another, with unknown consequences.

Match cases on important Z variables. For example, in the second stage of the *Culture and Entrepreneurship* study (Davidsson & Wiklund, 1997), I wanted to isolate the influence of cultural variation on regional-level entrepreneurship, after already having established the importance of a number of structural regional variables (Davidsson, Lindmark, & Olofsson, 1994a). Unable to collect individual-level primary data on values and beliefs (to be aggregated to the region level) for 100+ regions, I identified pairs of regions that were structurally similar, but where one region in each pair had higher and the other lower start-up rate than what would be predicted based on their structural characteristics. By matching on structural variables, the design gave at least some chance to isolate cultural influences using as few as six regions in the study. More on this in Chap. 5. A drawback with direct matching is that one can only match on one or a few characteristics. Recently, researchers have started to use *propensity score matching* in order to approach matching across a larger number of potential *Z* variables (e.g., Elert, Andersson, & Wennberg, 2014).

See, there are some options! We do not have to be overwhelmed by *Z* problems at the analysis stage if we pay some attention to this issue when designing the study.

Now back to the question of how to identify potential *Z* variables. As a starting point, we can imagine a theory suggesting a positive relationship between unemployment (X) and business start-ups (Y) because those faced with unemployment have reason to seek alternative ways to provide for themselves. The simple matrix in Fig. 4.5 can help us identify *Z* variables. In this figure, Δ denotes change. Quadrants I and III represent the cases that accord with the suggested explanation of variance in business start-up rates or inclination: when *X* (employment status) changes, *Y* (probability of creating a start-up) changes as well (Q1). When there is no change in *X*, no change in *Y* ensues (QIII).

Quadrants II and IV constitute counterfactual cases. In quadrant II the question is under what conditions the proposed relationship might not hold. This is a question to ask in the pursuit of *Z* suspects. For example, in a society or social stratum where most individuals are affluent by birth, the construct "unemployed" would not be equivalent with the same notion in mainstream societies. Where the institutional framework includes generous unemployment benefits and/or high bureaucratic or cultural barriers to firm formation, the relationship could also be weak or nonexistent. Similarly, for people close to retirement age, the response to layoffs may more rarely be to set up their own business. Importantly, if the theoretically focused

Fig. 4.5 A simple schema for identifying heterogeneity issues in design

	I	II
ΔX	ΔY	$\sim\Delta Y$
	IV	III
$\sim\Delta X$	ΔY	$\sim\Delta Y$

variable is correlated with another variable that has a negative effect on *Y*, the positive effect of *X* will not necessarily appear in empirical estimation. This would be the case here if the economic conditions that increase unemployment and therefore increase "necessity-based" entrepreneurship at the same time reduce "opportunity-based" entrepreneurship via decreased market demand (cf. Hamilton, 1989; Wennekers, Stel, Thurik, & Reynolds, 2005). It should be clear from the example that the counterfactual mental gymnastics has potential for informing our design.

Quadrant IV depicts the other counterfactual case. Where, when, and for whom might business start-ups occur for reasons other than unemployment? Here, the candidates are many, and those that are correlated (negatively or positively) with the risk of unemployment are particularly important to identify. Obviously, like in most social research, we are dealing with a phenomenon that has many possible causes. This is the root of the problem of unobserved heterogeneity. Therefore, the essence of systematic search for other causes is to find those correlated variables that either need to be included in the empirical design or made uncorrelated with *X* (or, more correctly, with x') via constant holding, randomization, or matching.

4.7 Design Pros and Cons of Different Types of Data

4.7.1 Primary Survey Data

The biggest allure of questionnaire-based surveys is the great versatility and potential for customization that they offer. Applying good sampling and operationalization strategies, in terms of Fig. 4.1, the survey designer can include a comprehensive set of well-operationalized variables x' and just the right y' variable(s) for the purpose, while controlling for all known *Z* through sample restriction (constant holding) and inclusion of control variables. These attractive features have long made survey research the clear "market leader" in entrepreneurship research.

This said, there are issues with survey research. Tell me about it; from my dissertation study to our current work on CAUSEE, I've probably done more survey research—including the frustrating mistakes that come with it—than most of my colleagues. In fact, out of my ten best-cited empirical works, nine are survey based, using data from seven different survey studies (Brown, Davidsson, & Wiklund, 2001; Davidsson, 1989a, 1989b, 1991, 1995a, 1995b; Davidsson & Honig, 2003; Delmar & Davidsson, 2000; Wiklund, Davidsson, & Delmar, 2003). So despite what I say below, it isn't like I or the field of entrepreneurship have turned our backs on survey research.

The big survey research issue that isn't much talked about is measurement validity. No, I don't mean presenting an acceptable Cronbach's alpha or performing some lame test to indicate the absence of common-method bias; I mean *VALIDITY*: does the paper-and-pencil (or mouse-clicking) behavior of survey respondents adequately capture their actual behavior (or characteristics) in the real world? That's a big leap of faith. We'll have more to say about that in Chap. 6.

There are other issues as well. Like any form of observational data, survey data reflect a complex reality with "too many moving parts," making it difficult to find the correct model and to prove causality. Technological development has made it easy to pull together and distribute a survey online, but our diverse communication habits and spamming (in a broad sense) have made it harder to find a representative sampling frame and to get the people in it to participate in the survey. Initial non-response plus attrition in later waves may make samples problematically small and nowhere near randomly selected, thus invalidating all the cherished significance tests. As a consequence of these shortcomings, survey research has recently become less popular as it has become increasingly longitudinal (Crook et al., 2010). However, the relative decline also reflects increases in the availability of archival data. It is mostly good development, because research in both categories is improving.

For entrepreneurship research, a particular challenge is that there are no sampling frames that cover hitherto unexploited "opportunities" or individuals that are in the process of creating new economic activities. However, regarding the latter, Paul Reynolds has made an enormous contribution by creating the sampling methodology used in GEM and PSED-type research (Davidsson, 2005; Reynolds, 2009). Future survey research can extend this great contribution by applying the PSED idea of early-capture-then-follow-over-time to more homogeneous, higher potential samples as well as to corporate venturing.

Some apparent shortcomings of survey research are not inherent to the approach but a matter of poor execution. If you want to find bad examples, look for poorly defined populations, ridiculously low response rates, cross-sectional design with data from a single respondent, an ad hoc home-brew or theoretical constructs and operationalizations, and dull, descriptive reporting of results.

Then again, these shortcomings in execution can be seen as a response to another issue with survey research: doing good survey research is costly in terms of both time and money. By and large, you won't get away with cross-sectional work anymore in higher-tier journals, and the time cost of prospective, longitudinal work may be unbearable for dissertation projects. So, what to do? One strategy is to tap into someone else's ongoing survey panel (which makes the data archival/secondary from your point of view, unless you can influence the design). Another is to focus on either *X* or *Y* rather than the relationship between them. Test the true validity of some of our favorite measures and/or develop new and better alternative operationalizations. Take on the conceptual and operationalization task of assessing outcomes for not-yet-operational businesses (cf. Chap. 7). In response to calls for sensitivity to context (Welter, 2011; Zahra & Wright, 2011) develop customized and "locally" better operationalizations of core constructs. There are many options (you know I can't say "opportunities").

One of my favorites using survey research is (still) Baum and Locke's (2004) study of psychological determinants of small-firm growth. They elegantly design away many of the *Z* that may hide the psychological influence by selecting one, narrowly defined industry: North American architectural woodwork firms. Further, they have enough of a time lag—6 years—for it to be reasonable to assume these psychological effects to show their impact. At least for its time, the operationalizations were generally superior to the contemporary standard, and the authors also developed a believable model of direct and indirect relationships. Finally, the research was succinctly reported and published in a highly respected outlet.

4.7.2 Archival or Secondary Data

This refers to any "already available" data that someone else has collected for purposes other than your academic research. This makes it a heterogeneous beast ranging from large registers of official business statistics to verbal data in annual reports and all the electronic traces we leave and others store for whatever Orwellian purpose or use. As such, it is a category about which it is somewhat difficult to make valid generalizations.

Archival data can have several strengths. One is that you may get historical, longitudinal data in an instance and that these data were collected at the time rather than retrospectively, as is often the case in questionnaires and interviews. The data may also have been collected in an unobtrusive manner, thus removing concerns about social desirability or impression management biasing the results (Davis, Thake, & Vilhena, 2010). Further, working with archival or "secondary" data often means working with data from an entire population or very large samples. This means that your conclusions will likely not be the fruit of stochastic variation and that you can use better truth criteria and stronger truth claims than what the contentious tool of statistical significance testing can offer (cf. Chap. 9).

Archival data are sometimes portrayed as cheap to get and easy to work with because someone else has already prepared them for you. In my experience, this is generally not the case with archival data as used in high-quality, scholarly research. Instead, you may have to work hard and/or pay significant amounts of money to check the quality of the data and to get them reorganized into analyzable form. Further, it is often necessary to combine data from several sources in order to achieve sufficient quality and completeness of the data. For example, this is usually the case with studies using GEM data—these data do not cover enough bases, so researchers need to combine them with other sources of "soft" and "hard" country-level data. Doing good research based on archival data is not about taking a data set off the shelf, run a few regressions, and send off to publication. It is hard work—as it should be. In the first edition of this book, I devoted two entire chapters (Chaps. 7 and 8) detailing what it entails to work with "secondary" data sets, so you can go there for more detail if you so wish.

The main issue with less satisfactory research based on archival data is that the data—which were collected for other purposes—simply cannot do the job. In terms of Fig. 4.1, the most interesting *X* variables may not be there, making researchers look at less interesting relationships or use rather distant "proxy variables," which in plain English means indicators that probably have very low validity. If the interesting *X* variables exist and are reasonably well operationalized, the problem may be on the *Y* side; there are no measures of the most important outcomes. Alternatively, the indicators are of questionable quality or assessed at the wrong time. If both *X* and *Y* are reasonably well represented, you can be almost sure that your main source of data will not contain information on some important *Z* variables, thus making reviewers and editors suspicious that your results misrepresent the true relationships. Alternatively, the database covers the wrong industry, as when Davis, Haltiwanger, and Schuh (1996) used very detailed analyses of the shrinking manufacturing industry in order to make claims about the dynamics of job creation in a country that had added dozens of millions of jobs during the studied period. Another sad example I related in the first edition was an effort back in the 1990s toward a harmonized European study of "high-growth firms" where the smallest common denominator was to use only manufacturing firms that were at least 10 years old and which had at least 20 employees. This effectively means removing all those categories of firm where most high growth occurred at the time: young, small, and service-producing firms. These examples represent quite massive misalignment between "Our Study" and "The World About Which We Wish to Know and Tell." Both examples remind of the old joke about the very drunk man who is seen rummaging about in the snow under the streetlight late at night. Another man approaches him and the following conversation takes place:

> Excuse me, can I help you? What are you looking for?
> I'm…schersching for my (hic) … keysch.
> Your keys? And you're sure this is where you dropped them?
> No (hic) …not at all. I'm schure I dropped them over there (the drunk man says, pointing into the darkness, almost falling over backwards).
> Over there? But why then are you looking for them over here!?
> Well…ischn't that obviousch? Over there it'sch scho damn dark you can't posschibly find anything…

As researchers, we do not want to be like this drunkard, right? Archival data are, to some extent, like streetlights. They are put there for general purposes or for some other purpose than yours. They do illuminate some area, but they do not necessarily cast light on the issues you are interested in. If the data cannot possibly answer your research questions—as is often the case in entrepreneurship since emerging phenomena usually do not appear in registers—don´t try!

Against this background, it is not necessarily a good thing that the use of archival data is increasing in entrepreneurship research (Crook et al., 2010). I found that there is a strong development in that direction in the stream of research using "opportunity" in title, abstract, or keywords (Davidsson, 2015). Which makes me wonder what archival data sets contain rich information on such elusive, early-stage phenomena as "opportunities"?

This said, the last decade has seen some fantastic developments regarding archival or "secondary" data for entrepreneurship research. This is the more positive reason for its increased use in published research. First, there are now rich, customized data sets in the public domain, specifically designed for entrepreneurship research. These include two initiatives, each of which has been the basis for more than 100 published research articles: GEM (Bergmann, Mueller, & Schrettle, 2014) and PSED and its counterpart studies (Davidsson & Gordon, 2012). Another example is the Kauffman Firm Survey (DesRoches et al., 2007). Second, especially in Europe, there are several examples of high-quality, multilevel data sets that can be used for entrepreneurship research purposes (Baptista & Mendonça, 2010; Campbell, 2005; Campbell, Ganco, Franco, & Agarwal, 2012; Coad, Frankish, Roberts, & Storey, 2013; Hellerstedt, 2009). Third, technological development has led to massive amounts of electronic traces being left everywhere, including those that can be used for innovative entrepreneurship research by those who show enough creativity as well as an ability to gain access.

On the positive side, I should also mention that some of my fondest and proudest moments as a researcher are associated with research on "secondary" data sets. This includes participation in a seven country, harmonized study of regional drivers of business start-ups, which arrived at comparatively strong results (Davidsson, Lindmark, & Olofsson, 1994b; Reynolds, Storey, & Westhead, 1994), enlightening the "small-firm job creation" debate and influencing the US Bureau of Labor Statistics to change its reporting standards (Butani et al., 2005; Davidsson, Lindmark, & Olofsson, 1998; de Wit & de Kok, 2013), developing a methodology for distinguishing between organic and acquisition-based growth and finding strong results related to this distinction (Davidsson & Delmar, 2006; Delmar et al., 2003; Lockett, Wiklund, Davidsson, & Girma, 2011), and providing some provocative evidence on the relationship between growth and profitability (Davidsson et al., 2009). Presently, I'm working with Rene Bakker and Dean Shepherd on a fantastic, longitudinal, multilevel and (in the context of scholarly research) virgin data set Rene dug up shortly after landing in Australia. In line with my previous experiences, in Sweden it took quite a bit of hard work to make the data set analyzable and useful.

A favorite in the archival genre is the study by Pe'er and Vertinsky (2008). This study puts a twist on Schumpeter's (1934) notion of "creative destruction" by arguing the importance of exits to make room for new start-ups. Without exit there aren't any idle resources to employ at reasonable cost. The authors show due care in developing and checking the data set and employ sophisticated analyses to empirically support their case through time and space.

4.7.3 Laboratory Research and Field Experiments

The share of published entrepreneurship research that is experimental is small but increasing (Crook et al., 2010)[13]. In some areas, experiments are prevalent. For example, my recent review of research on "entrepreneurial opportunities" showed that more than 50 % of the small stream that explicitly addresses Shane and

[13] The share was twice as high in 2005–2007 compared to 2000–2002. The authors report this doubling of the frequency as "not significant." However, the doubling is a fact about the population of articles they investigated, which was not a random sample from a larger population (cf. Chap. 9).

Venkataraman's "individual-opportunity nexus" idea is experimental (Davidsson, 2015). Laboratory research also includes approaches such as think-aloud protocol analysis (Grégoire, Barr, & Shepherd, 2010) and simulations. The latter come in two varieties: those that only involve a computer (and a programmer) and those where human participants interact with a simulated world, similar to computer games. The use of the former type—especially agent-based simulation (ABS)—has recently been propagated by influential management and entrepreneurship scholars (Davis, Eisenhardt, & Bingham, 2007; McKelvey, 2004). However, although there are published examples (Albino, Carbonara, & Giannoccaro, 2006; Wu, Kefan, Hua, Shi, & Olson, 2010), at this point it is probably fair to say that we are still waiting for the first truly influential simulation studies in entrepreneurship.

The first great strength of laboratory research is *researcher control*. The artificial situation allows the researcher to manipulate x' and measure a usually rather unambiguous y' (which clearly happens after x'), while Z is a complete nonissue (some types of simulation) or can be effectively dealt with through randomization, matching, and control variables, as discussed earlier in the chapter. This means that laboratory research is ideal for *theory testing* with the researcher having a much stronger basis for claiming *causality* compared to other approaches. The advantages do not stop there. With multi-period experiments and simulations, we can do longitudinal research in close to no time and often at much lower cost than when collecting real-world data. That's a great advantage for dissertation projects with a time component! Further, statistical significance is much less problematic in experimentation. The logic of statistical testing in experiments builds on *random assignment*—ascertaining that the allocation of participants to different experimental treatments is random—rather than *random sampling*. The inference is drawn not to an external, larger population but to all possible ways the participants could have been allocated to treatments. So the question we are testing is "Could we have been so unlucky that the result is due to a skew allocation of participants rather than to our manipulation of the X?"

It's not all rosy, though. The BIG limitation of laboratory research is *external validity*. Does what happens in the laboratory have much bearing at all on entrepreneurial processes in the real world? In particular, laboratory research cannot easily invoke the psychological reality of making high-stakes decisions. In addition, participants in experiments are typically volunteers rather than a random sample, adding to the questions about generalizability. This calls for more *field experiments* or *randomized trials*, i.e., experimentally controlled, real-world initiatives. And this is precisely what is now under way, for example, through the randomization of participation in GATE and other policy initiatives (Fairlie, Karlan, & Zinman, 2015) and the Innovation Growth Lab (www.innovationgrowthlab.org/innovation-growth-lab). Interesting developments, indeed!

Whether inside or outside the laboratory, however, experiments cannot cover much complexity, usually only a couple of X variables and their interaction. What minuscule fraction of real-world variance do the factors the experiment hones in on explain in the real world? Conjoint experiments (Lohrke et al., 2010) are somewhat more allowing in terms of number of manipulated variables, but this approach usually also makes the experimental situation an even more artificial and hypothetical matter of "paper-and-pencil" reactions or decisions. "Computer game" or "microworld"-type simulations and "think-aloud" exercises can capture more complexity but only at the cost of reducing researcher control and the unambiguous interpretation of causality that goes with it. Further, I mentioned covering many periods without needing much clock time as a strong point, but the caveat is that multi-period laboratory work may not capture the true working of time. Moreover, an especially problematic issue from an entrepreneurship point of view is that the laboratory approaches may not be able to capture creative behavior that goes beyond what the researcher was able to imagine when designing the research.

This said, the history of science is full of examples of researchers who have reached important new insights not only from experiments that "worked" but also from those that have "gone wrong," suggesting the possibility of insights beyond the limits of the creativity that went into the design. Further, agent-based simulation is argued to be an excellent tool for theory building rather than theory testing, especially as regards how macrolevel phenomena arise from microlevel behavior. Apparently, despite the simulation being a great simplification of reality and delimited by the designers' cognitive capacity, the results can often be both surprising and greatly illuminating. This may concern what microlevel behavioral rules are required to create known, aggregate level phenomena or what aggregate level changes would follow from seemingly small microlevel variations.

My own experience with laboratory research is less comprehensive than my experience of research based on observational data. However, it includes what might be the first application of conjoint analysis in the field of entrepreneurship (Davidsson, 1986; I was about 6 years old at the time, honestly!) as well as a very recent one (Steffens, Weeks, Davidsson, & Isaak, 2014). I have also worked with doctoral students who used experiments as main or supplementary method in their investigation of entrepreneurial expertise and decision-making (Gustafsson, 2004), celebrity entrepreneurship (Hunter, 2009), and determinants of the attractiveness of new venture ideas (Semasinghe, 2011). My current favorite example of a laboratory research is Denis Grégoire's stream on structural vs. superficial alignment in "opportunity identification" (Grégoire & Shepherd, 2012; Grégoire, Shepherd & Schurer Lambert, 2010; Grégoire et al., 2010). This research creatively builds on predecessors, develops nonobvious theory with practical potential, uses real-world participants and technologies, and is published in very good outlets.

4.7.4 Qualitative Process Research

I have already discussed qualitative vs. quantitative earlier in the chapter, and I think I may also have promised a couple of times that because I'm massively incompetent, I will shut up on qualitative research approaches. But obviously I cannot help myself! I just wanted to add the observation that there are signs of increasing appreciation of qualitative research among influential "mainstream" authors (McMullen & Dimov, 2013; Zahra & Wright, 2011). There is also increasing consensus that entrepreneurship is about *behaviors* in the *process* of new economic activity. This would seem to make *ethnography* a suitable method. Ethnographic case studies are longitudinal and have the capacity to capture processes. Accordingly, there have been calls (e.g., Aldrich & Baker, 1997) for intense study of "super-entrepreneurs" of the type that Mintzberg (1974) did on managers.

Although such an approach may seem interesting, there is a catch, which derives from our choice of perspective on entrepreneurship in Chap. 1. An intense but short in duration study like Mintzberg's would likely not be very illuminating, even if focused on someone well known to be a habitual and repeatedly highly innovative and successful entrepreneur. This is because the researcher would have to be extremely lucky or persistent in order to capture much behavior that is uniquely entrepreneurial. Arguably, managers perform managerial functions most of their working time. Not so with entrepreneurs; the "entrepreneurial function" is much less of the nature that it is constantly switched on. Further, a weeklong study would at best cover just a little glimpse of the entrepreneurial process from idea to successfully established venture.

In addition, in my experience the most fruitful contributions from qualitative entrepreneurship research focus on a few issues across a somewhat larger number of cases, rather than offering a very intense study and holistic understanding of very few cases. A current favorite is Fauchart and Gruber (2011) although I am old enough to remember that this is far from the first contribution on entrepreneur types and identities (Smith, 1967; Stanworth & Curran, 1973; Woo, Cooper, & Dunkelberg, 1991). Fauchart and Gruber study cases in a single industry, thus removing theoretically irrelevant noise (Z variables). Their work comes across as a really solid piece of qualitative work, and they include one little detail—a simple, quantitative analysis that indicates a strong relationship between type/identity on the one hand and performance on the other—that makes the whole exercise much more interesting to readers of my ilk. Recent developments make it possible to combine some of the rationale behind the intense, ethnographic study with the possibility of following a larger number—perhaps a couple of dozen—and the necessity of doing so over a longer period of time. The possibility I have in mind is the Experience-Sampling Method (ESM) which Uy et al. (2010) recently introduced to entrepreneurship research. I look forward to seeing more creative applications of this approach to collecting small snippets of data at frequent intervals through a device most of us never stray far from: the (smart) mobile phone.

4.8 Summary and Conclusion

The readers of this new edition have been punished with a much longer, but also hopefully much better, Chap. 4. I'm reasonably happy with the distinction between "The World About Which We Wish to Know and Tell" on the one hand and "Our Study" on the other, using this to link a range of design concepts and as the "glue" for the reminder of the chapter. I have argued that knowledge development in entrepreneurship benefits from different types of research—"qualitative" as well as "quantitative" and laboratory research as well as studies that rely on data from real settings. Preferably, these different types of research should be combined in comprehensive programs. At least, it would be to the advantage of knowledge development if the different forms of research informed and inspired one another, rather than different methodological camps of entrepreneurship researchers developing separate and noncommunicating discourses (van Burg & Romme, 2014).

I have further argued that entrepreneurship research can and should be conducted on different levels of analysis. However, in order to qualify as entrepreneurship research, the study has to take new venturing on the studied level into explicit consideration. In empirical entrepreneurship research, the focal phenomenon should not be reduced to an assumption. Regardless of the level of analysis chosen, it is important that it be properly matched with the theory, as discussed in the previous chapter.

We have also discussed implications of the process nature of entrepreneurship and ways of dealing with heterogeneity that otherwise threatens to blur our findings. Finally, I commented on trends, strengths, and weaknesses of different data approaches in relation to design issues. There is no such thing as "perfect research," but we can get better at matching the approach with the research problem and—given the approach chosen— using the best methods knowledge available in order to make the most of our research efforts.

References

Ahl, H. J. (2002). *The making of the female entrepreneur*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Ahl, H. J. (2006). Why research on women entrepreneurs needs new directions. *Entrepreneurship: Theory and Practice, 30*(5), 595–621.

Albino, V., Carbonara, N., & Giannoccaro, I. (2006). Innovation in industrial districts: An agent-based simulation model. *International Journal of Production Economics, 104*(1), 30–45.

Aldrich, H. E., & Baker, T. (1997). Blinded by the cites? Has there been progress in the entrepreneurship field? In D. Sexton & R. Smilor (Eds.), *Entrepreneurship 2000* (pp. 377–400). Chicago, IL: Upstart Publishing Company.

Amit, R., & Zott, C. (2001). Value drivers in e-business. *Strategic Management Journal, 22*, 493–520.

Amorós, J. E., Bosma, N., & Levie, J. (2013). Ten years of global entrepreneurship monitor: Accomplishments and prospects. *International Journal of Entrepreneurial Venturing, 5*(2), 120–152.

Anderson, J. R. (1990). *Cognitive psychology and its implications*. New York: W.H. Freeman and Co.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329–366.

Baptista, R., & Mendonça, J. (2010). Proximity to knowledge sources and the location of knowledge-based start-ups. *The Annals of Regional Science, 45*(1), 5–29.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173.

Baum, J. R., & Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology, 89*(4), 587–598.

Bergmann, H., Mueller, S., & Schrettle, T. (2014). The use of Global Entrepreneurship Monitor data in academic research: A critical inventory and future potentials. *International Journal of Entrepreneurial Venturing, 6*(3), 242–276.

Bhave, M. P. (1994). A process model of entrepreneurial venture creation. *Journal of Business Venturing, 9*, 223–242.

Björnsson, B. (2001). *En bank även för kvinnor? Småföretagares erfarenheter av rådgivningsmöten i bank (A bank also for women? Small-business customers' perceptions of counselling encounters in banks).* Doctoral dissertation. Göteborg University, Göteborg.

Brundin, E. (2002). *Emotions in motion. Leadership during radical change*. Doctoral dissertation, Jönköping International Business School, Jönköping.

Bruns, V. (2004). *Who receives bank loans? A study of lending officers' assessments of loans to growing small and medium-sized enterprises*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Butani, J., Clayton, R. L., Kapani, V., Spletzer, J. R., Talan, D. M., & Werking, G. S. J. (2005). *Business employment dynamics: Tabulations by employer size.* US Bureau of Labor Statistics, Working Paper 385.

Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology, 34*(2), 155–175.

Campbell, B. A. (2005). *Using linked employer-employee data to study entrepreneurship issues* (Handbook of Entrepreneurship Research, pp. 143–166). Berlin: Springer.

Campbell, B. A., Ganco, M., Franco, A. M., & Agarwal, R. (2012). Who leaves, where to, and why worry? Employee mobility, entrepreneurship and effects on source firm performance. *Strategic Management Journal, 33*(1), 65–87.

Capon, N., Farley, J. U., & Hoenig, S. (1990). Determinants of financial performance: a meta-analysis. *Management Science, 36*(10), 1143–1159.

Carter, N. M., Gartner, W. B., & Reynolds, P. D. (1996). Exploring start-up event sequences. *Journal of Business Venturing, 11*, 151–166.

Chandler, G. N., & Lyon, D. W. (2001). Methodological issues in entrepreneurship research: The past decade. *Entrepreneurship: Theory and Practice, 25*(4), 101–113.

Cialdini, R. B. (1980). Full cycle social psychology. In L. Beckman (Ed.), *Applied social psychology annual* (Vol. 1). Beverly Hills, CA: Sage.

Cliff, J. E., Jennings, D. P., & Greenwood, R. (2006). New to the game and questioning the rules: The experiences and beliefs of founders who start imitative versus innovative firms. *Journal of Business Venturing, 21*, 633–663.

Coad, A., Frankish, J., Roberts, R. G., & Storey, D. J. (2013). Growth paths and survival chances: An application of Gambler's Ruin theory. *Journal of Business Venturing, 28*(5), 615–632.

Cohen, J. (1990). Things I have learned (so far). *American Psychologist, 45*(12), 1304–1312.

Crawford, G. C., Aguinis, H., Lichtenstein, B., Davidsson, P., & McKelvey, B. (2015). Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.001.

Crook, T. R., Shook, C. L., Morris, M. L., & Madden, T. M. (2010). Are we there yet? An assessment of research design and construct measurement practices in entrepreneurship research. *Organizational Research Methods, 13*(1), 192.

Dahlqvist, J. (2007). *Assessing new economic activity: Process and performance in new ventures*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Dahlqvist, J., & Wiklund, J. (2012). Measuring the market newness of new ventures. *Journal of Business Venturing, 27*(2), 185–196.

Davidsson, P. (1986). *Tillväxt i små företag: En pilotstudie om tillväxtvilja och tillväxtförutsättningar i små företag* (Small Firm Growth: A Pilot Study on Growth Willingness and Opportunity for Growth in Small Firms) (Studies in Economic Psychology No. 120). Stockholm: Stockholm School of Economics.

Davidsson, P. (1989a). *Continued entrepreneurship and small firm growth*. Doctoral dissertation, Stockholm School of Economics, Stockholm.

Davidsson, P. (1989b). Entrepreneurship—and after? A study of growth willingness in small firms. *Journal of Business Venturing, 4*(3), 211–226.

Davidsson, P. (1991). Continued entrepreneurship: Ability, need, and opportunity as determinants of small firm growth. *Journal of Business Venturing, 6*(6), 405–429.

Davidsson, P. (1995a). Culture, structure and regional levels of entrepreneurship. *Entrepreneurship & Regional Development, 7*, 41–62.

Davidsson, P. (1995b). *Determinants of entrepreneurial intentions working paper 1995:1*. Jönköping: Jönköping International Business School. http://eprints.qut.edu.au/2076/1/RENT_IX.pdf.

Davidsson, P. (2005). Paul Davidson Reynolds: Entrepreneurship research innovator, coordinator and disseminator. *Small Business Economics, 24*(4), 351–358.

Davidsson, P. (2006). Method challenges and opportunities in the psychological study of entrepreneurship. In J. R. Baum, M. Frese, & R. A. Baron (Eds.), *The psychology of entrepreneurship* (pp. 287–323). Mahway, NJ: Erlbaum.

Davidsson, P. (2008). Strategies for dealing with heterogeneity in entrepreneurship research. In P. Davidsson (Ed.), *The entrepreneurship research challenge* (pp. 103–124). Cheltenham, UK: Edward Elgar Publishing.

Davidsson, P. (2013). Some reflection on research 'Schools' and geographies. *Entrepreneurship & Regional Development, 25*(1-2), 100–110.

Davidsson, P. (2015). Entrepreneurial opportunities and the entrepreneurship nexus: A reconceptualization. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.002.

Davidsson, P., & Delmar, F. (2006). High-growth firms and their contribution to employment: The case of Sweden 1987–96. In P. Davidsson, F. Delmar, & J. Wiklund (Eds.), *Entrepreneurship and the growth of firms*. Cheltenham, UK: Edward Elgar Publishing.

Davidsson, P., Gordon, S. R., & Bergmann, H. (Eds.). (2011). *Nascent entrepreneurship*. Cheltenham, UK: Elgar.

Davidsson, P., & Gordon, S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics, 39*(4), 853–876.

Davidsson, P., & Honig, B. (2003). The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing, 18*(3), 301–331.

Davidsson, P., & Klofsten, M. (2003). The business platform: Developing an instrument to gauge and assist the development of young firms. *Journal of Small Business Management, 41*(1), 1–26.

Davidsson, P., Lindmark, L., & Olofsson, C. (1994a). *Dynamiken i svenskt näringsliv (Business Dynamics in Sweden)*. Lund, Sweden: Studentlitteratur.

Davidsson, P., Lindmark, L., & Olofsson, C. (1994b). New firm formation and regional development in Sweden. *Regional Studies, 28*, 395–410.

Davidsson, P., Lindmark, L., & Olofsson, C. (1998). The extent of overestimation of small firm job creation: An empirical examination of the 'regression bias'. *Small Business Economics, 10*, 87–100.

Davidsson, P., & Steffens, P. (2011). Comprehensive Australian Study of Entrepreneurial Emergence (CAUSEE): Project presentation and early results. In P. D. Reynolds & R. T. Curtin (Eds.), *Business creation panel studies: An international overview*. New York, NY: Springer.

Davidsson, P., Steffens, P., & Fitzsimmons, J. (2009). Growing profitable or growing from profits: Putting the horse in front of the cart? *Journal of Business Venturing, 24*(4), 388–406.

Davidsson, P., Steffens, P., & Gordon, S. R. (2011). *Comprehensive Australian Study of Entrepreneurial Emergence* (*CAUSEE*): *Design*, *data collection and sample description*. I. In K. Hindle & K. Klyver (Eds.), *Handbook of new venture creation research*. Cheltenham, UK: Edward Elgar Publishing.

Davidsson, P., & Wiklund, J. (1997). Values, beliefs and regional variations in new firm formation rates. *Journal of Economic Psychology, 18*, 179–199.

Davidsson, P., & Wiklund, J. (2001). Levels of analysis in entrepreneurship research: Current practice and suggestions for the future. *Entrepreneurship: Theory and Practice, 25*(4), 81–99.

Davis, C. G., Thake, J., & Vilhena, N. (2010). Social desirability biases in self-reported alcohol consumption and harms. *Addictive Behaviors, 35*(4), 302–311.

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. *Academy of Management Review, 32*(2), 480–499.

Davis, S. J., Haltiwanger, J., & Schuh, S. (1996). *Job creation and destruction*. Boston, MA: MIT Press.

de Wit, G., & de Kok, J. (2013). Do small businesses create more jobs? New evidence for Europe. *Small Business Economics,* 1–13.

Delmar, F., & Davidsson, P. (2000). Where do they come from? Prevalence and characteristics of nascent entrepreneurs. *Entrepreneurship & Regional Development, 12*, 1–23.

Delmar, F., Davidsson, P., & Gartner, W. B. (2003). Arriving at the high-growth firm. *Journal of Business Venturing, 18*(2), 189–216.

Delmar, F., & Shane, S. A. (2004). Legitimating first: Organizing activities and the survival of new ventures. *Journal of Business Venturing, 19*, 385–410.

Derbyshire, J., & Garnsey, E. (2014). Firm growth and the illusion of randomness. *Journal of Business Venturing Insights, 1*, 8–11.

DesRoches, D., Barton, T., Ballou, J., Potter, F., Zhao, Z., Santos, B., & Sebastian, J. (2007). *Kauffman Firm Survey (KFS) baseline methodology report mathematica policy research*: Retrieved from http://www.mathematicampr.com/~/media/publications/PDFs/labor/KFS Baseline method.pdf.

Dimov, D. (2010). Nascent entrepreneurs and venture emergence: Opportunity confidence, human capital, and early planning. *Journal of Management Studies, 47*(6), 1123–1153.

Dimov, D. (2011). Grappling with the unbearable elusiveness of entrepreneurial opportunities. *Entrepreneurship: Theory and Practice, 35*(1), 57–81.

Edmondson, A. S., & McManus, S. (2007). Methodological fit in management field research. *Academy of Management Review, 32*(4), 1155–1179.

Edwards, J. R., & Lambert, L. S. (2007). Methods for integrating moderation and mediation: A general analytical framework using moderated path analysis. *Psychological Methods, 12*(1), 1.

Elert, N., Andersson, F., & Wennberg, K. (2014). The impact of entrepreneurship education in high school on long-term entrepreneurial performance. *Journal of Economic Behavior & Organization, 111*(2), 209–223.

Evanschitzky, H., Baumgarth, C., Hubbard, R., & Armstrong, J. S. (2007). Replication research's disturbing trend. *Journal of Business Research, 60*(4), 411–415.

Fairlie, R. W., Karlan, D., & Zinman, J. (2015). Behind the GATE experiment: Evidence on effects of and rationales for subsidized entrepreneurship training. *American Economic Journal: Economic Policy, 7*(2), 125–61.

Fauchart, E., & Gruber, M. (2011). Darwinians, communitarians, and missionaries: The role of founder identity in entrepreneurship. *Academy of Management Journal, 54*(5), 935–957.

Garonne, C. (2014). *Business planning in emerging firm: Uses and effects.* Doctoral dissertation, Queensland University of Technology, Brisbane.

Gartner, W. B., & Carter, N. (2003). Entrepreneurial behavior and firm organizing processes. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research*. Dordrecht, NL: Kluwer Academic.

Gartner, W. B., & Shaver, K. G. (2012). Nascent entrepreneurship panel studies: Progress and challenges. *Small Business Economics, 39*(3), 659–665.

Gartner, W. B., Shaver, K. G., Carter, N. M., & Reynolds, P. D. (2004). *Handbook of entrepreneurial dynamics: The process of business creation*. Thousand Oaks, CA: Sage.

Garud, R., & Karnøe, P. (2003). Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship. *Research Policy, 32*(2), 277–300.

Garvi, M. (2007). *Capital for the future: Implications of founding visions in the venture capital setting*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Gordon, S. R. (2012). *Dimensions of the venture creation process: Amount, dynamics, and sequences of action in nascent entrepreneurship*. Doctoral dissertation, Queensland University of Technology, Brisbane.

Grégoire, D. A., Barr, P. S., & Shepherd, D. A. (2010). Cognitive processes of opportunity recognition: The role of structural alignment. *Organization Science, 21*(2), 413–431.

Grégoire, D. A., & Shepherd, D. A. (2012). Technology-market combinations and the identification of entrepreneurial opportunities: An investigation of the opportunity-individual nexus. *Academy of Management Journal, 55*(4), 753–785.

Grégoire, D. A., Shepherd, D. A., & Schurer Lambert, L. (2010). Measuring opportunity-recognition beliefs. *Organizational Research Methods, 13*(1), 114–145.

Gustafsson, V. (2004). *Entrepreneurial decision-making*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Hamilton, R. T. (1989). Unemployment and business formation rates: reconciling time-series and cross-section evidence. *Environment and Planning A, 21*(2), 249–255.

Hang, M. (2007) *Media business venturing: A study on the choice of organizational mode*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Hellerstedt, K. (2009). *The composition of new venture teams: Its dynamics and consequences*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Hill, S. A., & Birkinshaw, J. M. (2010). Idea sets: Conceptualizing and measuring a new unit of analysis in entrepreneurship research. *Organizational Research Methods, 13*(1), 85–113.

Horsten, L., & Weber, E. (2005). *INUS conditions*. Wiley StatsRef: Statistics Reference Online. Doi: 10.1002/9781118445112.stat06759.

Hunter, E. J. (2009). *Celebrity entrepreneurship and celebrity endorsement.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Hunter, E. J., Burgers, J. H., & Davidsson, P. (2009). In G. T. Lumpkin & J. Katz (Eds.), *Celebrity capital as a strategic asset: Implications for new venture strategies* (Advances in entrepreneurship, firm emergence, and growth, Vol. 1, pp. 137–160). Bingley, UK: Emerald.

Jenkins, A. (2012). *After firm failure: Emotions, grief, and re-entry*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Karlsson, T. (2005). *Business plans in new ventures: An institutional perspective*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Katz, J., & Gartner, W. B. (1988). Properties of emerging organizations. *Academy of Management Review, 13*(3), 429–441.

Kish, L. (1987). *Statistical design for research*. New York, NY: John Wiley & Sons, Inc.

Klein, K. J., & Kozlowski, W. J. (2000). *Multilevel theory, research, and methods in organizations*. San Francisco, CA: Jossey-Bass Inc.

Kozlowski, S. W. J., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design capturing the dynamics of emergence. *Organizational Research Methods, 16*(4), 581–615.

Lee, L. F. (1982). Specification error in multinomial logit models: Analysis of the omitted variable bias. *Journal of Econometrics, 20*(2), 197–209.

Liao, J., & Welsch, H. (2008). Patterns of venture gestation process: Exploring the differences between tech and non-tech nascent entrepreneurs. *Journal of High Technology Management Research, 19*(2), 103–113.

Liao, J., Welsch, H., & Tan, W. L. (2005). Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns. *Journal of High Technology Management Research, 16*(1), 1–22.

Lichtenstein, B. B., Carter, N. M., Dooley, K. J., & Gartner, W. B. (2007). Complexity dynamics of nascent entrepreneurship. *Journal of Business Venturing, 22*(2), 236–261.

Lieberson, S. (1985). *Making it count: The improvement of social research and theory*. Berkeley, CA: University of California Press.

Lockett, A., Wiklund, J., Davidsson, P., & Girma, S. (2011). Organic and acquisitive growth: Re-examining, testing and extending Penrose's growth theory. *Journal of Management Studies, 48*(1), 48–74.

Lohrke, F. T., Holloway, B. B., & Woolley, T. W. (2010). Conjoint analysis in entrepreneurship research: A review and research agenda. *Organizational Research Methods, 13*(1), 16–30.

Lövstål, E. (2008). *Management control systems in entrepreneurial organisations: A balancing challenge.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Markowska, M. (2011). *Entrepreneurial competence development: Triggers, processes & consequences*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

McKelvey, B. (2004). Toward a complexity science of entrepreneurship. *Journal of Business Venturing, 19*(3), 313–341.

McKelvie, A. (2007). *Innovation in new firms: Examining the role of knowledge and growth willingness.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden

McMullen, J. S., & Dimov, D. (2013). Time and the entrepreneurial journey: The problems and promise of studying entrepreneurship as a process. *Journal of Management Studies, 50*(8), 1481–1512.

Menzies, P. (1989). Probabilistic causation and causal processes: A critique of Lewis. *Philosophy of Science, 56*(4), 642–663.

Mintzberg, H. (1974). *The nature of managerial work*. New York, NY: Harper & Row.

Naldi, L. (2008). *Growth through internationalization: A knowledge perspective on SMEs.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

O'Boyle, E. H., Banks, G. C., & Gonzalez-Mulé, E. (2014). The chrysalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*. doi:10.1177/0149206314527133.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, UK: Cambridge University Press.

Pe'er, A., & Vertinsky, I. (2008). Firm exits as a determinant of new entry: Is there evidence of local creative destruction? *Journal of Business Venturing, 23*(3), 280–306.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903.

Rauch, A., & Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology, 16*(4), 353–385.

Reynolds, P. D. (2007). New firm creation in the United States: A PSED I overview. *Foundations and Trends in Entrepreneurship, 3*(1), 1–150.

Reynolds, P. D. (2009). Screening item effects in estimating the prevalence of nascent entrepreneurs. *Small Business Economics, 33*(2), 151–163.

Reynolds, P. D., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies, 28*(4), 443–456.

Rönkkö, M., Peltonen, J., & Arenius, P. (2013). Selective or parallel? Toward measuring the domains of entrepreneurial bricolage. In A. Corbett & J. Katz (Eds.), *Advances in entrepreneurship, firm emergence and growth* (15th ed., pp. 43–61). Bingley, UK: Emerald.

Ross, W. D. (1951). *Plato's theory of ideas*. Oxford, UK: Clarendon Press.

Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. *Research in Organizational Behavior, 7*, 1–37.

Ruef, M., Aldrich, H. E., & Carter, N. M. (2003). The structure of organizational founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review, 68*(2), 195–222.

Russell, B. (1912). On the notion of cause. *Proceedings of the Aristotelian Society-New Series, 13*, 1912–1913.

Samuelsson, M. (2004). *Creating new ventures: A longitudinal investigation of the Nascent venturing process*. Doctoral dissertation, Jönköping International Business School, Jönköping

Sarasvathy, S. D. (2001). Causation and effectuation: Towards a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–288.

Schaffer, B. S., & Riordan, C. M. (2003). A review of cross-cultural methodologies for organizational research: A best-practices approach. *Organizational Research Methods, 6*, 169–215.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, UK: MA: Harvard University Press.

Semasinghe, D. M. (2011). *The role of idea novelty and relatedness on venture performance.* Doctoral dissertation, Queensland University of Technology, Brisbane.

Senyard, J., Baker, T., Steffens, P., & Davidsson, P. (2014). Bricolage as a path to innovativeness for resource-constrained new firms. *Journal of Product Innovation Management, 31*(2), 211–230.

Shane, S. A. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science, 11*(4), 448–469.

Shane, S. A., & Delmar, F. (2004). Planning for the market: Business planning before marketing and the continuation of organizing efforts. *Journal of Business Venturing, 19*, 767–785.

Shugan, S. M. (2006). Errors in the variables, unobserved heterogeneity, and other ways of hiding statistical error. *Marketing Science, 25*(3), 203–216.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.

Smith, N. R. (1967). *The entrepreneur and his firm: The relationship between type of man and type of company*. East Lansing, MI: Michigan State University.

Stanworth, J., & Curran, J. (1973). *Management motivation in the smaller business*. Epping, UK: Gower Press.

Steffens, P. R., Weeks, C. S., Davidsson, P., & Isaak, L. (2014). Shouting from the ivory tower: A marketing approach to improve communication of academic research to entrepreneurs. *Entrepreneurship: Theory and Practice, 38*(2), 399–426.

Ucbasaran, D., Westhead, P., & Wright, M. (2006). *Habitual entrepreneurs*. Cheltenham, UK: Edward Elgar Publishing.

Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2009). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing, 26*(3), 341–358.

Uy, M. A., Foo, M. D., & Aguinis, H. (2010). Using experience sampling methodology to advance entrepreneurship theory and research. *Organizational Research Methods, 13*(1), 31.

van Burg, E., & Romme, A. G. L. (2014). Creating the future together: Toward a framework for research synthesis in entrepreneurship. *Entrepreneurship: Theory and Practice, 38*, 369–397.

Van de Ven, A. H. (2007). *Engaged scholarship: A guide for organizational and social research*. Oxford, UK: Oxford University Press.

Van de Ven, A. H., Polley, D., Garud, R., & Venkataraman, S. (1999). *The innovation journey*. Oxford, UK: Oxford University Press.

van Weezel, A. (2009). *Entrepreneurial strategy-Making mode and performance: A study of the newspaper industry.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review, 92*, 548–573.

Welter, F. (2011). Contextualizing entrepreneurship—conceptual challenges and ways forward. *Entrepreneurship: Theory and Practice, 35*(1), 165–184.

Wennekers, S., Stel, A., Thurik, A. R., & Reynolds, P. D. (2005). Nascent entrepreneurship and the level of economic development. *Small Business Economics, 24*, 293–309.

Western, B. (1998). Causal heterogeneity in comparative research: A Bayesian hierarchical modelling approach. *American Journal of Political Science, 42*, 1233–1259.

Wigren, C. (2003). *The spirit of Gnosjö. The grand narrative and beyond.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Wiklund, J. (1998). *Small firm growth and performance: Entrepreneurship and beyond.* Doctoral dissertation, Jönköping International Business School, Jönköping.

Wiklund, J., Davidsson, P., & Delmar, F. (2003). What do they think and feel about growth? An expectancy-value approach to small business managers' attitudes towards growth. *Entrepreneurship: Theory and Practice, 27*(3), 247–269.

Woo, C. Y., Cooper, A. C., & Dunkelberg, W. C. (1991). The development and interpretation of entrepreneurial typologies. *Journal of Business Venturing, 6*(2), 93–114.

Wu, D. D., Kefan, X., Hua, L., Shi, Z., & Olson, D. L. (2010). Modeling technological innovation risks of an entrepreneurial team using system dynamics: An agent-based perspective. *Technological Forecasting and Social Change, 77*(6), 857–869.

Zahra, S. A., & Wright, M. (2011). Entrepreneurship's next act. *Academy of Management Perspectives, 25*(4), 67–83.

Sampling and Case Selection Issues

5

Abstract

Who and what should we study in entrepreneurship research? What type of entities, how many, and which particular ones should we study in order to effectively answer our entrepreneurship-related research questions? Starting from the axiom that social science is not like opinion polling, this chapter provides a somewhat unorthodox view on sampling and case selection which focuses on the theoretical relevance of the selected entities. Specific sampling challenges are discussed for entrepreneurship research focusing on the individual, venture, firm, industry, and spatial (region/country) levels of analysis.

5.1 A Different Look at Sampling

This is not a conventional sampling chapter. A conventional sampling chapter builds on statistical inference theory and deals primarily with two issues. First, how can we create a sampling frame and sampling mechanism that allow us to draw a statistically representative sample from the empirical population in question? Issues here are over and undercoverage of the sampling frame relative to the population: techniques for drawing a random sample or, to be more precise, one for which the sampling probability of each element in the sampling frame is known and (possibly) techniques for minimizing nonresponse. Second, how large does the sample have to be for us to detect the differences and effects our theory predicts? Based on assumptions of variances and effect sizes, this involves calculating the statistical power (Cohen, 1988) of different sample sizes. The more conventional side of statistical inference theory, i.e., *statistical significance testing*, deals with the opposite risk that effects found in the sample may be due to random sampling error rather than reflecting effects that are true for the population from which the sample was drawn.

My agenda is different. As there already exist a plethora of books and chapters on sampling written by statistical experts, I aim instead to offer a sampling chapter

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_5

based on what the statistical experts do not provide: sampling and case selection viewed from the perspective of theory- and curiosity-driven social science research and backed with extensive practical research experience. What I want to discuss here, then, is how—on different levels of analysis—we can obtain data from a sample of cases that are *theoretically relevant*. By this I mean that the sample is composed of cases that reflect the theoretical unit of analysis and the theoretically relevant variance in the characteristics of these cases. Discussion of *boundary conditions* has increased a lot over the past 10 years, and this gives at least some implicit guidance on the issues I aim to discuss. However, I'd like to make these links more explicit. I also want to discuss the extent to which the sample is *workable* from a practical point of view, i.e., that it is possible without breaking one's back (or budget, although not all my suggestions will be for the most frugal research design) to obtain data from or about units in the target population. Indeed, large parts of my argumentation are not about "sampling" in the statistical sense at all but apply equally well to studies of entire populations and to an extent also to selection of cases in "small *n*" research. What I address in this chapter is how we determine what are to be the *cases* in our data matrix, whereas the next chapter (on operationalization) will deal with the *variables* in the matrix. As will be argued below, for most research questions every accessible empirical population is a sample relative to the theoretically relevant population, i.e., the category for which we hope our results have some validity.

Below, I will first expand on the theme just introduced. I will then discuss sampling problems and solutions for research on different levels of analysis—individual, venture, firm, industry, and spatial units (region, nation). As it turns out, this will be enough, or even more than enough, for a chapter.

5.1.1 Social Science Is Not Opinion Polls

Sampling aiming for representativeness and associated significance testing are important safeguards against ignoring relevant parts of empirical populations, giving undue weight to atypical cases or ascribing substantive meaning to results that can easily be generated by chance factors. However, for the statistical inference apparatus to be applicable in a strict sense, the population should be well defined and the sample should reflect the composition of this population in a probabilistically known manner. These are ideals that are rarely achieved in social science research. For one thing, the painful fact is that response rates in published research typically fall in the 5–35 % range. This alone makes application of statistical inference highly dubious. Even if you can show there is no statistically significant difference between respondents and nonrespondents on a few sociodemographic variables, we know for a fact that they differ on one important behavioral dimension: the propensity to participate in surveys. To believe that this behavioral difference is uncorrelated to the characteristics, behaviors, and outcomes our research concerns is naïve, or dumb.

To make matters worse, statistical inference theory is a tool that is tailor made for opinion polls and industrial quality control rather than for the true needs of a social science researcher (cf. Cohen, 1994; Hubbard & Lindsay, 2013a; 2013b; Oakes, 1986; Schwab, Abrahamson, Starbuck, & Fidler, 2011). Consider political opinion polls. Here, we have a clearly defined population, which in most countries is also reasonably reachable: all eligible voters. What we want to know are their political preferences on the day of investigation. Hence, we can draw a random sample and ask the selected individuals about their preferences. Applying statistical inference theory, we can with high accuracy estimate with what uncertainty our sample results are associated. This allows us to determine whether the difference between two political parties or the change for one party over time deserves a substantial interpretation or is likely to be due to chance (i.e., random sampling error). Clearly, probability sampling and significance testing are useful tools in this situation (assuming high response rates in the opinion polls). We can say much more on the basis of this probability sample than on the basis of just any voter sample of unknown origin.

In entrepreneurship research, I can really only think of one major study that comes close to this situation, namely the country comparisons of the prevalence of "nascent entrepreneurs" or "early-stage entrepreneurial activity" in the Global Entrepreneurship Monitor (GEM) (Álvarez, Urbano, & Amorós, 2014; Amorós, Bosma, & Levie, 2013; Bergmann & Stephan, 2013). Here, what we want to know is what proportions of the adult population in various countries are involved in business start-ups at a given point in time and how uncertain are the estimates that we get from samples of a certain size from the adult populations in those countries? For the most part, however, social science research is *not* like opinion polls, and theories are not built by democratic vote. That is, it is not a given that every empirical case in the sample at hand is relevant to our theoretical research question or that all the cases in the sample are equally important. What we are really after in social science research is *theoretical representativeness*—that the studied cases are relevant for the theory we try to test or develop. There is no way we can draw a random sample directly from *the* theoretical population, because that population does not exist in one place at one time. This is why I argue that every empirical population, even if we investigate it in its entirety, is a nonrandom sample from the theoretically relevant population. The questions are: What is the theory about and what should therefore be represented in the sample? Can we find *an* empirical population that is theoretically relevant (to study in its entirety or from which to sample)?

Figure 5.1 sketches the reality we are facing. Assume our theoretical interest is in "the new venture start-up process." Unless we further delimit our interest, this means all such processes—past, current, and future and independent and corporate, in any country. Regardless of how precisely we delimit our theoretical interest, we are likely to always face a situation where (a) the entire, relevant population is not accessible for probabilistic sampling and (b) the theoretically interesting relationships are unlikely to be exactly the same for all subpopulations within the theoretically relevant population. This implies that searching for "the truth" that is

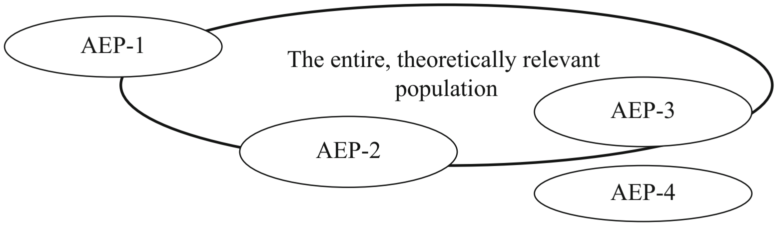


Fig. 5.1 Theoretically relevant vs. empirically accessible populations (AEP)

representative for the entire theoretical population may be futile and not necessarily even desirable, at least from the perspective of one, individual study.

A theory-driven logic instead gears the researcher to ask "Can I find some accessible empirical population (AEP), which provides a relevant testing ground for my theoretical ideas?" This question of theoretical relevance is the more important concern, which should be asked and answered before any issues of statistical representativeness are raised. If we are sloppy with the first question, we may end up working with statistically representative samples from theoretically inappropriate populations like AEP-1 or even AEP-4. In relation to our interest in "the new venture start-up process," the latter could be, for example, interview-based case studies of the development processes for young firms after they have already become operating businesses or the study of experimental or simulated processes that are so artificial that they are not informative in relation to the real-world phenomenon in which we are interested.

If we are lucky and/or do our job well, we may be able to identify a population like AEP-3. However, even if we managed to investigate this entire subpopulation with a perfectly designed study, this would not allow us to draw inference *on statistical grounds* to the entire, theoretically relevant population in which we are interested. This is because AEP-3 is not a probability sample from the larger population. More likely, we could not afford to undertake a population study of AEP-3. Instead, we would draw a sample from AEP-3 that is large enough for statistical purposes. It is thus for this smaller task—to draw inferences from our sample to a subpopulation like AEP-3—that statistical inference has a role in observational studies. To convince us about the general or variable applicability of our theory across the entire, theoretically relevant population, we need something else: replication using several different samples from different subpopulations like AEP-3. When results across such studies pull in the same direction, we have some basis for the broader inference we really want to make. This would give us some justified confidence that what we have learnt from research represents neither the idiosyncrasies of a particular (albeit relevant) subpopulation nor the unfortunate incidence of probabilistic sampling yielding a nonrepresentative sample. Thus, we would have a good basis for giving advice to or about business start-ups that were not part of the research.

In most cases, accessible empirical populations do not come as neat as AEP-3. A more realistic situation is something like the stylized AEP-2, where part of the accessible population is, in fact, theoretically irrelevant noise. It is then important that the researcher can withstand the temptation to keep the sample as large as possible just in order to maximize statistical power. Instead, the theoretically irrelevant cases should be trimmed from the sample. For example, in a study of entrepreneurial failure, it is important not to confuse matters by including those cases that exit for reasons unrelated to financial distress or other signs of failure (Wennberg, Wiklund, DeTienne, & Cardon, 2010). Such trimming of the sample should normally strengthen the magnitude of estimated relationships. Besides, if support for the theory is contingent on including theoretically irrelevant cases, the "positive" result isn't really support for the theory after all, is it?

5.1.2 Judgment, Empirical Ubiquity, and Theoretical Relevance

Figure 5.1 is also useful for reflections around the notions of *probabilistic* (or random) sample, *judgment* sample, and *convenience* sample. A probabilistic sample fulfills certain statistical criteria, but the notion as such does not safeguard us from the possibility that the sample was drawn from something like AEP-4. A judgment sample worthy of that label should be drawn from populations like AEP-3 or, after trimming, from AEP-2. However, this label also suggests that the researcher was unable to apply a probabilistic sampling mechanism in drawing cases from the population in question. The notion of convenience sample clearly signals the risk that the sample is derived from AEP-1- or AEP-4-type populations. Overall, I would argue we should prefer a judgment sample to a convenience sample but also to a probabilistic sample from a dubious population. No level of statistical significance can fix the problem that the population was wrong in the first place.

The above concerns the relevance of the studied population and hence the sample drawn from that population (which, again, is usually but a small subpopulation of the theoretical population about which we *really* want to make statements). A related concern is the theoretical relevance and empirical ubiquity of subtypes of elements within AEPs. This is a nonproblem in political opinion polls or industrial quality control. Each voter has one vote; hence, they are all equally important for the accuracy of the results of the opinion poll. Same with quality control, the functionality of each item of a batch of identical engine parts is equally important to the decision of whether the batch can be shipped or not to a valued customer. Not so with most samples of interest to entrepreneurship researchers. Our samples are often numerically dominated by a "modest majority" (Davidsson & Gordon, 2012). In a random sample of nascent ventures, imitative start-ups are much more numerous than are innovative ones (Samuelsson & Davidsson, 2009), and those that are destined for VC backing and future IPOs are so few that they are practically nonexistent in such samples. If you draw a simple random sample of small firms, here, meaning commercially active firms with fewer than 50 employees, you are likely to end up with something like this: 62 % self-employed without employees, just short of 35 % micro-firms with 1–9 employees, and a remainder of less than 4 % firms with 10–49 employees (NUTEK, 2002; cf. also Chap. 10). Similarly, in studies of regions or industries, there are typically many cases with small populations and a few with very large populations. What we are encountering here is, of course, a special case of the general feature of *heterogeneity* and how it should be handled. I dare ask, are those categories that there are many of in the AEPs we have access to necessarily more important from a theoretical point of view? I dare answer no, for most conceivable research questions they are not!

Several implications follow from the above. First, the most important sampling issue is not statistical but *theoretical* representativeness. That is, it should be carefully ascertained—and communicated—that the elements in the sample represent the type of entities and/or phenomena that the theory makes statements about. This is, by the way, equally relevant for case study research. Second, simple random sampling is not necessarily the ideal. Stratified and deliberately "narrow" statistical samples and even judgment samples may on theoretical grounds be preferable in many situations. Third, *replicability*—not statistical significance alone—is the crucial theory test. The development and testing of sound theory requires replication within AEPs to rule out statistical artifacts and across AEPs to provide a basis for the broader inference we are ultimately striving for. This, then, further reinforces the importance of replications like those that will be presented in Chap. 9.

5.2 Sampling Individuals

How can we obtain a theoretically relevant—and hopefully statistically representative—sample of individuals for an entrepreneurship study? This is *not* a simple matter. Consider the first backbone reaction: let's study entrepreneurs! The first problem with this is that this is not a well-defined population but a hazy and moving target. Hence, it is not possible to create an indisputable sampling frame. Some people sometimes engage in entrepreneurial activities as we have defined them in Chap. 2. At other times they don't—but then other people are active in entrepreneurial endeavors. Whom should we include in the sampling frame? Self-employed? Owner-managers? Current venture champions? All who have ever engaged in any behavior we define as "entrepreneurship"? By the way, how and where do we obtain contact information for these people?

Assume we have defined a sampling frame of current "entrepreneurs" that we can live with, as well as a comparison group. Now we compare the two groups and find some differences. How should these differences be interpreted? As causal factors that make people engage in entrepreneurship, right? Well, so we would like to believe. The problem is that when we compare people "currently active in entrepreneurship" with those who currently are not, we confound several different factors:

- The propensity to *engage* in such behavior. Those with higher propensity should, ceteris paribus, have a higher likelihood of ending up in our "entrepreneur" sample.
- The propensity to *persist* in self-employment (Patel & Thatcher, 2014), sometimes in the face of failure. Those who try again or stay in business despite substandard performance (cf. DeTienne, Shepherd, & De Castro, 2008) should,

ceteris paribus, have a higher likelihood of ending up in our "entrepreneur" sample.
- The ability to *succeed* in such behavior. Those who are successful in entrepreneurial endeavors should, ceteris paribus, have a higher likelihood of still being members of the group(s) we sample as "entrepreneurs" and therefore end up in that sample.

In addition, broadly based studies relating individual characteristics to engagement and success in relation to one particular venture would run considerable risk of not being able to tease out the person-related factors from the influence of the characteristics of the environment and the emerging venture itself (Aldrich & Ruef, 2006; Shane & Venkataraman, 2000; Shaver, 2010). Moreover, during the last 10–15 years, entrepreneurship researchers have become increasingly aware that roughly half of all business founders start their ventures in teams (Delmar & Shane, 2006; Ruef, Aldrich, & Carter, 2003; Steffens, Terjesen, & Davidsson, 2012). To put it mildly, this idea of cross-sectional comparison of "entrepreneurs" with others maybe wasn't as great as we thought it was. At best, it would give confounded answers to the questions that are implicitly or explicitly underlying the chosen design, namely "What makes some individuals more likely than others to (a) engage and (b) succeed in entrepreneurial activities?"

This said, the individual difference approach to entrepreneurship should not be completely counted out. The fact is that research over the last decade has strengthened the evidence for personality-based influence on entrepreneurial behavior and success (Brandstätter, 2011; Rauch & Frese, 2007). Sampling (and following) individuals remains relevant and suitable for the following types of studies:

1. Longitudinal, career-oriented studies where individuals' engagement in and success at entrepreneurship can be aggregated and compared over time, across particular venture ideas and teams. This vastly increases the chances of correctly attributing effects to enduring characteristics of the individual as such. New, linked employer-employee data sets have made this type of study possible (Amaral, Baptista, & Lima, 2011; Campbell, 2005; Sørensen, 2007).
2. The study of expert (habitual, repeatedly successful) vs. novice entrepreneurs (Sarasvathy, 2008; Ucbasaran, Westhead, & Wright, 2006). This type of study holds some promise of generating knowledge about the teachable and learnable skills that signify successful entrepreneurship. In terms of sampling, an indisputable sampling frame is not to be hoped for, but stringent criteria for being classified as "expert" and "novice" should be employed—mere experience does not necessarily imply expertise. The use of cognitive theory on expertise may help this type of research in general, including the identification of sampling criteria (Gustafsson, 2004).
3. The study of fit between individual(s) and new venture idea(s) Shane (2000) is a famous example in this category (although he called such ideas "opportunities"). See further Chap. 8.

4. Studies of how structural and situational factors influence entrepreneurial behavior. I have argued above and elsewhere (Davidsson, 1992; 2006) that if the explanation for entrepreneurial behavior is not innate characteristics of individuals, then entrepreneurship research on the individual level can, in principle, use any sample of individuals. We can here think of laboratory research where the researcher manipulates hypothetical situations, so as to induce entrepreneurial attitudes, beliefs, decision-making, and behaviors. If the theoretical prediction is that *humans* should react in particular ways to external stimuli, then any sample of humans is relevant. However, the participants need to be capable and motivated to provide quality data, and serious scholarly journals are unlikely to be impressed by samples consisting of easily accessible students or cheap online panels.

I have not yet discussed the intricate problems of sampling "nascent entrepreneurs." This is because at the entry point, sampling individuals and sampling emerging new ventures coincide. This type of sampling will be described in some detail below. For now, we may note that the study of nascent entrepreneurs can be combined with studying experts vs. novices, and fit between individual and idea, as described above. We may also sum up that drawing a theoretically relevant sample of individuals is not an easy task.

5.3 Sampling Emerging New Ventures

5.3.1 Identifying an Eligible Sample of On-Going Independent Venture Start-Ups

I have pointed out as particularly important and promising for entrepreneurship research the type of study that uses the venture idea, including the activity and organization that evolve around it, as the level of analysis. For a long time, this was a relatively neglected type of study, no doubt in part because it is a tricky one from a sampling point of view. To begin with, it requires one to define a criterion for what it means to be an emerging venture, a lower limit below which no "emerging venture" can be meaningfully said to exist. What should that criterion be? That an individual nurtures a dream or intention to start some kind of venture at some future point in time? That an individual is pondering a specific idea for a new venture? That concrete action has been taken toward the realization of such an idea? If the latter, what and/or how much action should be required?

The ideal study would capture all the cases at the exact moment that they transition from nonexistence to fulfilling a minimum criterion that can be agreed upon and then follow their journey from that point forward (Delmar, 2015). In reality, the initiation of ideas or intentions cannot be captured in real time. As regards actions, it was realized early on that no single action would work as a valid marker of process initiation across a majority of start-ups; the sequences of actions undertaken in the start-up process are simply too diverse (Liao, Welsch, & Tan, 2005; Reynolds & Miller, 1992).

So if an action criterion were to be included, it would have to be each case's first "start-up activity" whatever precise activity that might be.

Further, it would be unrealistic to try to capture each case at the exact time the first activity is undertaken. This requires one to also define an upper limit beyond which the case no longer is an "emerging new venture." When is the process of emergence completed, so that the case now is a functional young firm rather than an emerging venture?

Luckily, smart people have invested quite a bit of energy into the problems of minimum and maximum criteria for qualifying as an emerging new venture (Katz & Gartner, 1988; McMullen & Dimov, 2013; Reynolds, 2007; Reynolds, 2009; Reynolds & Curtin, 2008; Schoonhoven, Burton, & Reynolds, 2009; Shaver, Carter, Gartner, & Reynolds, 2001). The PSED and GEM research programs have carefully developed and honed a screening methodology to identify valid samples of emerging new ventures (and/or "nascent entrepreneurs"). Vast experience has been invested into these efforts. Therefore, anyone who starts from scratch in order to try to do better on some dimension is likely to end up doing far worse in other ways—and only learn this the hard way, in arrears.

Consequently, we built closely on these experiences when we designed the abovementioned CAUSEE project (Davidsson & Reynolds, 2009; Davidsson & Steffens, 2011; Davidsson, Steffens, & Gordon, 2011). Figure 5.2 (originally drawn by Steffens for the works just cited) illustrates the screening procedure, which is designed to capture two categories: emerging new ventures (i.e., nascent ventures) and young firms (i.e., those that are overqualified as nascent ventures). Households were sampled through random digit dialing over the phone, and the screener questions were directed at the adult in the household who was next up for their birthday.

Although the figure may appear complex, it is actually a simplification of the screener interview.[1] In short, the procedure aims to (a) quickly screen out most non-eligible cases while casting a broad net so as to capture all cases that qualify according to *the researcher's* definitions (first three questions), (b) correctly classify eligible cases as either nascent ventures or young firms (first three questions plus later possibility of reclassification), and (c) test additional min and max criteria for the respective categories so as to exclude non-eligible cases from initial suspects (activity, ownership, positive cash flow, age).

The establishment of this sampling mechanism is quite an achievement. Previously, in order to find an "early-stage" sample, researchers would have had to rely on a decidedly nonrandom sampling frame like would-be founders contacting some support agency or the like (e.g., Gatewood, Shaver, & Gartner, 1995). Alternatively, they would use the first visible trace that the new venture leaves in some type of register (e.g., Kessler & Frank, 2009) which typically happens too late to include early failures and to follow the creation process in real time. However, no

[1] For example, a "control group" subsample was drawn from the non-eligible cases, and "tie-breaker" mechanisms were employed for the situation where a respondent was involved in more than one nascent and/or young firm or ended up in an otherwise never-ending loop.

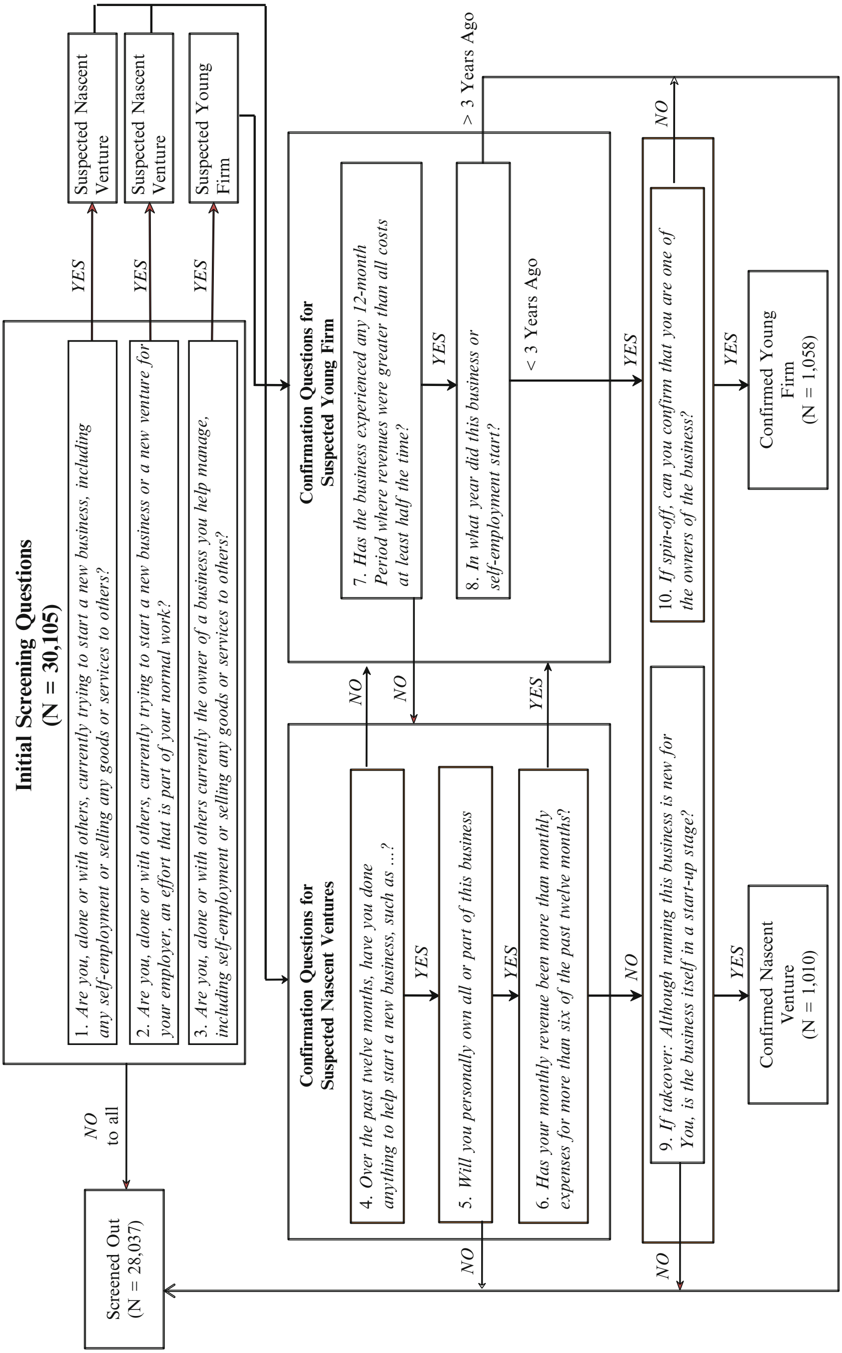


Fig. 5.2 CAUSEE screening procedure (adapted from Davidsson et al., 2011)

procedure is perfect, and over time we have been able to identify quite a number of potential issues with the PSED-GEM-CAUSEE approach (Davidsson & Gordon, 2012). Depending on the purpose of the research, each of these issues may be more or less serious. In assessing these shortcomings, it is also useful to carefully consider whether other, available alternatives would do any better on the particular issue or overall.

1. It is costly. Phone interviewing costs more than some most contact formats, and a very large number of cases—mostly non-eligible ones—have to be conducted in order to arrive at a valid sample.
2. The sample will be very heterogeneous along a number of dimensions, e.g., industry, motivation and goals, resources, and stage of development when first captured. This means that the heterogeneity problems discussed in Chap. 4—potential unobserved heterogeneity, causal heterogeneity, and uneven validity—will apply and have to be dealt with.
3. Under ideal circumstances, the procedure will yield a representative sample of nascent entrepreneurs, i.e., *individuals* currently involved in an emerging new venture. However, like for any survey research, there are sources of potential sampling bias. First, there is nonresponse at two stages: (1) not participating at all and (2) not participating in the full survey after having been identified as eligible (in CAUSEE, this second step led to a 40 % drop relative to the numbers of eligible cases in Fig. 5.2). Second, despite the refined procedure, there is still some risk that individuals interpret the questions differently and therefore under- or overreport eligibility for inclusion. Third, while random digit dialing to households was still viable when CAUSEE was initiated in 2007, it is not so today. Achieving random sampling individuals via mobile phone numbers (or any other common contact vehicle) has its own challenges and may or may not be feasible (Steffens, Tonelli, & Davidsson, 2011).
4. Regarded as a sample of emerging *new ventures*, the procedure oversamples start-up processes of long duration, because these are available for sampling over a longer period of time. Assume the entire population of start-up processes in a given year consists of 40 cases. Ten of those are "slow" start-ups, which are initiated on January 1 and completed on December 31. The other 30 are "quick" start-ups, which take 4 months from initiation to completion. Ten each are initiated on January 1, May 1, and September 1, and consequently ten "quick" start-ups are completed on April 30, August 31, and December 31. Although the proportion of "quick" to "slow" start-ups is three to one on a yearly basis, we will sample from a population with a 50/50 distribution no matter what date we select for our sampling. This is potentially a serious bias, and different remedies have been tried and suggested (Delmar, 2015; Delmar & Shane, 2004; Yang & Aldrich, 2012). The procedure will also oversample ventures started by teams, because more team members mean higher sampling probability. However, the fact that many teams consist of members from the same household (Ruef et al., 2003) reduces this bias. Further, as per my earlier argument, a sample with 25 %

team start-ups may not be theoretically superior to one with 50 %, even if the former better represents the true empirical distribution.

5. Despite our best efforts, the procedure seems to capture a significant subgroup of people who do enough to formally qualify but who do not seem to ever put their start-up effort to an acid test. Hence, they tend to remain "still trying" cases more or less forever (Reynolds & Curtin, 2008). These cases may dilute estimated relationships and should perhaps preferably be excluded—if satisfactory criteria for doing so could be found—unless, of course, the research question is what drives people to think and say they are starting a new venture when really they aren't.
6. Retrospection. It is expensive enough to identify an analyzable number of cases that are currently somewhere in the start-up process, without yet having completed it. To only include cases that just became eligible is inconceivable from a practical point of view. As a result, although the sampling aims at allowing concurrent study of start-up processes, the reality is that on average, close to half of that process has already happened prior to the first interview. This means that the design is not free from potential retrospection bias and that some analyses may treat concurrent and retrospective reporting as equally reliable.
7. Finally, the design is subject to the "modest majority" issue discussed above. For this reason, we added separate judgment samples of "high(er) potential" nascent and young firms (100+ of each). In order to increase representativeness, we sourced these from as many different sources as we could think of: research laboratories, incubators, patent attorneys, etc. This would likely reduce any particular bias pertaining to each single source, thus approximating the canceling out of (unwanted) biases inherent in random sampling. Identifying or defining "high potential" at an early stage is challenging, though. We found the sampling source or any other single criterion to be insufficient. Therefore, in addition to the regular screener, the cases had to score highly enough on a "high potential" screener, combining criteria based on human capital, aspirations, and technological sophistication (Gordon & Davidsson, 2013). Although this may result in a sample that is more theoretically relevant for some purposes, the combination of several criteria comes at a cost. It makes some group comparisons tautological, and in relational analyses the imposed range restrictions may affect estimated relationships (Johns, 2006).

Despite all these issues, the sampling mechanism developed through PSED and GEM remains the best effort we yet have seen. The simple fact is that there is no fully satisfactory solution to the challenge of obtaining a representative sample of ongoing, independent start-up processes. The seriousness of the listed issues varies by purpose, and many of them can be dealt with through weighing and other corrections introduced in the analysis.

Some features of the PSED methodology that definitely deserve being retained in future studies are (a) catching cases as early as possible, according to researcher-defined criteria for being under- or overqualified for inclusion, (b) following the continuation of the process through repeated data collection, and (c) using phone or other person-to-person means of data capture. In my experience, trying inexpensive but impersonal and non-engaging modes of data collection would lead to disastrous attrition and poor data quality.

Now that the basics about the prevalence and composition of nascent ventures have essentially been answered (e.g., Delmar & Davidsson, 2000; Reynolds, Carter, Gartner, & Greene, 2004), the "catch early and follow over time" can preferably be applied in contexts that offer lower cost, less heterogeneity, greater representation of "high potential" ventures, and perhaps an even earlier and more equal starting point. New means of data capture can also be tried, although impersonal means are likely to work only for supplementary purposes. Here are some ideas:

– Researchers can agree on pooling data from current data sets in order to study large enough subsamples that are more homogeneous in terms of industry, type of location, type of founder(s), and stage of development.
– New studies may be able to use institutions such as incubators, crowdfunding websites, or events like *Startup Weekend* in order to sample at lower cost at an earlier point, at the same time possibly achieving more homogeneity and greater theoretical relevance.
– In such studies, researchers may want to combine founder interview/survey data with experience sampling (Uy, Foo, & Aguinis, 2010), data from other informed individuals (e.g., incubator managers), and any data the emerging ventures generate in electronic or else retrievable form as part of their regular course of action.

I am very much looking forward to creative developments by a new generation of entrepreneurship researchers!

5.3.2 Sampling Ongoing Internal Venture Start-Ups

The PSED screening questions make it possible to also identify "nascent intrapreneurs" and hence new internal ventures (see item 2 in Fig. 5.2, and Parker, 2011). Similarly, these days GEM captures also "employee entrepreneurship" (Steffens, 2013). However, I am not convinced that starting from a sample of individuals is ideal for sampling of new internal ventures. While I was still based in Sweden, we tried instead to start from an existing, large sample of (young, small, and owner-managed) firms, the *1994 Start-Up Cohort* (Dahlqvist, Davidsson, & Wiklund, 2000), to identify internal new ventures with a PSED-like approach. Because our firms had previously been approached with mail questionnaires (where the first few questions were mandatory data collection for a government agency, thus yielding high response rates), we choose a mail questionnaire directed at the CEO for the screening. Under other circumstances, phone interviewing would yield a higher response rate. The focal screening questions (asked in year 2000) were the following:

1. After the start of this company in 1994, have you started any new venture within the company, which during some period has provided income to the company? We are interested in new business initiatives in your company, which have led or could lead to new income-generating activities. NB! Not mergers or acquisitions.
2. Do you have a business initiative in progress now, which you or others in the company have devoted time and possibly other resources to develop, but where the new activity does not yet yield a steady income?

Additional questions asked *when* the new initiative in (2) was initiated and whether the respondent had started any additional *firms* (separate from the sampled one) since 1994. The first question above is intended to define "new initiative" (or new internal venture) and to separate up-and-running initiatives from ongoing ones. The critical screening question is (2). Those who answered this question affirmatively were later contacted for a phone interview. In that interview, the eligibility of the initiative was double-checked with the following question:

3. By initiative toward new business activity, we mean attempts to change or expand the business, for example, developing new products or services, aiming for completely new customers, or entering new markets. We are interested in all such changes, which could affect your future income to a non-negligible degree. With this clarification, would you say that you today have any new initiative toward business activity in progress, which you or others in the company have devoted time and possibly other resources to develop, but where the new activity does not yet create a steady income?

This strategy for sampling ongoing internal ventures seems to have worked satisfactorily and has led to useful contributions (e.g., Dahlqvist, 2007; Dahlqvist & Wiklund, 2012). With adaptations, it should be possible to use with larger businesses as the screening sample. However, the two-step procedure, and in particular the phone contact, turned out to be very important. Because of the dual checks, many non-eligible cases could be eliminated after clarifying interaction over the phone. This indicates that a single-point screener would need some refinement.

The approach shares some of the issues identified with regard to sampling emerging independent ventures. The procedure is somewhat costly. In our case, 4950 firms were contacted for a yield of only 250 eligible cases; however, this would improve if larger firms were sampled. This said, problems of identifying relevant respondents and selecting a focal venture if several were under way would be increased if larger firms were interviewed. In our study, it made sense to assume the CEO had all the relevant information. A study starting from a sample of large firms would either have to give up ambitions toward statistical representativeness or develop a procedure for first locating a sufficient number of relevant informants representing different roles in the company. Further, above a certain firm size, almost every sampled firm would likely have more than one new internal venture under way, calling for a sophisticated procedure for choosing among them or—if several ventures per firm are included in the sample—techniques for adjusting for statistical dependence between cases with the same origin.

Something which is probably more of an issue with small (and independently owned) firms than with large ones is whether the new venture is going to form part of the original firm or become a legally separate business. These two possibilities should be acknowledged in the design of the study and considered in the analysis. I personally see no reason to decide a priori to include only one type or only the other. On the contrary, this choice of "mode of exploitation" (Shane & Venkataraman, 2000; Wiklund & Shepherd, 2008) can be an interesting research question in itself.

In summary, sampling emerging ventures at an early stage can be cumbersome and costly, but it is also a way to get to the heart of entrepreneurship, which is possible and important (Shane, 2012). For those researchers who wish to make important future contributions, there is a solid grounding to start from, and technological developments would seem to bring important next steps within reach. Go for it!

5.4 Sampling Firms

After this Golgotha walk of sampling such elusive entities as emerging new ventures, many a cautiously natured reader are likely to have turned a deserter, already halfway to the safe haven of conventional, firm-level study. Compared to "emerging new ventures," sampling firms should be a piece of cake, right? Wrong! You ain't seen nothin' yet! Although the firm apparently remains the most common sampling unit in entrepreneurship research (Brush, Manolova, & Edelman, 2008; Chandler & Lyon, 2001; Davidsson & Wiklund, 2001), this is *not* because the firm is the most relevant or the most unproblematic level of analysis. The reason why it is so often selected while so rarely analytically dissected is simply, I would argue, that we are blinded by our conventions.

One rather common use of the firm as level of analysis in entrepreneurship studies used to be a wish to study their own emergence. For reasons discussed above, this retrospective approach is inferior to the concurrent study of processes of emergence. What I have in mind for the present section is instead the use of firm-level study in order to study entrepreneurship within or by the established firm ("intrapreneurship," "corporate entrepreneurship"), such as the launching of new products, entering into new markets, etc. (Naldi & Davidsson, 2014).

Although usually regarded a microlevel unit, the firm is (often) already an aggregate of different decision-making individuals and business activities. In discussing the sampling of such aggregate units, starting with firms and continuing with industries and spatial units, I will organize the discussion around the *relevance*, *size*, *size distribution*, and *heterogeneity along other dimensions* of the units to be sampled. I will let the longer discussion of relevance wait till last and start with a combined discussion of the other three criteria.

5.4.1 Size, Size Distribution, and Heterogeneity Along Other Dimensions

This entity we call "firm" comes in a variety of sizes from part-time, home-based businesses with minuscule sales to multinational giants with hundreds of thousands of employees and a budget larger than the GDP of many a small nation. As researchers, we have to handle this variability; otherwise, we risk working with inferior data and comparing apples with oranges. As regards the absolute *size*, the question to ask about minimum size is: *Will the sampled units be big enough for the entrepreneurial behavior we are investigating to have sufficient likelihood of occurring within the studied time frame?* As regards the *upper* limit for absolute firm size, the question instead becomes: *Can we obtain reliable data on entrepreneurship within this entity with the data collection method we are going to use?*

If most of the sampled firms are so small that what we operationalize as entrepreneurship (e.g., launching a new, internal venture) almost never happens, we will end up with a dependent variable with limited and very erratic variance. As a result, we are likely to get weak and confusing results. Although the reasons for doing so were not necessarily as well articulated at the time, I have in fact used a minimum size criterion in all the firm-level studies I have been involved in. For example, in my dissertation project *Continued Entrepreneurship and Small Firm Growth* (Davidsson, 1989a, 1989b, 1991), it was set at two employees. In the *Entrepreneurship in Different Organizational Contexts* study (Brown, Davidsson, & Wiklund, 2001), we used ten employees, and in our study of *High-Growth Firms* (Delmar & Davidsson, 2006; Delmar, Davidsson, & Gartner, 2003), we set the minimum at 20 employees.

In the *1994 Start-Up Cohort* study (Dahlqvist et al., 2000), the minimum criterion was that there was proof that the firm was commercially active, as indicated by registration as employer and/or for sales tax and/or corporate tax. As a result, this sample includes many very small firms. This came to illustrate the problem of insufficient minimum size when screening for cases for our *New Internal Ventures* study (Dahlqvist, 2007; Dahlqvist & Wiklund, 2012). As reported above, only about 250 new internal ventures were found in a screening of close to 5000 firms. The issue of minimum size, then, overlaps with the question of *relevance* (cf. below). In order to have relevant variation in the dependent variable (i.e., some aspect of entrepreneurship), the firms in the sample may have to be of a certain size. Alternatively, the time span for which to report entrepreneurial behaviors can be extended, but in concurrent studies this increases time and cost, and in retrospective studies it aggravates the problem of bias from hindsight and memory decay.

Let's now return briefly to the upper limit for absolute firm size. Using secondary data, the possible choices are restricted by how the provider of the data organizes the information. Indicators of corporate entrepreneurship activity—such as filing for patents, registration of new establishments, etc.—may be linked to a certain level of aggregation in a corporate hierarchy of establishments and companies (cf. below). However, as long as the ownership links between establishments and companies in a corporate hierarchy is known, the researcher can aggregate the data to any higher level than the original, according to her preferences. The problem with large absolute size is actually worse when primary (interview or questionnaire) data are collected. Not only does increasing size make it more unlikely that a single respondent can adequately report for the entire firm, it also becomes increasingly unlikely that the CEO is willing to participate in the study.[2] For this reason, I have never used a single informant for units larger than 250 employees, and I strongly recommend using multiple informants and other means of triangulation for firms this big or bigger. It is actually desirable for much smaller units as well (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003).

As regards the *size distribution* problem, we noted above that a simple random sample of "firms" would be dominated by entities that are tiny in size. In order to ascertain representation of somewhat larger firms, one could stratify the sample by size to avoid results being totally dominated by the micro-firms. In relation to most research questions, the researcher should probably give up aspirations to achieve exact statistical representativeness across all industries. It is simply not desirable. Instead, we should work with samples that reduce unwanted heterogeneity or those that *acknowledge* the heterogeneity by having a reasonable and balanced representation of different kinds of valid empirical manifestations of the theoretical concept "firm." In order to make both full sample and subsample analyses meaningful, we should also *limit and control* the heterogeneity.

Therefore, I recommend that samples on the firm level be either narrowed down to a more homogeneous category of firm or stratified along relevant dimensions so as to represent several such more homogenous categories. Hopefully, in the latter case, the strata can collectively be regarded as a valid representation of the theoretical "firm" concept in all its richness. Stratification is the main strategy I have followed in my own research. Returning to the studies referred to above, the *Continued Entrepreneurship and Small Firm Growth* study was stratified by size class (2–4, 5–9, and 10–19 employees) and industry (manufacturing of metal products and machinery, manufacturing of high-tech products, repair services, and retailing in clothing and home equipment). The *Entrepreneurship in Different Organizational Contexts* study was stratified along three dimensions, viz, size class (10–49 and 50–249 employees), industry (manufacturing, knowledge-intensive services, retail and wholesale, and other services), and governance (independent, part of company group with less than 250 employees in total and ditto with 250 or more employees). In my more recent work on *Growth and Profitability*, the two data sets were also trimmed and/or stratified by size and industry (Davidsson, Steffens, & Fitzsimmons, 2009; Steffens, Davidsson, & Fitzsimmons, 2009). In all of those cases, the stratification served us well in terms of subsample comparison and validation of results (Brown et al., 2001; Davidsson et al., 2009; Wiklund, Davidsson,

[2] I can tell those who believe they collect mail/email/online survey data from CEOs of large firms that probably they do not. Long before I became a researcher, I learnt from my father who really filled out the questionnaires addressed to the CEO. At the time he had an idiosyncratic position as speechwriter and expert on business cycles as well as communist block barter trade—and questionnaire filler—for the CEO of a multinational (Sandvik AB). Yes, I sometimes trust samples of one!

& Delmar, 2003). The *Business Platform* study (Davidsson & Klofsten, 2003; not mentioned above) used a narrow sample of technology- and knowledge-based firms located at Swedish technopoles.

To sum up, when sampling firms, I have argued that the units must be large enough for the investigated aspect of entrepreneurship to have a reasonable likelihood of occurring but not too large for obtaining reliable data through the chosen data collection method. Heterogeneity in size and along other dimensions ought to be both acknowledged and/or controlled, so that the resulting sample adequately reflects the theoretical concept of "firm" while at the same time it should make it possible to apply the meaningful operationalizations that are required in order to arrive at strong findings. Carefully thought through stratification, then, is the key to successful sampling of firms for entrepreneurship studies.

5.4.2 Relevance

If you thought we were done I can ensure you that the fun has just begun! *What is an empirical "firm"*? Well, we do have a few alternatives to choose from. What *legal forms* should be included? Limited liability companies (of which there are several different types in many countries) only? Partnerships? Sole proprietorships? All of the above? Well, there are also foundations, cooperatives, and various other types of associations that can be commercially active. What legal forms should be included may differ from country to country and from study to study. In the firm-level studies, I have conducted I have sampled either limited liability companies solely, or added also partnerships and sole proprietorships, in order to arrive at a theoretically relevant and workable sample.

Sometimes researchers use a unit that is not even a legal entity to represent "the firm," namely the *establishment*. An establishment is a place of work. In a retail chain, each outlet would be an establishment whether or not it was also a separate company. In the manufacturing industry, each geographically separated plant is an establishment. The reason researchers turn to establishment data is almost without exception practical considerations rather than the question of theoretical relevance. In the first part of the *Business Dynamics in Sweden* project (Davidsson, Lindmark, & Olofsson, 1994a, 1994b), we felt we had to use establishment data for data quality reasons. Establishment-level data tend to be more current and reliable, and establishment codes were not changed as easily as were company codes, so establishments were less subject to artificial births and deaths due to mere reregistration. In the second part of the project (Davidsson, Lindmark, & Olofsson, 1996, 1998), we found a way to aggregate establishment data to the theoretically more relevant firm level. However, this project used the region (and to some extent industry and nation) as the level of analysis, so the question of defining "firm" was an operationalization issue rather than a matter of sampling.

There are at least two other empirical structures that researchers use when they sample "firms." These are the company (or enterprise) and the company group (or multicompany corporation). Figure 5.3 illustrates some of the complexity of the matter—but only some.

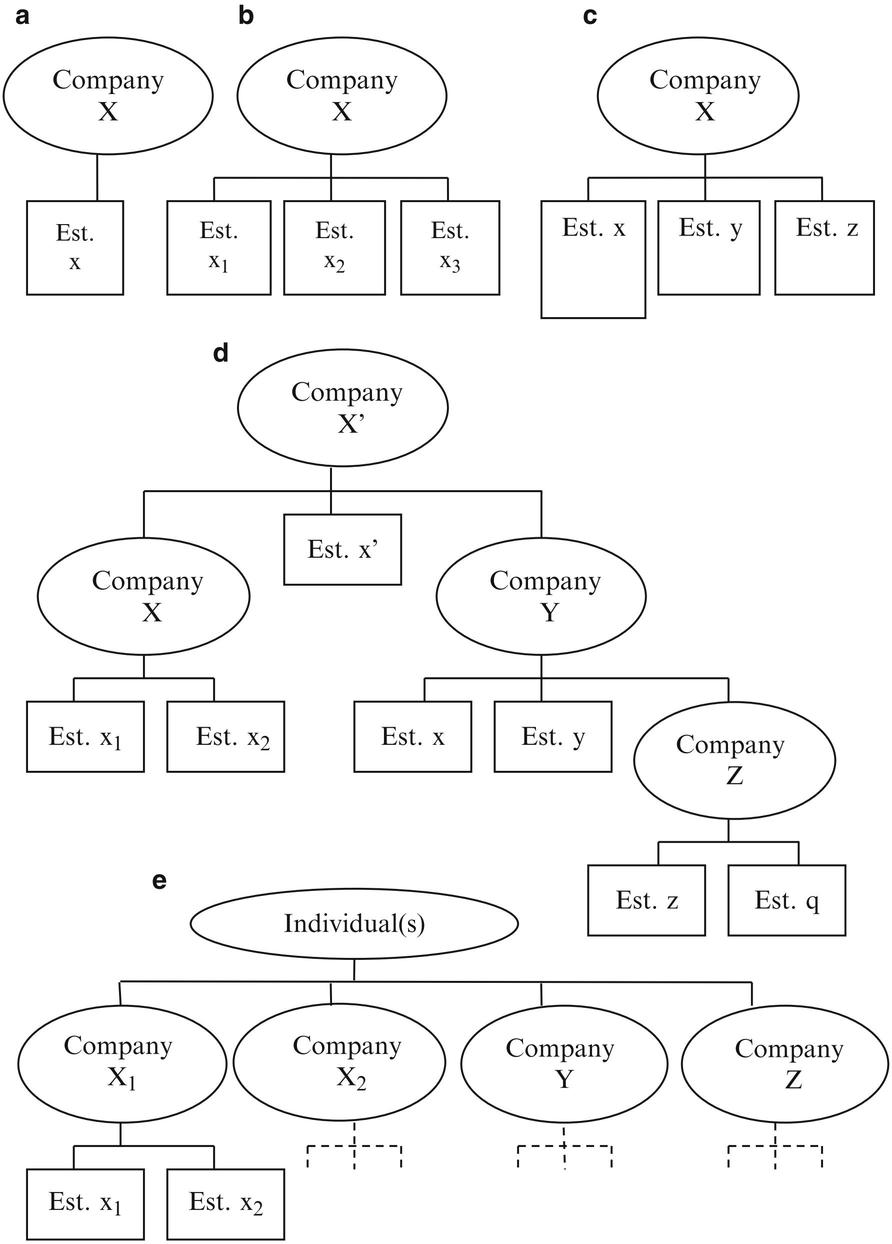


Fig. 5.3 Hierarchies of possible "firms"

The first case, (a), is an easy one. This is the single-establishment firm. Whether we sample establishments or companies, we end up with the same entity. With exhibit (b), we start to make things complicated. Here, we have an independent company with three establishments. In this case, we assume that all three establishments are active in the same industry (that's why they share index "x"). What is here the firm—the company or each establishment separately? As will be argued below, the answer can be contingent on what specific conceptualization of "firm" we are working with. Suffice it here to note that it is not a simple matter to determine what the right entity is, because behind this sketchy representation we can find different realities. The three establishments could be three branch offices of a consultancy business, not operating under a strong brand name or common concept but all enjoying a great deal of freedom and relying on local knowledge and contacts. Or it could be three semi-independent outlets in a retail chain, each with its own manager but working under the same brand name and the restrictions of strong company policies regarding marketing and purchasing, for example. Alternatively, it could be three production plants in a manufacturing firm, producing to order and with a minimum of decision-making discretion.

Case (c) is similar to (b) but the establishments are here active in different industries, according to their industry classifications. While classified on the firm level by its dominant industry (X), this company is in fact a conglomerate, operating in several industries. What is the firm here? Perhaps theory can guide us? Does the theory conceive of the firm as a power structure, or as an entity whose *raison d'être* is an aptitude to perform a certain type of activities? We will return to theoretical conceptualization of the firm later on, and the issue of matching theoretical and empirical firm definition.

With panel (d), we approach the reality of large corporations, although the figure merely portrays the principles of a much greater complexity. This structure combines the issues discussed so far. Is it the individual establishments or the companies that are firms? Should the "firm" concept be used for an entity with a logically coherent set of activities or can any disparate conglomerate under the same boss/owner(s) qualify? Is a unit like company Y independent enough to deserve the "firm" label, or is it but a mechanical servant of higher or lower levels in the hierarchy? The structure in (d) also adds another possibility: to use the entire structure that appears under the same ownership as the entity called "the firm." This may be suitable for some purposes but adds, among other things, the problem of determining how to deal with partly owned units further down the hierarchy.

In the *High-Growth Firms* study, the choice between company and company group level was particularly tricky. The main (policy) purpose of the project was to investigate the prevalence and job contributions of rapidly growing firms in the Swedish economy, preferably as compared to other countries. The establishment level could easily be ruled out as irrelevant. There was no policy interest in that level and prior claims that a tiny *x* percent of the business population (so-called gazelles) created a massive *y* percent of all new jobs clearly referred to the firm level (Henrekson & Johansson, 2008). Regardless of whether we choose the company or the company group level, we would risk serious underestimation of the prevalence and job contributions of high-growth firms. Growing *companies* are likely to eventually form *company groups*—which a single-level study would not note. If, instead, we choose the company group level, we would likely miss some spectacular growth firms within corporations because other parts of the company group were shrinking or divested during the same time. Ericsson (remember Ericsson phones, anyone??) was a case in point during the period studied (1987–1996)—while mobile phones and systems skyrocketed, other parts of the corporations shrunk quite dramatically. Our solution was to create data sets on both levels, as well as some ability to analyze across them (Davidsson & Delmar, 2003).

Panel (e) illustrates the case where an individual or a team owns a series of separate businesses, so like in (d), there is common ownership and ultimate control of all levels. However, there is no cross-ownership within the group, and therefore the entire group will not turn up as one "firm" in any sampling frame. Is this a problem? Well, if we accept the entire group as a firm in the (d) structure, why shouldn't we in the (e) structure just because the owners have chosen the latter form of internal organization of their empire? Although this problem may not be our biggest concern, I know for sure that it exists. For example, we noted in the sampling (of companies) for the *Entrepreneurship in Different Organizational Contexts* study there was at least a couple of very successful entrepreneurs who got into the sample repeatedly precisely because they had structured their company groups in this way.[3]

So, there are several answers to the question: *What is an empirical firm?* Now let's turn to the parallel question: *What is a theoretical firm?* Again, there are plenty of suggestions (see e.g., Coase, 1937; Conner & Prahalad, 1996; Cyert & March, 1963; Daft & Weick, 1984; Foss, 1993; Mueller, 1972; Seth & Thomas, 1994; Wernerfelt, 1984, 1995; Williamson, 1999). There is thus no shortage of ideas about what a firm is, and some of this conceptual work has actually been undertaken from the perspective of a particular interest in entrepreneurship (Alvarez & Barney, 2004; Foss & Klein, 2012; I. Zander, 2007). The problem is that these different conceptualizations highlight different aspects of "the firm" and therefore only partially overlap one another—and the empirical "firm" definitions we have discussed above. Sometimes the concept "firm" is used for different entities that it is difficult to see much common ground at all. "'The firm' is not a firm," as Edith Penrose (1959) put it, referring to the difference between real-world commercial organizations and "the firm" in microeconomic theory.

It would certainly be pretentious of me to claim sufficient mastery of theories of the firm for providing a fully informed discussion of the matching of theoretical and empirical delineations of "the firm" in entrepreneurship studies. I hope that the

---

[3] Luckily, during my student days many years ago, I had worked one summer for one of these guys who had been sampled half a dozen times, when he was setting the foundations for his hotel empire to be. He remembered my name when he got the cover letter and therefore generously shared his time when he was later contacted by an interviewer. There are many odd ways to minimize nonresponse! However, we were sensible enough not to have him go through the same questions six times.

admittedly crude treatment below can be of some direct assistance but also that highlighting the problem can inspire researchers to consult more solid sources as they think this through more carefully.

If we start with microeconomic theory and similar conceptualizations, we find a firm that is portrayed as a *production function* or as the cost structure of a specific production process. In the base model, the firm produces only one type of goods. If this or other theories with a similar "firm" notion are used, I would argue that the empirical entity that best matches this conceptualization is the *establishment*. This is an entity that is aligned with an assumption of one or a limited number of outputs, and little freedom to make its own decisions. By contrast, this type of firm concept does not seem to align well with the more complex empirical structures in Fig. 5.3 and, therefore, seems to be a mismatch.

Conceptualizations emphasizing the firm as a unique bundle of resources, knowledge, or routines (Barney, 1991; Kogut & Zander, 1992; Nelson & Winter, 1982; Penrose, 1959; Wernerfelt, 1984) are also difficult to match with conglomerate firms such as structures (c, d, e) in Fig. 5.3. Resources, knowledge, and capabilities are relevant for performing certain types of tasks. They are usually not resources or routines that make the firm good at just anything. Although much closer to the realities of real-world companies and corporations, these theories seem to share with microeconomics a rather narrow view on what "a firm" offers on the market. However, strategic theories assume a degree of discretion in decision-making, which is not necessarily present within establishments that form part of larger structures. Ecological and evolutionary theory, which also emphasize distinct competence as the firm's reason for being, put less emphasis on the deliberate actions of the decision-maker and are therefore more compatible with establishment-level analysis. The best match for strategic knowledge- and resource-based theoretical perspectives seems to be firms like (a) and (b) in Fig. 5.3. For the other types of structures, one would like to cut out the different parts—perhaps "strategic business units"—of the total structure that make logical units from a resource- or knowledge-based perspective. This, of course, poses quite a difficult sampling challenge for "large n" studies. Nonetheless, taking on that challenge may be a prerequisite for arriving at valid tests of the theory.

Although externally rather than internally oriented, Porter's (1980, 1985) strategic theory seems to share the same type of matching problem. Much of Porter's theorizing is about industry attractiveness. Hence, "firms" are assumed to operate in *an* industry and apply *a* strategy that is suitable for maintaining or improving the firm's position in that industry. This does not seem to be a theory that works for empirical entities with a high degree of both horizontal and vertical heterogeneity. And therein lies the core of the matching problem for the strategic theories we are discussing. The problem arises when the theory is used for making predictions and discussing implications for *the entire structure*. If the internal heterogeneity of the sampled units is carefully considered in the design and derivation of hypotheses, the match between these theories and "firms" like the structures (c) and (d) in Fig. 5.3 can be quite good. That is, in many cases the dependent variable should perhaps not be the profitability or growth rate of the entire organization but instead focus on *what parts* of the organization will grow and yield a surplus.

Theories that emphasize transaction costs or agency problems, governance and power issues, behavioral firm theory, and theories emphasizing organizational structure and incentive systems (Coase, 1937; Daft, 1983; Jensen & Meckling, 1976; Williamson, 1975) seem to be less sensitive to heterogeneity concerning the firm's outputs and markets. They can therefore probably be meaningfully applied to different levels of "firm"—establishment, company, or the multicompany corporation—as long as these different creatures are not mixed too indiscriminately in the same sample. An issue for some of these perspectives would be whether a unit like company Y in Fig. 5.3 (d) is independent enough to qualify as a firm and whether it can be included alongside entities like company X in (a) or (b) without carefully distinguishing between the different hierarchical positions when making predictions and interpretations. Another concern that goes for these types of conceptualizations as well as for all those previously discussed is whether it can be reasonably assumed, as the type of empirical firm grows in size and internal complexity, that the entities sampled have any consistent characteristics throughout the organization along those dimension highlighted by the theory. For example, when a researcher wants to investigate whether organizational structures and incentive systems influence the occurrence and success of internal venturing in a firm, can it then be meaningful to work with "firms" that have thousands of employees working in spatially and legally separated companies that are also producing different types of products for likewise different types of customers? My spontaneous answer would be that most questions about entrepreneurship at the firm level are better researched in smaller and more homogenous units, be they independent firms or parts of a larger structure.

To sum up, this section has shown that sampling on the firm level is far from unproblematic. This is a reason for researchers who want to make a unique contribution to consider alternative levels of analysis. If the firm level is chosen, the different theoretical and empirical definitions of "firm," and especially the match between them, ought to be taken into careful consideration.

5.5 Sampling Industries (or Populations)

As a bridge between firm and industry levels, let me first discuss the increasing popularity of *single-industry studies.* Such studies are not industry *level* in the sense of comparing industries or estimating effects of industry characteristics, but they involve an industry sampling problem: is the industry context chosen for the single-industry study a relevant setting for testing a particular theory, or for generalizable, empirical fact-finding about the broader phenomenon of entrepreneurship?

Single-industry studies make a lot of sense for reasons discussed in Chaps. 3 and 4. If the industry in question is one example of a context where the theory is supposed to be valid, then the industry is suitable for a theory test: If the theory is any good, its predictions should come true in the selected industry. Further, a lot of heterogeneity that is irrelevant to the theory is designed away. In addition, the narrow context may allow better, custom-made operationalizations of theoretical constructs. All this would point to a more reliable test of the theory (Shugan, 2007).

On the other hand—and to build on two of my favorite, single-industry studies—few of us have any particular interest in *North American architectural woodwork firms* (Baum & Locke, 2004) or *law firms created in the Greater Vancouver area 1990–1998* (Cliff, Jennings, & Greenwood, 2006). It is not *the* particular context that makes us take delight in these studies. Rather, it is the choice of *a* suitable context, into which they also have particularly deep insights, that allows the researchers to please us with exemplary studies.

Contrasting the two studies also illustrates differences in likely boundary conditions. Baum and Locke (2004) study the influence of psychological characteristics of founder-owners on the growth of the firms. It is not immediately evident that these relationships should vary markedly by industry context. Therefore, we may be inclined to infer broad applicability of the results, although replications in other contexts would be necessary for our ultimate conviction. Cliff et al. (2006) face a greater challenge in this regard. Their study asks the question whether insiders or outsiders to an industry are more innovative (cf. Schumpeter, 1934). Here, at least I would be hesitant to make broad generalizations. Wouldn't the answer likely be contingent on the type of industry? And also on the type of innovation? Cliff et al. (2006) study *organizational* innovation and signal confidence that their results would apply also to "other mature, highly institutionalized professional fields"—and in this case, it is perhaps within these parameters one should dare make tentative generalizations.

Recently, I've been involved in entrepreneurship research in the mining sector (Bakker, Shepherd, & Davidsson, 2014; Sakhdari, Burgers, & Davidsson, 2014). I have to confess that the first time I came across the suggestion to focus on this sector, I wasn't thrilled. I had no particular interest in mining, and it seemed a dubious or at least highly atypical context for studying entrepreneurship issues. It is dominated by multinational giants and seemingly an industry with pretty well-established routines. It is extremely capital intensive compared to most other sectors, with no chance at all of generating revenue before very large upfront investments have been made. Wouldn't seem the ideal setting for a study of effectuation (Sarasvathy, 2008) or entrepreneurial bricolage (Baker & Nelson, 2005; Senyard, Baker, Steffens, & Davidsson, 2014).

But I warmed up to it. The context actually has its strong points. First, it is a relatively neat industry setting—it is pretty clear what different actors in the industry are supposed to do. Somewhat akin to highly abstracted theory or a computer simulation, we can trust the actors have comparable tasks and task environments. Second, it is (somewhat ironically, given my position on "opportunities") a context where this odd idea of objective, preexisting and actor-independent "opportunities" (Eckhardt & Shane, 2010; 2013; Shane, 2012; Shane & Venkataraman, 2000) actually makes some sense.[4] In mining, "opportunities" can be represented by not-yet-discovered

[4] However, Chap. 8 will establish that I remain firmly unconvinced that "objective opportunity" is an empirical entity we should try to sample and study.

(or exploited) ore bodies, which undoubtedly exist before they are found and which have value that one could argue guarantees an above-zero probability of successful exploitation (Shane, 2012). Third, mining progresses through comparatively well-defined stages. We found this to facilitate theorizing about how antecedents differentially influence further progress at different points in the process. Although processes in other industries are not equally easy to subdivide, the simpler example of mining may inspire us to develop complex criteria for stage transitions also in other settings, allowing theorizing and testing that is sensitive to process stage (Bakker & Shepherd, 2015).

In short, the value of single-industry studies will be contingent on how well the chosen context suits the task at hand. To achieve a good fit one can alter the chosen context—or the task.

Broadening the view now to real industry-level studies, I have not used the industry as the primary sampling unit in any of my studies, although in *Business Dynamics in Sweden* it was of the variables by which the data set could be aggregated. So my expertise in this area is limited. This said, many of the issues related to sampling industries fit under the headings *relevance*, *size*, *size distribution*, and *heterogeneity along other dimensions.* I will therefore reuse these organizing categories here.

5.5.1 Size, Size Distribution, and Heterogeneity Along Other Dimensions

Industry statistics typically use a hierarchical classification system. These standards are similar across countries, but in order to make life interesting for researchers, they are not identical. In order to make life *really* interesting for researchers involved in longitudinal research—and to reflect real changes in the economy—the systems are also revised periodically. At the crudest level, these systems subdivide the economy into about ten categories, such as *primary industries*, *manufacturing*, *wholesale* and *retail*, *education* and *health-care services*, and other broad categorizations like these. These industries are then successively subdivided down to a five- or six-digit level. For example, according to the North American NAICS system, a firm that produces nuts and bolts is included in the industry aggregate on all of the following levels:

3	Manufacturing
33	Metal manufacturing
332	Fabricated metal product manufacturing
3327	Machine shops; turned product; and screw, nut, and bolt manufacturing
33272	Turned product; and screw, nut, and bolt manufacturing
332722	Bolt, nut, screw, rivet, and washer manufacturing

This gives the researcher a great deal of freedom of choice as regards what level of aggregation should be used. Often the final decision does not have to be made at the design stage. If data are collected at finer levels of disaggregation, they can always be aggregated later, whereas the converse is not true.

Similarly to the firm level, the design question to ask concerning minimum absolute *size* of the industry units in the sample is: *Will the sampled units be big enough for the entrepreneurial behavior we are investigating to have any likelihood of occurring within the studied spatial unit and time frame used?* For example, the total number of start-ups, patents, or other indicators of entrepreneurship per annum in Nova Scotia may not be very high in 332722. Again, if the studied units are too small, variation in the dependent variable will appear stochastic and hard to explain. If, on the other hand, too high a level of aggregation is used, we will run into other problems. First, we may simply end up with too few industries to compare in order for the analysis to yield interesting results. Second, very different types of firm will be assigned to the same industry, which is a problem of (internal) *heterogeneity* and therefore a threat to *relevance* relative to the theoretical industry concept being employed. Thirdly, expanding and shrinking subindustries or niches may cancel out within very broadly defined industries.

As regards *size distribution*, we again run into a similar problem as on the firm level. Assume that we decide to work with industries on the three-digit level. Some of the resulting industry categories may have thousands of firms in them, whereas others only contain a few dozen. Should they weigh equally in the analysis? Perhaps yes, perhaps not. As noted above, the fact that a particular category is rare or numerous in a particular country at a particular time does not necessarily mean it should be granted greater or lesser theoretical significance. The researcher should at least make an informed decision. If for some reason, more equally sized industries are deemed desirable, it may be worth forming categories that in some cases are on the two-digit level and in other cases on the three- or four-digit level.

As regards *heterogeneity* across industries, we need to ask, again, whether we are about to compare apples with oranges. Some of the questions to ask oneself are:

– Does our theory apply to all industries?
– Do all the industries that result from application of the standard industry codes yield categories that correspond to the same theoretical "industry" (or "population") construct?
– Do the operationalizations of variables that we plan to use work for all industries, and are values on those variables meaningfully comparable across those industries?

As I see it, a big threat here is that our conceptualizations—sometimes explicitly, but even more so implicitly—are modeled on the manufacturing industry. Time and again in my research I have come across instances where operationalizations fit better and results were stronger for the manufacturing subsample, presumably because the thinking behind the research and the tools used in it were "manufacturing biased." In this day and age, the manufacturing firm really should not be the implicit model, but I suspect that sometimes it still is. So this is a pitfall worth watching out for.

5.5.2 Relevance

The questions asked above link to the issue of relevance. Also regarding relevance, it is not a given that simple aggregation according to the industry hierarchy used in the standard system will leave us with industry categories that are maximally relevant. Depending on our theory and research questions, we may want to compare young vs. mature industries, contrast growing with contracting industries, compare entrepreneurial activity for industries with high vs. low entry barriers or which are in a business-to-business vs. a business-to-consumer situation, get special insights into industries that are research- or knowledge-intensive, or investigate the effects of firm size structure and capital intensity on innovative or entrepreneurial activity on the industry level. In order to arrive at industry groupings suitable for these types of contrasts, one might want to combine subindustries that, according to the standard classification, belong to different main groups. For example, for *Business Dynamics in Sweden*, we created the following industry categories:

1. High-tech manufacturing
2. Wood-based, paper, and pulp manufacturing
3. Engineering industries
4. Mining and steel manufacturing
5. Other manufacturing
6. Technology-related services
7. Other knowledge-intensive services
8. Financial services
9. Construction
10. Accommodation and food services
11. Wholesale and retail
12. Transportation and communication
13. Other services
14. Education and health care
15. Agriculture, forestry, and fishing
16. Government sector

Of those, only industries 9, 11, 12, and 15 are aggregated strictly by the logic of the original standard classification. All other categories were more or less customized for our purposes. Depending on the specific context, we sometimes worked with even more aggregated industry sectors. For example, industries 1–5 were combined to "manufacturing," 6–8 to "knowledge-intensive services," etc.

My final issue concerning relevance has to do with the concept of "population" (of species of organizations, cf. Aldrich & Ruef, 2006). We noted in Chap. 3 that empirically, membership of a population is often equated with having the same industry code, and "populations" then become equal to what industrial economists call "industries." There are two problems with this approach. If it is possible at all to achieve a good match between the theoretical and empirical concepts of "population," it is probably only possible at very fine levels of detail (five- or six-digit groups). We may then quickly run into problems related to small size (cf. above). Second, from an entrepreneurship perspective, the most interesting questions are related to how truly *new* species and populations come into being. The problem with this is time lag in the industry classification system. When McDonald's started to revolutionize the fast-food industry, there was probably no unique code for "(franchised) fast-food restaurant belonging to a chain," and during the Internet boom, there was no unique code to assign to "dot-coms." Similarly, I suspect ventures developing "apps" or games for online devices do not appear in easily identifiable classification categories. Therefore, other approaches may be more relevant for the study of emerging populations. For an example of a study that covers the emergence, growth, and disappearance of what can truly be called a distinct population of organizations, see Gratzer's fascinating study of the automated restaurant industry (Gratzer, 1996). Unfortunately, only fractions of this exemplary study are available in English (Gratzer, 1999).

Throughout this subsection on sampling of industries, the assumption has been that industry codes in secondary data sets somehow be used. When collecting primary data from firms, it is, of course, possible to collect any information relevant to one's theoretical concept of industry, which can then be used for post-stratification of firms into industries.

5.6 Sampling Spatial Units

It is a very common occurrence that the first research proposal presented by international Ph.D. students or applicants at my current institution (Queensland University of Technology in Brisbane, Australia) involves some idea of international comparison. Specifically, the prospective student plans to compare two countries: their original home country and their country of study, however, rarely if ever do they present a strong theoretical rationale as to why it is interesting to compare these two countries. In other words, the suggestion is completely driven by individual curiosity and convenience. That's not how we make our research interesting to others and hence not how we make it into prestigious scholarly journals.

It can work, though. Recently, a student of mine—Kamal Sakhdari, who is Iranian—chose to focus on the mining equipment, technology, and services (METS) sector in Iran and Australia. It is not an irrelevant choice, considering the importance of the mining industry in those two countries. Further, he built on the fact that these two countries were close to opposite extremes on "institutionally market-oriented context" according to objective data. Based on this, he derived interesting hypotheses and arrived at likewise interesting results regarding the moderating effects of this contextual variable for corporate entrepreneurship (Sakhdari et al., 2014). The lesson? As with any research, there ought to be a *theoretical* reason for the case selection and sampling we undertake.

A lot has happened in the last decade regarding theory and analysis approaches in research on spatial units. I am not sure the development has left me completely behind on the issue of sampling, though, so I dare take the risk of leaving this section in, with some editing. Two of the major projects I have been involved in—*Business Dynamics in Sweden* and *Culture and Entrepreneurship*—used regions as the cases in the data matrix. As both of these studies were comprehensive efforts, each involving the compilation of two successive data sets, I can claim much more experience here than for industries. Figures 5.4 and 5.5 give an idea of what the main research questions were in those studies. I will use my experiences from these projects in discussing "sampling" of spatial units, while blending in observations from other research as well. As this chapter is getting long already, I will try to be concise. Another measure to keep the reader awake is that here I will shift to discussing relevance first and then turn to absolute size, size distribution, and heterogeneity along other dimensions.

5.6.1 Relevance

Much spatially based work focuses on regions rather than countries. Most countries are spatially subdivided in a number of different ways. The first to come to mind may be municipalities and counties or close equivalents to those, but there are several other subdivisions as well. The rationales for these subdivisions tend to be (historical) administrative needs for political, judicial, military, or religious purposes. Apart from issues related to size and variability (below), there are important practical issues to take into account when choosing among these, such as for what units statistics are compiled and what units have been used in other studies that we

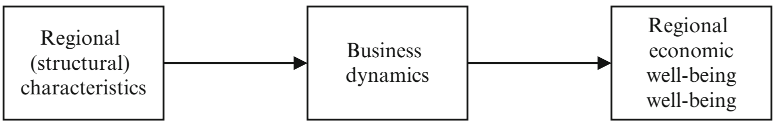


Fig. 5.4 Region-related research questions in Business Dynamics in Sweden

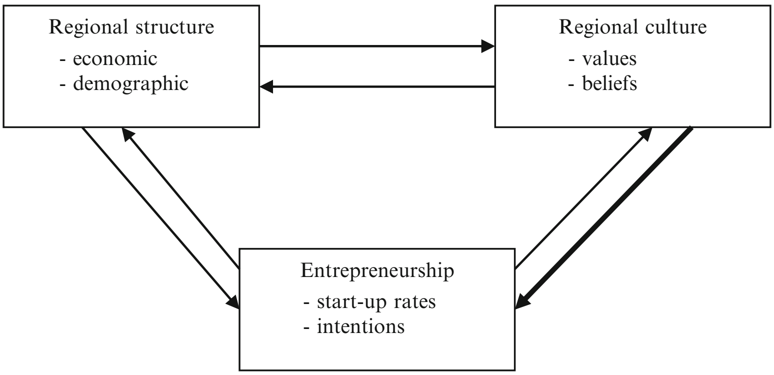


Fig. 5.5 Research questions in *Culture and Entrepreneurship*, main interest emphasized

want to compare our results with? Unless such practical concerns strongly suggest administrative units of this kind, other types of subdivisions may be better suited for entrepreneurship research. In short, we want to have spatial units that make sense from an *economic* point of view rather than being based on some administrative criterion. For example, following Reynolds's work in the USA, we used labor market areas (LMAs) as the regional units of choice for the abovementioned studies (Reynolds & Maki, 1990; Reynolds, Miller, & Maki, 1993). This type of spatial subdivision has been made also in a range of other countries. We choose LMAs because of the following, distinctive advantages:

- Being aggregates of municipalities, they form units for which statistics from a lower level of analysis exist and can be aggregated.
- Being based on travel-to-work statistics—municipalities with high levels of commuting between them are combined—it can be argued that they are natural economic entities.
- They are not artificially bound by county or (within-nation) state limits. When the commuting so suggests, LMAs can be defined across such borders.
- They can be clustered into *region types* on economic-structural criteria (Davidsson, 1995a; Reynolds et al., 1993), which gives additional input to design and interpretation of analyses.

It is not a given, though, that the LMA subdivision should be adopted as is. In *Business Dynamics in Sweden*, for example, we did not accept Greater Stockholm as one LMA because we knew that the northern parts of the capital were "hot" at the time while the southern parts fought with a dying industrial heritage and other problems. We therefore subdivided Greater Stockholm into three spatial units: North, Central, and South. In the second part of the study, we also subdivided the other two major cities, Gothenburg and Malmö, into separate center and hinterland units because we had reason to believe these parts were distinct regarding entrepreneurial activity and the structural characteristics that might affect it. Future researchers are advised to make adaptations of this kind based on whatever knowledge they have over and above the mechanical grouping of municipalities that results from the clustering based on travel-to-work data.

For the *Culture and Entrepreneurship* study, I had the little problem that no data were available on any level of analysis concerning the types of variables—prevailing values and beliefs—that I needed. Hence, I had to collect primary data from individuals and use the mean responses per region as the regional point estimates. It is probably not hard for the reader to imagine that collecting primary data from large enough representative samples of individuals in 111 LMAs is going to be prohibitively costly, to say the least. Statistics Sweden had clustered the 111 LMAs into *region types* based on structural criteria. The workable solution I found was to sample individuals representing these region types rather than LMAs. In order to capture variability and assure relevance, I drew samples from the following region types (see Davidsson, 1993, 1995a for structural descriptions):

1. Greater Stockholm
2. Regional centers
3. Average communities
4. Rural LMAs
5. One-company (industrial) towns
6. The Gnosjö/Gislaved industrial district

Based on the original clustering results, the nonselected types were suspected not to stand out from the selected ones on any dimensions. Category 6 is a special adaptation. Rather than sampling from the entire region type characterized by strong emphasis on small-scale manufacturing, I included what is Sweden's most famous industrial district. The essence of its fame is its alleged "entrepreneurial spirit" (Wigren, 2003). Including this region is obviously relevant in a study on *Culture and Entrepreneurship*. Again, this is a "manual" adaptation based on knowledge outside of the mechanical subdivision of spatial units, and I think researchers should be encouraged to make such adaptations. In a regional study of entrepreneurship in any country, one should, of course, try to include the regions that are the most interesting from an entrepreneurship perspective and delineate such regions as precisely as possible. This might include also the regions thought to have the biggest problems with low levels of entrepreneurship, which was the case with the category "one-company towns" selected for *Culture and Entrepreneurship*.

5.6.2 Size, Size Distribution, and Heterogeneity Along Other Dimensions

The problems relating to size, size distribution, and other heterogeneity for regions are similar to those discussed for firms and industries above. The original grouping of 111 LMAs in Sweden yielded units that ranged from a population of a couple of thousand to over a million. In order to increase minimum size and decrease size variability in *Business Dynamics in Sweden*, regions that were small (<10,000) in population and geographically adjacent were combined to larger units if they were also structurally similar. At the other end of the spectrum, we split the largest agglomerations into two or three units. As a result of these changes, we worked with 80 (83 in the second study) somewhat more equally sized regions rather than the original 111.

With *Culture and Entrepreneurship*, I faced two types of heterogeneity problems. First, for budgetary reasons, I had used region types rather than LMAs. Being aggregates of structurally similar but not necessarily geographically adjacent LMAs, the units may have become too large and internally heterogeneous to find any distinct cultural differences between the types. Second, with only six cases (region types) and considerable variation in three variable groups (structure, culture, and entrepreneurship), there was the (calculated) risk that the results would indicate that all the arrows in Fig. 5.5 had something to them, without the possibility of sorting out relative strength or dominant causal direction. The latter is exactly what happened (Davidsson, 1993, 1995a). Overall, there seemed to be a positive relationship between the prevalence of "entrepreneurial values" and regional start-up rates. However, the same variation appeared at least as explainable by structural variation. In short, where the structural (pull) conditions for (independent) entrepreneurship were favorable, the culture also tended to favor entrepreneurship. Recently, however, Obschonka et al. (2015) found direct effects of cultural variation as well as such variation moderating the effects of aggregate level differences in human capital. The same researchers have also provided insights into the possible, historical reason for regional variation in entrepreneurship culture (Stuetzer et al., 2015).

Because the design of the original study had led to possible dilution of regionally distinct cultures as well as confounding of structural and cultural explanations, I employed a different sampling strategy for the second study. The sampling criteria for this second study were (Davidsson, 1995b):

1. The regional units should be small enough so that cultural variation did not cancel out within them.
2. It should be possible to obtain data from them regarding relevant cultural and structural variables as well as on regional entrepreneurship indicators.
3. There had to be variation in entrepreneurship among them (here operationalized as regional start-up rates for independent businesses).
4. They should be as homogeneous as possible on variables other than the cultural variation in values and beliefs, which are the key explanatory variables.

In order to achieve (1), I selected LMAs rather than region types. Structural and entrepreneurship measures for LMAs were available from *Business Dynamics in Sweden*, whereas cultural variation had to be obtained through primary data collection from individuals (2). In order to fulfill (3) and (4), I cluster analyzed all LMAs in the *Business Dynamics in Sweden study* on those (seven) structural characteristics that according to a regression model had substantial influence on start-up rates. From the resulting clusters, I chose three matched pairs, where both LMAs in each pair belonged to the same structural cluster, whereas one had a much higher and the other a much lower start-up rate than predicted by the regression analysis. The logic was that unmeasured cultural variation might be the explanation for deviations from the values predicted by the structure model. After measuring the cultural variation, the conclusions were the following. The results were more for than against a separate, causal effect of cultural variation. However, with these more distinct regional units, the cultural variation still appeared small, relative to the structural variation in the country. For this reason, structural variation seemed to account for relatively more of the variation in regional entrepreneurship [in Sweden during the studied period] (Davidsson, 1995b; Davidsson & Wiklund, 1997). Although the small number of cases prohibited definitive conclusions, the structurally matched sampling procedure helped take the analysis much further than otherwise possible.

In the above example, information from samples of individuals was used to represent characteristics of spatial units. When this approach is chosen, the logic of statistical inference theory applies to the full. For us to conclude that the average value in the sample is representative for the spatial unit, we have to work with an unbiased sample of individuals from the relevant (sub)population within that spatial unit. Practical and budgetary concerns may make probabilistic sampling impossible, but when the individuals in the samples are drawn from a particular company, association, or educational group, and/or if nonresponse is high and uneven across spatial units, one should be aware that there are great risks that erroneous conclusions be drawn (cf. Hofstede, 1980; Lynn, 1991; Scheinberg & MacMillan, 1988). For example, if samples of MBA students are used across countries, one should be aware that this is a group that represents an extremely small social elite in some countries, whereas in other countries having or undertaking an MBA is not a reason for others to engage in eyebrow-raising exercises. Hence, resorting to this type of convenience sampling is likely to cause serious bias or distortions.

The issues of (relative) size and heterogeneity are also important concerns for studies on the country level. Countries are very different animals, and it can be validly asked whether, for example, causes and effects of internationalization (or of national competition, cf. Porter, 1990) can be meaningfully investigated and validly generalized across spatial units that are extremely different in terms of size and internal heterogeneity (think of, e.g., Australia, Croatia, Indonesia, Japan, Luxemburg, Singapore, Switzerland, and the USA). This, of course, is a very important issue for research on “international entrepreneurship” as well as other international-comparative work in our field.

In order to reduce size variability and other heterogeneity in international comparisons, Swedish researchers have contrasted Sweden with Ohio rather than with the USA as a whole (Braunerhjelm & Carlsson, 1999; Braunerhjelm, Carlsson, Cetindamar, & Johansson, 2000; Fridh, 2002). Sweden and Ohio are relatively similar in size and shared the same traditional industry structure. Therefore, comparing this “matched pair” should be a better ground than Sweden vs. the USA for comparing institutional factors and their effect on entrepreneurship and industrial renewal. Strategies like this appear worthy of following by researchers interested in other countries and regions as well.

Summing up, I have argued in this section that ideally, regional units that make economic sense should be used rather than administrative subdivisions. If labor market areas have been defined for the country in question, this can be a good choice. However, considerations of size, size distribution, and other heterogeneity, as well as the adding of prior knowledge about spatial variation in entrepreneurship-related issues, may call for adaptations of the regional subdivision offered by statistics providers. For studies comparing countries, the very large variation in size and other characteristics have to be taken into account. The more relevant comparison can sometimes be between an entire country and a region within another country. Finally, when using a particular subpopulation like MBA students or IBM employees, one should carefully consider whether this subgroup is equally representative of each country.

5.7 Sampling Other Units of Analysis

Chandler and Lyon (2001) as well as Davidsson and Wiklund (2001) show that published entrepreneurship research is dominated by studies on the individual and firm levels of analysis. Some use the aggregate levels industry or region. Very few use other levels of analysis. This does not mean they are less relevant. One very relevant but rarely used level of analysis has been treated rather elaborately above: the emerging new venture. But there are others that could be considered: *the patent*, *the invention/innovation*, *the team*, *the dyad*, *the community of practice*, *the network*, and *the cluster*, to name a few.

These alternatives share the characteristic that it is difficult to obtain a sampling frame and/or secondary data on them. So why make life difficult? Well, if you ask me you could just as well ask "Why make life interesting?" Challenges are fun! Besides, why shouldn't researchers believe their favorite strategic recipes for business success—be it Portnerian "diversification strategy" or an RBV emphasis on "sustainable competitive advantage through unique knowledge and capabilities"—have analogous applicability in research? Be different!

5.8 Summary and Conclusions

I opened this chapter by arguing that theoretical relevance is the most important criterion for sampling. The composition of units in the sample should match the theory used. Statistical representativeness is desirable when possible to achieve at all, but secondary to theoretical relevance. We have further observed that entrepreneurship research can be conducted on many different levels of analysis and that each level has its problems that have to be dealt with. It is not always possible to overcome those problems—there is no such thing as "perfect" research—but it is certainly worth trying to solve as many as possible and to be aware of the remaining shortcomings of one's sample.

Importantly, the most conventional levels of analysis in entrepreneurship research—the individual and the firm—are not markedly less problematic than are other alternatives. This insight should provide incentive for researchers to consider leaving the most trodden paths and apply other levels of analyses than those that first come to mind.

References

Aldrich, H. E., & Ruef, M. (2006). *Organizations evolving*. Newbury Park, CA: Sage.

Álvarez, C., Urbano, D., & Amorós, J. E. (2014). GEM research: Achievements and challenges. *Small Business Economics, 42*(3), 445–465.

Alvarez, S. A., & Barney, J. B. (2004). Organizing rent generation and appropriation: Toward a theory of the entrepreneurial firm. *Journal of Business Venturing, 19*(5), 621–635.

Amaral, A. M., Baptista, R., & Lima, F. (2011). Serial entrepreneurship: Impact of human capital on time to re-entry. *Small Business Economics, 37*(1), 1–21.

Amorós, J. E., Bosma, N., & Levie, J. (2013). Ten years of global entrepreneurship monitor: Accomplishments and prospects. *International Journal of Entrepreneurial Venturing, 5*(2), 120–152.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329–366.

Bakker, R. M., & Shepherd, D. A. (2015). Pull the plug or take the plunge: Multiple opportunities and the speed of venturing decisions in the Australian mining industry. Paper accepted for publication in *Academy of Management Journal*.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Baum, J. R., & Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology, 89*(4), 587–598.

Bergmann, H., & Stephan, U. (2013). Moving on from nascent entrepreneurship: Measuring cross-national differences in the transition to new business ownership. *Small Business Economics, 41*(4), 945–959.

Brandstätter, H. (2011). Personality aspects of entrepreneurship: A look at five meta-analyses. *Personality and Individual Differences, 51*(3), 222–230.

Braunerhjelm, P., & Carlsson, B. (1999). Industry clusters in Ohio and in Sweden. *Small Business Economics, 1*(4), 279–293.

Braunerhjelm, P., Carlsson, B., Cetindamar, D., & Johansson, D. (2000). The old and the new: The evolution of polymer and biomedical clusters in Ohio and Sweden. *Journal of Evolutionary Economics, 10*(5), 471–488.

Brown, T., Davidsson, P., & Wiklund, J. (2001). An operationalization of Stevenson's conceptualization of entrepreneurship as opportunity-based firm behavior. *Strategic Management Journal, 22*(10), 953–968.

Brush, C. G., Manolova, T. S., & Edelman, L. F. (2008). Separated by a common language? Entrepreneurship research across the Atlantic. *Entrepreneurship: Theory and Practice, 32*(2), 249–266.

Campbell, B. A. (2005). *Using linked employer-employee data to study entrepreneurship issues* (Handbook of Entrepreneurship Research, pp. 143–166). Berlin: Springer.

Chandler, G. N., & Lyon, D. W. (2001). Methodological issues in entrepreneurship research: The past decade. *Entrepreneurship: Theory and Practice, 25*(4), 101–113.

Cliff, J. E., Jennings, D. P., & Greenwood, R. (2006). New to the game and questioning the rules: The experiences and beliefs of founders who start imitative versus innovative firms. *Journal of Business Venturing, 21*, 633–663.

Coase, R. H. (1937). The nature of the firm. *Economica, 4*, 386–405.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 47*(12), 997–1003.

Conner, K., & Prahalad, C. K. (1996). A resource-based theory of the firm: Knowledge vs. opportunism. *Organization Science, 7*(5), 477–501.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall Inc.

Daft, R. L. (1983). *Organization theory and design*. New York, NY: West Publishing Co.

Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretations systems. *Academy of Management Review, 9*(2), 284–295.

Dahlqvist, J. (2007). *Assessing new economic activity: Process and performance in new ventures*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Dahlqvist, J., Davidsson, P., & Wiklund, J. (2000). Initial conditions as predictors of new venture performance: A replication and extension of the Cooper et al. study. *Enterprise and Innovation Management Studies, 1*(1), 1–17.

Dahlqvist, J., & Wiklund, J. (2012). Measuring the market newness of new ventures. *Journal of Business Venturing, 27*(2), 185–196.

Davidsson, P. (1989a). *Continued entrepreneurship and small firm growth*. Doctoral dissertation, Stockholm School of Economics, Stockholm.
Davidsson, P. (1989b). Entrepreneurship—and after? A study of growth willingness in small firms. *Journal of Business Venturing, 4*(3), 211–226.
Davidsson, P. (1991). Continued entrepreneurship: Ability, need, and opportunity as determinants of small firm growth. *Journal of Business Venturing, 6*(6), 405–429.
Davidsson, P. (1992). *Entrepreneurship and small business research: How do we get further? (BA-publications No. 126)*. Umeå: Umeå Business School.
Davidsson, P. (1993). *Kultur och entreprenörskap - om orsaker till regional variation i nyföretagande (Culture and Entrepreneurship - On the Determinants of Regional Variation in New Firm Formation)*. Stockholm: NUTEK.
Davidsson, P. (1995a). Culture, structure and regional levels of entrepreneurship. *Entrepreneurship & Regional Development, 7*, 41–62.
Davidsson, P. (1995b). *Kultur och Entreprenörskap - en uppföljning (Culture and Entrepreneurship - A Follow-up)*. Örebro: Stiftelsen Forum för Småföretagsforskning.
Davidsson, P. (2006). Method challenges and opportunities in the psychological study of entrepreneurship. In J. R. Baum, M. Frese, & R. A. Baron (Eds.), *The psychology of entrepreneurship* (pp. 287–323). Mahway, NJ: Erlbaum.
Davidsson, P. (2014). Getting published—and cited—in entrepreneurship: Reflections on ten papers. In A. Fayolle & M. Wright (Eds.), *How to get published in the best entrepreneurship journals. A guide to steer your academic career*. Cheltenham, UK: Edward Elgar Publishing.
Davidsson, P., & Delmar, F. (2003). Hunting for new employment: The role of high-growth firms. In D. Kirby & A. Watson (Eds.), *Small firms and economic development in developed and transition economies: A reader* (pp. 7–20). Aldershot, UK: Ashgate.
Davidsson, P., & Gordon, S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics, 39*(4), 853–876.
Davidsson, P., & Klofsten, M. (2003). The business platform: Developing an instrument to gauge and assist the development of young firms. *Journal of Small Business Management, 41*(1), 1–26.
Davidsson, P., Lindmark, L., & Olofsson, C. (1994a). *Dynamiken i svenskt näringsliv (Business Dynamics in Sweden)*. Lund, Sweden: Studentlitteratur.
Davidsson, P., Lindmark, L., & Olofsson, C. (1994b). New firm formation and regional development in Sweden. *Regional Studies, 28*, 395–410.
Davidsson, P., Lindmark, L., & Olofsson, C. (1996). *Näringslivsdynamik under 90-talet (Business Dynamics in the 90s)*. Stockholm: Nutek.
Davidsson, P., Lindmark, L., & Olofsson, C. (1998). The extent of overestimation of small firm job creation: An empirical examination of the 'regression bias'. *Small Business Economics, 10*, 87–100.
Davidsson, P., & Reynolds, P. D. (2009). PSED II and the Comprehensive Australian Study of Entrepreneurial Emergence [CAUSEE]. In P. D. Reynolds & R. T. Curtin (Eds.), *New firm creation in the United States: Preliminary explorations with the PSED II data set*. New York, NY: Springer.
Davidsson, P., & Steffens, P. (2011). Comprehensive Australian Study of Entrepreneurial Emergence (CAUSEE): Project presentation and early results. In P. D. Reynolds & R. T. Curtin (Eds.), *Business creation panel studies: An international overview*. New York, NY: Springer.
Davidsson, P., Steffens, P., & Fitzsimmons, J. (2009). Growing profitable or growing from profits: Putting the horse in front of the cart? *Journal of Business Venturing, 24*(4), 388–406.
Davidsson, P., Steffens, P., & Gordon, S. R. (2011). *Comprehensive Australian Study of Entrepreneurial Emergence* (*CAUSEE*): *Design*, *data collection and sample description*. I. In K. Hindle & K. Klyver (Eds.), *Handbook of new venture creation research*. Cheltenham, UK: Edward Elgar Publishing.
Davidsson, P., & Wiklund, J. (1997). Values, beliefs and regional variations in new firm formation rates. *Journal of Economic Psychology, 18*, 179–199.

Davidsson, P., & Wiklund, J. (2001). Levels of analysis in entrepreneurship research: Current practice and suggestions for the future. *Entrepreneurship: Theory and Practice, 25*(4), 81–99.

Delmar, F. (2015). A response to Honig and Samuelsson (2014). *Journal of Business Venturing Insights, 3*, 1–4.

Delmar, F., & Davidsson, P. (2000). Where do they come from? Prevalence and characteristics of nascent entrepreneurs. *Entrepreneurship & Regional Development, 12*, 1–23.

Delmar, F., & Davidsson, P. (2006). High-growth firms and their contribution to employment: The case of Sweden 1987-96. In P. Davidsson, F. Delmar, & J. Wiklund (Eds.), *Entrepreneurship and the growth of firms* (pp. 158–178). Cheltenham, UK: Edward Elgar Publishing.

Delmar, F., Davidsson, P., & Gartner, W. B. (2003). Arriving at the high-growth firm. *Journal of Business Venturing, 18*(2), 189–216.

Delmar, F., & Shane, S. A. (2004). Legitimating first: Organizing activities and the survival of new ventures. *Journal of Business Venturing, 19*, 385–410.

Delmar, F., & Shane, S. A. (2006). Does experience matter? The effect of founding team experience on the survival and sales of newly founded ventures. *Strategic Organization, 4*(3), 215–247.

DeTienne, D. R., Shepherd, D. A., & De Castro, J. O. (2008). The fallacy of "only the strong survive": The effects of extrinsic motivation on the persistence decisions for under-performing firms. *Journal of Business Venturing, 23*(5), 528–546.

Eckhardt, J. T., & Shane, S. A. (2010). An update to the individual-opportunity nexus. In Z. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research* (2nd ed., pp. 47–76). New York: Springer.

Eckhardt, J. T., & Shane, S. A. (2013). Response to the commentaries: The individual-opportunity (IO) nexus integrates objective and subjective aspects of entrepreneurship. *Academy of Management Review, 38*(1), 160–163.

Foss, N. J. (1993). Theories of the firm: Contractual and competence perspectives. *Journal of Evolutionary Economics, 3*, 127–144.

Foss, N. J., & Klein, P. G. (2012). *Organizing entrepreneurial judgment: A new approach to the firm*. Cambridge, UK: Cambridge University Press.

Fridh, A. C. (2002). *Dynamics and growth: The health care industry*. Doctoral dissertation, The Royal Institute of Technology, Stockholm.

Gatewood, R. D., Shaver, K. G., & Gartner, W. B. (1995). A longitudinal study of cognitive factors influencing start-up behaviors and success at new venture creation. *Journal of Business Venturing, 10*, 371–391.

Gordon, S. R., & Davidsson, P. (2013) *Capturing gazelles: Features of high potential firms and new venture growth. Business creation in Australia, 06*. Queensland University of Technology Business School, Brisbane, QLD. Retrieved from http://eprints.qut.edu.au/62936/

Gratzer, K. (1996). *Småföretagandets villkor. Automatrestauranger under 1900-talet (Conditions for Small Firms. Automated Restaurants During the Twentieth Century)*. Stockholm, Sweden: Almqvist & Wicksell.

Gratzer, K. (1999). The making of a new industry – the introduction of fast food in Sweden. In B. Johannisson & H. Landström (Eds.), *Images of Entrepreneurship Research -- Emergent Swedish Contributions to Academic Research* (pp. 82–114). Lund, Sweden: Studentlitteratur.

Gustafsson, V. (2004). *Entrepreneurial decision-making*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Henrekson, M., & Johansson, D. (2008). Gazelles as job creators: A survey and interpretation of the evidence. *Small Business Economics, 1*, 1–18.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage Publications.

Hubbard, R., & Lindsay, R. M. (2013a). From significant difference to significant sameness: Proposing a paradigm shift in business research. *Journal of Business Research, 66*(9), 1377–1388.

Hubbard, R., & Lindsay, R. M. (2013b). The significant difference paradigm promotes bad science. *Journal of Business Research, 66*(9), 1393–1397.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*, 305–360.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386–408.

Katz, J., & Gartner, W. B. (1988). Properties of emerging organizations. *Academy of Management Review, 13*(3), 429–441.

Kessler, A., & Frank, H. (2009). Nascent entrepreneurship in a longitudinal perspective: The impact of person, environment, resources and the founding process on the decision to start business activities. *International Small Business Journal, 27*(6), 720–742.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science, 3*(3), 383–397.

Liao, J., Welsch, H., & Tan, W. L. (2005). Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns. *Journal of High Technology Management Research, 16*(1), 1–22.

Lynn, R. (1991). *The secret of the miracle economy. Different national attitudes to competitiveness and money*. London: The Social Affairs Unit.

McMullen, J. S., & Dimov, D. (2013). Time and the entrepreneurial journey: The problems and promise of studying entrepreneurship as a process. *Journal of Management Studies, 50*(8), 1481–1512.

Mueller, D. C. (1972). A life cycle theory of the firm. *Journal of Industrial Economics, 20*(3), 199–219.

Naldi, L., & Davidsson, P. (2014). Entrepreneurial growth: The role of international knowledge acquisition as moderated by firm age. *Journal of Business Venturing, 29*(5), 697–703.

Nelson, R. R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge, MA: Belknap Press.

NUTEK. (2002). *Företagens villkor och verklighet 2002. Dokumentation och svarsöversikt (Conditions and reality of small firms 2002. Documentation and overview of responses)*. Stockholm: NUTEK.

Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioural sciences*. Chichester, UK: John Wiley & Sons, Inc.

Obschonka, M., Stuetzer, M., Gosling, S. D., Rentfrow, P. J., Lamb, M. E., Potter, J., et al. (2015). Entrepreneurial regions: do macro-psychological cultural characteristics of regions help solve the "knowledge paradox" of economics? *PloS One, 10*(6), e0129332.

Parker, S. C. (2011). Intrapreneurship or entrepreneurship? *Journal of Business Venturing, 26*(1), 19–34.

Patel, P. C., & Thatcher, S. M. (2014). Sticking it out: Individual attributes and persistence in self-employment. *Journal of Management, 40*(7), 1932–1979.

Penrose, E. (1959). *The theory of the growth of the firm*. Oxford, UK: Oxford University Press.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903.

Porter, M. E. (1980). *Competitive strategy*. New York, NY: Free Press.

Porter, M. E. (1985). *Competitive advantage*. New York, NY: Free Press.

Porter, M. E. (1990). *The competitive advantage of nations*. London, UK: Macmillan.

Rauch, A., & Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology, 16*(4), 353–385.

Reynolds, P. D. (2007). New firm creation in the United States: A PSED I overview. *Foundations and Trends in Entrepreneurship, 3*(1), 1–150.

Reynolds, P. D. (2009). Screening item effects in estimating the prevalence of nascent entrepreneurs. *Small Business Economics, 33*(2), 151–163.

Reynolds, P. D., Carter, N. M., Gartner, W. B., & Greene, P. G. (2004). The prevalence of nascent entrepreneurs in the United States: Evidence from the Panel Study of Entrepreneurial Dynamics. *Small Business Economics, 23*(4), 263–284.

Reynolds, P. D., & Curtin, R. T. (2008). Business creation in the United States: Panel Study of Entrepreneurial Dynamics II initial assessment. *Foundations and Trends in Entrepreneurship, 4*(3).

Reynolds, P. D., & Maki, W. R. (1990). *Business volatility and economic growth. Final report*. Washington, DC: Small Business Administration.

Reynolds, P. D., & Miller, B. (1992). New firm gestation: Conception, birth and implications for research. *Journal of Business Venturing, 7*(5), 405–417.

Reynolds, P. D., Miller, B., & Maki, W. R. (1993). Regional characteristics affecting business volatility in the United States, 1980–94. In C. Karlsson, B. Johannisson, & D. Storey (Eds.), *Small business dynamics. International, national and regional perspectives* (pp. 78–114). London, UK: Routledge Peterson & Co.

Ruef, M., Aldrich, H. E., & Carter, N. M. (2003). The structure of organizational founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs. *American Sociological Review, 68*(2), 195–222.

Sakhdari, K., Burgers, H., & Davidsson, P. (2014). Capable but not able: The effect of institutional context and search breadth on the absorptive capacity-corporate entrepreneurship relationship. In P. Davidsson (Ed.), *Australian Centre for Entrepreneurship research exchange conference 2014 proceedings* (pp. 954–974). Sydney, NSW: Queensland University of Technology.

Samuelsson, M., & Davidsson, P. (2009). Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics, 33*(2), 229–255.

Sarasvathy, S. D. (2008). *Effectuation: Elements of entrepreneurial expertise*. Cheltenham, UK: Edward Elgar Publishing.

Scheinberg, S., & MacMillan, I. C. (1988). An 11 country study of motivations to start a business. In B. A. Kirchhoff, W. A. Long, W. E. McMullan, K. H. Vesper, & W. E. Wetzel (Eds.), *Frontiers of entrepreneurship research 1988* (pp. 669–687). Wellesley, MA: Babson College.

Schoonhoven, C. B., Burton, M. D., & Reynolds, P. D. (2009). Reconceiving the gestation window: The consequences of competing definitions of firm conception and birth. In P. D. Reynolds & R. T. Curtin (Eds.), *New firm creation in the United States* (pp. 219–237). New York, NY: Springer.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, UK: MA: Harvard University Press.

Schwab, A., Abrahamson, E., Starbuck, W. H., & Fidler, F. (2011). Researchers should make thoughtful assessments instead of null-hypothesis significance tests. *Organization Science, 22*(4), 1105–1120.

Senyard, J., Baker, T., Steffens, P., & Davidsson, P. (2014). Bricolage as a path to innovativeness for resource-constrained new firms. *Journal of Product Innovation Management, 31*(2), 211–230.

Seth, A., & Thomas, H. (1994). Theories of the firm: Implications for strategy research. *Journal of Management Studies, 3*, 165–191.

Shane, S. A. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science, 11*(4), 448–469.

Shane, S. A. (2012). Reflections on the 2010 AMR Decade Award: Delivering on the promise of entrepreneurship as a field of research. *Academy of Management Review, 37*(1), 10–20.

Shane, S. A., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review, 25*(1), 217–226.

Shaver, K. G. (2010). The social psychology of entrepreneurial behavior. In Z. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research: An interdisciplinary survey and introduction* (pp. 359–386). New York, NY: Springer.

Shaver, K. G., Carter, N. M., Gartner, W. B., & Reynolds, P. D. (2001). Who is a nascent entrepreneur? Decision rules for identifying and selecting entrepreneurs in the panel study of entrepreneurial dynamics (PSED) [summary]. In W. D. Bygrave, E. Autio, C. G. Brush, P. Davidsson, P. G. Green, P. D. Reynolds, & H. J. Sapienza (Eds.), *Frontiers of entrepreneurship research 2001* (p. 122). Wellesley, MA: Babson College.

Shugan, S. M. (2007). Errors in the variables, unobserved heterogeneity, and other ways of hiding statistical error. *Marketing Science, 25*(3), 203–216.

Sørensen, J. B. (2007). Bureaucracy and entrepreneurship: Workplace effects on entrepreneurial entry. *Administrative Science Quarterly, 52*(3), 387–412.

Steffens, P. R. (2013). Culture as a driver of entrepreneurship: Contrasting independent entrepreneurship versus employee entrepreneurship. *Paper presented at the ACERE Conference*, Brisbane, Feb 5–8. Retrieved from http://eprints.qut.edu.au/59472/.

Steffens, P. R., Davidsson, P., & Fitzsimmons, J. (2009). Performance configurations over time: Implications for growth-and profit-oriented strategies. *Entrepreneurship: Theory and Practice, 33*(1), 125–148.

Steffens, P. R., Terjesen, S., & Davidsson, P. (2012). Birds of a feather get lost together: New venture team composition and performance. *Small Business Economics, 39*(3), 727–743.

Steffens, P. R., Tonelli, M., & Davidsson, P. (2011) How do we reach them? Comparing random samples from mobile and landline phones. In *Proceedings of AGSE Entrepreneurship Research Exchange* 2011, Swinburne University of Technology, Melbourne, VIC.

Stuetzer, M. Obschonka, M., Audretsch, D.B., Wyrwich, M., Rentfrow, P.J., Coombes, M., Shaw-Taylor, L & Satchell, M. (2015). Industry structure, entrepreneurship, and culture: An empirical analysis using historical coalfields. *European Economic Review* (forthcoming).

Ucbasaran, D., Westhead, P., & Wright, M. (2006). *Habitual entrepreneurs*. Cheltenham, UK: Edward Elgar Publishing.

Uy, M. A., Foo, M. D., & Aguinis, H. (2010). Using experience sampling methodology to advance entrepreneurship theory and research. *Organizational Research Methods, 13*(1), 31.

Wennberg, K., Wiklund, J., DeTienne, D. R., & Cardon, M. S. (2010). Reconceptualizing entrepreneurial exit: Divergent exit routes and their drivers. *Journal of Business Venturing, 25*(4), 361–375.

Wernerfelt, B. (1984). A resource based view of the firm. *Strategic Management Journal, 5*, 171–180.

Wernerfelt, B. (1995). The resource-based view of the firm: Ten years after. *Strategic Management Journal, 16*, 171–174.

Wigren, C. (2003). *The spirit of Gnosjö. The grand narrative and beyond.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Wiklund, J., Davidsson, P., & Delmar, F. (2003). What do they think and feel about growth? An expectancy-value approach to small business managers' attitudes towards growth. *Entrepreneurship: Theory and Practice, 27*(3), 247–269.

Wiklund, J., & Shepherd, D. A. (2008). Portfolio entrepreneurship: Habitual and novice founders, new entry, and mode of organizing. *Entrepreneurship: Theory and Practice, 32*(4), 701–725.

Williamson, O. E. (1975). *Markets and hierarchies*. New York, NY: The Free Press.

Williamson, O. E. (1999). Strategy research: Governance and competence perspectives. *Strategic Management Journal, 20*, 1087–1108.

Yang, T., & Aldrich, H. E. (2012). Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing, 27*(4), 477–492.

Zander, I. (2007). Do you see what I mean? An entrepreneurship perspective on the nature and boundaries of the firm. *Journal of Management Studies, 44*(7), 1141–1164.

Operationalization Issues

6

Abstract

How can we measure things like "entrepreneurial self-efficacy" or "start-up success" so that we can test whether our theories hold up? This chapter starts with an extended discussion of standard issues pertaining to operationalization of theoretical constructs (levels of measurement, measurement validity, and measurement error) in an integrated and reflective fashion. After discussing different approaches to measurement (single-item vs. reflective vs. formative indices), it proceeds to discuss the balancing of generic vs. customized operationalizations as well as concrete ways to assess measurement validity. Operationalization challenges and solutions are illustrated with examples such as entrepreneurial bricolage, venture idea novelty, and entrepreneurial action.

6.1 A 90-Degree Turn

Operationalization concerns the "translation" of theoretical concepts into measured empirical variables. In broadly based, quantitative research, operationalization thus concerns the columns in the data matrix, whereas the sampling issues dealt with in the previous chapter concerned the rows. However, the main problem of operationalization—correspondence between theoretical constructs and empirical observations or *measurement validity*—is relevant to all types of research. If you want to gain insights into entrepreneurial behavior through interviews with key informants, you may not think you are engaging in measurement, but you probably want the information they provide to reflect their actual behavior rather than their inclination to forget, lie, or try to impress. When you design an experiment, you want your experimental manipulation to reflect the theoretical notion and/or real-world condition you are interested in rather than, e.g., (also) capturing some confounding factor (didn't I say something about "impressing the Sheila" in a previous chapter?). This

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_6

said, the operationalization issues pertaining to research based on surveys and archival data will dominate the chapter.

Despite the 90-degree turn, there is a certain amount of overlap between sampling and operationalization, and some important operationalization issues have therefore been dealt with already. For example, when I described in Chap. 5 how cases were sampled for the CAUSEE project, this was at the same time a description of how the concept "nascent venture" was operationalized. Likewise, the procedure for sampling new internal ventures in existing firms can, at the same time, be regarded as a firm-level operationalization of entrepreneurial action. However, these are just examples out of many possible ways of operationalizing these abstracted concepts. In the present chapter, we will go much deeper into these issues.

A lot has happened in the area of operationalization in entrepreneurship since the first edition of this book. Some 10–15 years ago, there weren't many operationalizations—or even theoretical constructs—that were reapplied in study after study. Today, reusing previously defined constructs and existing operationalization is the norm, and new operationalization suggestions and evaluations are the topic of entire journal articles (e.g., Dahlqvist & Wiklund, 2012; Grégoire, Shepherd, & Schurer Lambert, 2010). Recently, a single, regular issue of *Entrepreneurship Theory and Practice* had three articles discussing measurement of core constructs in our field (Covin & Wales, 2012; Perry, Chandler, & Markova, 2012; Runyan, Ge, Dong, & Swinney, 2012). We have even seen formal evaluations of the status of construct measurement in entrepreneurship (Bouckenooghe, De Clercq, Willem, & Buelens, 2007; Crook, Shook, Morris, & Madden, 2010). For this reason, this chapter has essentially been rewritten from scratch, only saving a few bits and pieces from the first edition.

As usual, I will not strive to be complete. I will try to cast somewhat different light on the basic measurement and operationalization issues that are covered in standard textbooks. In doing this, I wish to invite you to *think and reflect* about operationalization issues rather than learning to apply rules that some authority says will remedy this or that problem. I will also try to go deeper into operationalization issues that are particular to entrepreneurship. The most important operationalization issues arguably pertain to our explananda, i.e., the phenomena we wish to explain, or the "dependent variables." In fact, I find this so important that I give them a separate chapter immediately after this one.

I will assume that the reader has some prior knowledge of the vocabulary associated with operationalization and measurement issues, but as a refresher—and because there is quite a jungle out there with a plethora of terms used somewhat differently by different authors—let's start with a list of terms and my own definitions or explanations of them:

True score: An underlying assumption is that the theoretical construct is meaningful and that each assessed entity has a true value for that construct which can be compared to other entities. This is something few of us would debate when we talk about how tall we currently are, whereas the notion can be questioned for a construct like "entrepreneurial skill" (can it be compressed into one dimension?) and even more so for "attitude toward entrepreneurs" (do people even have one before you ask, and is "entrepreneur" a clear enough notion for the comparable "true score" to exist?). For now, we will *assume* that it is meaningful to *assume* the existence of a "true score."

Random measurement error: The extent of random, nonsystematic deviations of the observed measurement from the true score.

Systematic measurement error: The extent of systematic deviations of the observed measurement from the true score (e.g., respondents' tendency to exaggerate or underreport true income in tax office data).

Measurement validity or *construct validity:* An operationalization's degree of absence of random and systematic measurement error; that it measures the theoretical construct it is supposed to measure. A plethora of sub-concepts are associated with measurement validity, for example:

Face validity: Does the measure appear to capture the theoretical construct according to a superficial first impression (which is not necessarily invalid in itself; see Gladwell, 2007)?

Content validity: According to a systematic assessment, does the measure cover the entire theoretical construct and as little as possible of anything outside of it?

Convergent validity: Does the measure correlate highly with some other measures of the theoretical construct, which we know has high validity (but which may be too costly or cumbersome to use for our purposes)?[1]

Predictive validity: Does the measure correlate in theoretically expected ways with measures of other theoretical constructs?

Discriminant validity: Is the measure sufficiently distinct from measures of other theoretical constructs?

Measurement reliability: Relative absence of random measurement error, precision. If we could repeat the measurement on the same entities (without them undergoing any real change in the meantime or being affected by the initial measurement), how much would the results vary? A measuring instrument can be highly reliable without being valid (by being "exactly wrong" every time). Some sources see high reliability as a prerequisite for high validity; others see validity and reliability as separate issues by defining measurement validity as freedom from *systematic* measurement error, i.e., the measure is valid if in repeated assessments it is correct "on average," even if the reliability is low.

Level of measurement (not to be confused with "level of analysis") concerns how precise or sophisticated a type of measurement is, which influences what types of mathematical and statistical operations they allow. Following Suppes and Zinnes (1963) and Stevens (1946), measurement scales are often divided into the following four types in ascending order of sophistication:

[1] An interesting, real example is Guzman and Stern's (2015a) two versions of the Entrepreneurship Quality Index, one of which only requires data available at registration whereas the other makes use of additional indicators that come with a 1–2 year lag.

Nominal scales: The scale values are merely labels for categories; no order or relative size/amount is implied (e.g., in categorizing different mining sites by which metal they are extracting, we may arbitrarily label them 1 = silver, 2 = gold, 3 = tin, 4 = lead, 5 = iron, and 6 = copper.

Ordinal scales: The scale value gives ordinal information, but the distance between each ordered object may be unequal (e.g., ordering the following six metals by melting point in descending order, we would get 1 = iron, 2 = copper, 3 = gold, 4 = silver, 5 = lead, and 6 = tin, while the difference in degrees would be different at each step.

Interval scales: If we don't just have a furnace to determine the order in which they melt but a heavy-duty thermometer to assess the temperature at which they do so, we have the following scale values: iron = 1538 °C, copper = 1085 °C, gold = 1064 °C, silver = 961.8 °C, lead = 327.5 °C, and tin 231.9 °C.[2] Here, a degree is a degree is a degree, so we can validly compute a mean melting point for these metals or use their melting points as a variable in a regression analysis.

Ratio scales: The Rolls Royces of measurement are like interval scales with an indisputable zero point. Height, weight, and a number of other physical measures qualify here—as do sales and number of (full-time equivalent) employees in our own domain. Our thermometer scales are not. If it is 20 °C today and it was 10 °C yesterday, I can validly claim that the average was 15 °C, but I cannot validly claim it is twice as warm today as it was yesterday. Why? Because if you switch to Fahrenheit, the claim is obviously false (68 vs. 50 degrees in this example).

6.2 On Course Toward Validity

To illustrate some operationalization issues, let's talk golf! Table 6.1 displays the top ten players in the 2012 Frys.com competition, which is part of the PGA Tour. No, I did not choose this particular one because a Swede won it—I'm just as much an Aussie these days, remember? Neither did I choose golf because it's my favorite pastime (it isn't; I've completely lost my game and hardly ever play anymore). It just so happens that this competition ended the day before writing this,[3] and the results happen to illustrate a number of measurement issues, some of which are normally discussed in methods texts, whereas others are regularly forgotten both in such texts and in applied research. As it turns out, the seemingly comparatively straightforward task of measuring "golfing ability" is not that simple. This is the reason to reflect some more on the arguably much more delicate measurement problems we encounter in entrepreneurship and other social science research.

[2] If you live in Burma (Myanmar), Liberia, or that third country which has not yet officially accepted the SI system, multiply °C by 1.8 and add 32 to get Fahrenheit equivalents.

[3] Which reveals that this was the very first section of the book that I started to revise. I also confess to cheating with the WGR score, which I did not save at the time (unlike the WGR rank). For example, I have imputed the WGR scores for the corresponding ranks on April 11, 2015.

Table 6.1 Results for top ten players in the Frys.com golf tournament, 2012

Final position	Player	Country	WGR score	WGR rank	Drive dist.	R1	R2	R3	R4	Total score
1	Jonas Blixt	SWE	1.81	75	281.2	66	68	66	68	268
T2	Tim Petrovic	USA	0.39	421	280.1	70	68	67	64	269
T2	Jason Kokrak	USA	0.84	205	300.4	68	66	67	68	269
T4	Jimmy Walker	USA	1.47	103	304.8	73	68	67	62	270
T4	Vijay Singh	FIJ	2.06	67	290.5	70	66	66	68	270
T4	Alexandre Rocha	BRA	0.44	384	280.0	69	67	66	68	270
T4	John Mallinger	USA	1.14	154	266.4	66	62	70	72	270
8	Jeff Overton	USA	1.71	83	282.6	68	69	68	66	271
T9	Gary Woodland	USA	1.49	102	303.0	66	72	66	68	272
T9	Russell Knox	SCO	0.52	331	292.3	70	68	65	69	272

The table illustrates different types of measurement. First, we have the nominal scale "country." These players originate from Brazil, Fiji, Scotland, Sweden, and the USA. We could assign the values 1, 2, 3, 4, and 5 to these countries, or the reverse, or 5, 45, 90, 2, and 12,345, or even A–E. It does not really matter; the numbers are just labels. To make country a useful variable for analysis, we would convert the information into four ($n-1$) dichotomous variables, where each player had the value "1" on one country dummy and "0" on all others (except players from the reference country—in this case, likely the USA as the PGA Tour is USA based—who would have "0" on all country dummies).

Second, we have the *ordinal* scale "Final Position," from first to (tied) ninth. By itself, this ranking only tells who played better than whom, not by how much. In this particular case, a sneak glimpse at the total scores over the four rounds of play reveals that for each position there is one more stroke. That is, the ranks are equidistant by happenstance. If we instead rank the players by their performance in round 2 (column R2), we see the usual "problem" with ordinal scales, as the number of strokes separating first, tied second, fourth, tied fifth, ninth and tenth players is here 4, 1, 1, 1, and 3, respectively. Ordinal position is what matters most in sports. In golf, this is what determines your cut of the prize money, and in middle-distance running in the Olympic Games, the winner may be a tactical genius to win, but very far from any world record. In entrepreneurship research, however, truly ordinal scales are of little interest. It is not with ordinal scales that we test theory.

The organizer's intention is supposedly to give the winner's check to the player who shows the greatest "golfing ability" over the 4 days of play. This is a theoretical construct which is arguably most precisely measured by the *ratio* scale "Total Score." We can see that Total Score is a ratio scale because zero on that scale really means zero (no shots have been played) whereas zero on the common temperature scales cannot be interpreted as "no temperature." Consequently, we can make ratio calculations and find that Blixt won by using roughly one third of a percent fewer strokes than the closest competitor (OMG! That's "all" it takes to walk away with all that money!?). But this ratio scale also has some interesting properties. For example, the difference between 269 and 268 strokes is the same as the difference between 270 and 269 strokes, namely one stroke. But if we think a little harder, we realize that a "stroke" in golf can be a 300 m[4] drive, a short "hack" into a more playable position, a put that lips out or lips in, a penalty, or a 200+ meters approach shot that bounces right into the hole for an albatross. This makes our conclusion about equidistance of scores rest on the acceptance of a whole range of assumptions and conventions,[5] such as various arbitrary rules of the game including the agreed upon sizes of golf balls and golf hole cups. What are we adding up in our ratio scales in entrepreneurship and management research? Employees added from acquisitions and organic growth? Justified? Dollars of revenue from operations and one-off divestments? Justified? Do each dollar and each employee have the same theoretical meaning?

Not least, the tight final standings give reason to reflect on the validity and reliability of the Total Score as indicator of golfing ability. As regards validity, if we refrain from questioning the rules of the game for a while and consider the presence of spectators, marshals, and TV cameras, it appears there is no systematic measurement error involved (at least not in the form of cheating). Thus, the score in a round of tournament golf should be a highly valid indicator of a player's golfing ability at that time and on that type of course, which is exactly what that competition is about. So there are no problems with validity. This might be different when you ask the average social player about their score, and it cannot be completely ruled out that the player's memory is somewhat selective and hence their report might contain a dash of self-serving bias (i.e., systematic measurement error). There would also be a clear risk of systematic measurement error if we were to compare scores across players playing at different courses. Because some courses are harder than others, scores from different courses do not achieve measurement equivalence (see the discussion of uneven validity in Chap. 4). This is an important notion to keep in mind also when trying to compare measures across countries and industries (Runyan et al., 2012). Does introducing "a lot of new products"—or even the same objective number of new products—within a specified time frame mean the same amount of "entrepreneurial orientation" across these contexts, and does it thus deserve the same numerical value?

What about reliability? If we look at each round reported in the table as a separate event (columns R1–R4), we actually find six different winners (three of them tied on the first day). Why aggregate over 4 days? A player's score variation from

---

[4] If you live in Burma (Myanmar), Liberia, or that third country which has not yet officially accepted the SI system, a meter is 1.0936133 yards. At this point, I should perhaps also warn you that just a few lines down I am going to use the term "albatross" for what you may know as a "double eagle." And further down still, I use "athletics," which you may know as "track and field."

[5] The knowledgeable reader may protest that this is because I choose a count variable to illustrate ratio scales. Fair enough, this increases the problem, but if you dig deeply into it, you'll find it isn't completely absent even in something as straightforward as human physical height.

day to day may be seen mainly as random measurement error.[6] Such errors cancel out in the aggregate, so it seems a good idea to add rounds in order to arrive at a worthy winner. Thus, although all of these players show absolutely wonderful reliability in their golfing ability compared to us regular hackers and duffers, we may need to aggregate an array of separate measures in order to assess them fairly, not least because they are all so damn good. For the same reason, we use multiple-item batteries of questions to measure tricky, entrepreneurship-related variables. By this logic, Blixt is a more worthy winner than John Mallinger (who “won” over the first three rounds) or Jimmy Walker (who “won” over the last three rounds) because the score over four rounds contains less random measurement error. Similarly, a more comprehensive measure of entrepreneurial behavior is usually a more reliable one.

Golf tournaments typically run over four rounds to determine the champion for the week. However, the score from one tournament is hardly a valid measure of who is the best golfer in the world—who has the greatest golfing ability in the entire relevant population—at that point in time. This is because particularities of the course design, weather, and sheer luck—and who chose to participate or not—may have unduly affected the results. To achieve greater content validity in a more global measure of golfing ability, one would like to have results from different types of courses and under different weather conditions, so that proficiency at all parts of the game—driving distance and accuracy; strategy; play in rough, sand, wind, cold, and heat, across water, and on high altitude; accuracy in approach; putting on even and undulated surfaces as well as different types of grass—was given due weight in the overall measure. This is essentially what is done in the comprehensive assessment underlying the World Golf Ranking (WGR).[7] As is demonstrated by the “WGR rank” column in Table 6.1, the very best players at the time either did not participate or did not do very well in this tournament. The ranks also show the results of this one competition are not a good measure of these players’ absolute or relative standing in the world of golf. What about our entrepreneurship measures? What do they need to capture in order to achieve sufficient content validity? Further, despite the comprehensive assessment, WGR ranks do not seem strikingly stable over time.[8] This highlights the question of *when* we should assess our entrepreneurship constructs. At the theoretically relevant time for showing the expected effects, or what? A question too rarely asked?

In addition, is Total Score—which can alternatively be expressed as average score per round—the theoretically relevant measure? By the rules of this tournament it is, but could one not defend other criteria for having the best “golfing ability,” such as having the lowest score in any round, or the “least bad” one-day score across the four rounds? That is, should we not let variance enter the picture? In this particular tournament, it so happens that many of the top players also have very low

[6] There could also be an element of systematic measurement error due to shifting weather conditions or some stocky, blond golfer having a drink too many on one of the nights.

[7] www.owgr.com/Ranking

[8] Apart from the historical parenthesis when Tiger Woods held the No. 1 position for 281 weeks straight (or 683 in total…and possibly counting, although I very much doubt it).

variance. So perhaps consistency has already been duly rewarded? If we examine the table closely, we find that Jason Kokrak in third place has even lower variance than Blixt, so lowest variance per se did not quite cut it (and neither should it; think 74–74–74–74). Stopping after three rounds might have overly rewarded Mallinger for his "fluke" second round. By the same token, perhaps we should put Overton before Mallinger based on the former's proven ability to break 70 four days straight? By contrast, Mallinger actually performed the worst out of this lot of ten on each of the last 2 days. What is more important for our entrepreneurship operationalizations? Is it to minimize the average measurement error or to avoid instances of large error?

There is also reason to ask whether a player who always plays to their average ability—even if that ability is high—would ever win a tournament? Perhaps, given the same average (which more or less all of the players on the PGA Tour have, if you're not too picky), that higher variance is better? That is, perhaps we should reward peak performance rather than average score? That's exactly what we do in many other sports, for example, the long jump or discus throw in athletics. It does not matter what you did with your other five attempts; if your best one is the longest overall, you're the master. Similarly, other golf tournaments, like the World Match Play or the Reno-Tahoe Open, apply a different view on averages and variance, as does the WGR score, which gives considerable extra weight to wins and other top positions regardless of the score. If peak performance were the criterion, Walker and Mallinger would share the title with rounds of 62, whereas Blixt would not only have these two and Russell Knox ahead of him but also another 11 players who scored 65 or lower in one round but did not make the top 10 overall. Aussie Nick O'Hearn put in a round of 62, on day 1, thus sharing the win according to our revised system, although he ended up tied 22nd overall by the rules actually applied.

Where am I going with this? Well, at this stage, there is reason to think about the entrepreneurship equivalent. What is the relevant indicator? Should we use the normal (mode) or average value for the entity to predict engagement in or success at entrepreneurship, or is it the ability to really stick out once in a while that matters? Conversely, is it sometimes perhaps the ability to avoid disasters that is salient (think "affordable loss," Sarasvathy, 2008)? The analogous problem has been observed in research in marketing concerning consumers' decision rules (e.g., Hauser et al., 2010) but is commonly forgotten when discussing operationalization of theoretical constructs. That does not make it unimportant. We need to measure what is theoretically relevant.

This is also the place to have another think about levels of measurement, measurement error, and permissible statistics. According to conventional wisdom, interval and ratio scales have equal distance between, e.g., 5 and 6, and 134 and 135, respectively, and that therefore arithmetic operations are deemed permissible (which means we can calculate means, compute correlations, and use any sophisticated techniques using these as inputs). This makes ratio (and interval) scales the preferred level of measurement. Ordinal scales have ambiguous and probably varying intervals. There is *no intent* to have my degree of preference of 1 over 2 match that of my preference for 2 over 3, and therefore fewer statistical methods apply. And then comes the big mistake: to confound *rating scales* with ordinal level of measurement. Because when using a rating scale (e.g., agree-disagree statements with responses ranging from, e.g., 1 to 7 or −3 to +3) we cannot know whether the distance from 2 to 3 is "the same difference" as that between 4 and 5, rating scales are bundled with ranking scales as "ordinal" in many methods textbooks.

Now, let's see where concerns for equidistant intervals might take us. A wealth of quantitative information is collected about elite golfers. One of them is average driving distance, and nothing gets more metric than that; it is actually a measure in *meters*.[9] Each player's average driving distance in this tournament has its own column in Table 6.1. This is a very precise, ratio measure. It will have some error from rounding and occasional mistakes, but this error is likely to be infinitesimal, as meters are something that is vastly much easier to assess than almost anything we have a theoretical interest in within the social sciences. The only problem is that the measure assesses driving distance and nothing else. If you examine the table, you will find that this seems to have very little association with success in this tournament or with what we are really interested in: golfing ability. As a case in point, the ten players who ended up last in the tournament (not included in the table) have exactly the same average driving distance across the 4 days and 72 holes as have the ten top players in the table: 286 m.[10]

Here's the point: a wealth of information goes into the World Golf Ranking. Much of that information is quantitative (of sorts), and it arguably reflects all aspects of the theoretical construct golfing ability. Before converting it into an overall WGR rank (which is decidedly ordinal), all that information is compiled and weighted into an overall WGR score, which looks very quantitative. In fact, it is contingent on an array of imputed values that could be debated, like the relative weights assigned to different tournaments and to different positions within them. That is, no matter where we turn, at the end of the day, the WGR score is a *rating* of the golfer's ability. There is no valid, direct, interval, or ratio measure of global golfing ability, and there never will be one.

Do we believe that such a thing as golfing ability exists? Do we believe that in principle, each player has a "true score" on a meaningfully comparable theoretical scale of such ability? We can think of this as a measure of the total scores we would obtain if we could make all these players play all golf courses in the world under a variety of conditions. If we believe in the existence of a true score (or perhaps rather in the meaningfulness of assuming such existence), we also believe there is an underlying, continuous distribution with interval-scale properties. The *intent* of the WGR score (on which the rank is based) is to capture that interval-scale distribution. I would say exactly the same applies to firm-level *entrepreneurial orientation*,

[9] If you are still living in Burma, Liberia, or that third country which has not yet officially accepted the SI system and you have already forgotten, here's a reminder: a meter is 1.0936133 yards. Hey, I'm really adapting here, writing "meter" rather than "metre."

[10] Hmmm…coming to think of it, these stats are probably in yards, but at this point, I'd better keep a straight face. Doesn't matter: believe it or not, yards have metric properties, every one of them is equally long! We should also admit that across all types of golfers, there would be a fairly strong correlation between average driving distance and overall golfing ability.

individual-level *entrepreneurial self-efficacy*, and a whole range of other constructs of interest to entrepreneurship researchers. What matters is not whether or not *the measuring device* undoubtedly has interval-scale properties or not; what matters is which available measure is likely to yield the lesser, total measurement error. Choosing between driving distance (a ratio scale) and a comprehensive WGR score, the measure that would have the smaller measurement error in relation to the theoretical variable global golfing ability is undoubtedly the latter.

The above should serve to illustrate that "measuring the nonmeasurable" is no easy task. Even such an elaborate assessment as WGR score can be questioned. The above should also demonstrate that we are justified in measuring many theoretical constructs in terms of rating scales which can never reach indisputable, interval-scale properties. Our task is to minimize the measurement error. This is achieved through hard thinking and hard work but not by turning to something that is measurable on an interval or ratio scale, but lacks content validity.

The above examples should also be humbling in relation to the precision with which we can expect to measure theoretical constructs with Likert scales in questionnaires. Consider the example of the discus throwing competition at an athletics event. With the electronic measurement that is used today, the measurement error is very small. However, the results are rounded to whole centimeters, and sometimes it happens that the length is measured from the wrong divot. The true lengths of the throws, however, are exactly what they are regardless of our mistakes and our rounding conventions. The same is true with measuring tape; using such instead of electronic measurement increases measurement error but does not alter the true length of the throws. Now, imagine that neither electronic measurement nor measuring tape was ever invented. Instead, we would have the athletes judge the length of their throws on a five-point scale (1 = very short, 2 = short, 3 = average, 4 = long, 5 = very long). This is a much cruder type of measure, prone to considerable measurement error. Again, the true lengths of the throws, however, are not affected by this, and unless the athletes' judgments are completely unreliable, there will be a substantial positive correlation between the true lengths and the reported assessments. It is, for example, highly likely (in the men's event at world-class level) that every throw longer than 70 m would get a "5" and every throw shorter than 50 m a "1." Because of the crudeness of the measuring device, it would probably not discriminate between a throw of 68.48 and one of 69.36. Admittedly, it would sometimes occur that a throw of 66.50 was assigned a "4" while another of 65.14 got a "5." However, the difference between this measurement method and the electronic measurement is the size of the measurement error. Both measurement methods try to represent an underlying, continuous distribution of lengths.

The example may seem odd, but this is in fact how crudely we measure attitudes and many other constructs in surveys. As there is no difference in principle among the less-than-perfect measures in the example, this reasoning is defense for performing arithmetic operations on rating scale measures. The comparison with electronic measurement of length is also reason to reflect, however, on just how rudimentary a type of operationalization self-reports of attitudes and the like are. We know that random measurement errors will decrease drastically, though, if we aggregate the estimates across several throws and/or across several "judges." This is why we use a battery of items and sum or average the scores into an index. By doing so, we get a more continuous measure and therefore presumably one that better reflects the underlying, continuous distribution. Secondly, whereas random measurement errors tend to cancel out through this procedure, systematic measurement errors are not worsened.

6.3 Different Approaches to Operationalization

Some variables can be measured rather precisely with a single measure on one, agreed-upon scale. As just discussed, however, many theoretical constructs are like "golfing ability" and require a combination of several indicators in order to get a valid and reliable measure. Figure 6.1 illustrates two different approaches one can take in such situations, both applied to the same theoretical construct: ability.

Panel (a) in the figure displays a *reflective index* where the theoretical construct is regarded as an underlying, nonmeasurable characteristic, which causes the variation in (i.e., *reflected* in) the manifest variables that we use as operationalization. This is the dominant, psychometric measurement paradigm, and it is what we implicitly applied above when arguing that this week's golfing ability is reflected in each of the daily scores over the four rounds. It is for this type of index that it is meaningful to demand that items load highly on the same factor in a factor analysis and yield a high Cronbach's alpha coefficient. This measurement logic also weighs heavily in covariance structure-based structural equation modeling (SEM) techniques. Because the different items have the same underlying cause, they should have high intercorrelations and thus load on a common factor and yield a high alpha.

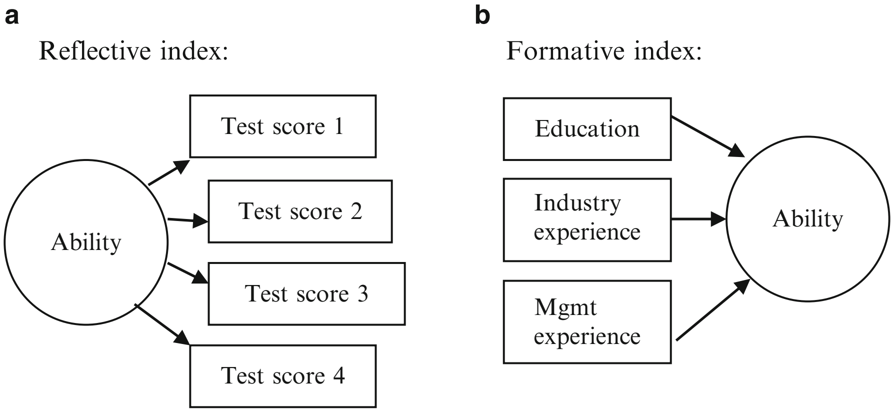


Fig. 6.1 Two types of summated indices (latent or composite variables) (**a**) Reflective index. (**b**) Formative index

We can also operationalize "ability" as a *formative index* (Diamantopoulos, Riefler, & Roth, 2008) as in panel (b). For other constructs, this approach may be the most suitable or even the only reasonable option. In this case, education and experience are seen as complementary *causes* of increased ability. Hence, the items are what build up (or *form*) the theoretical variable. The different components that contribute to ability need not be highly correlated. Formative index building is also an option in SEM models applying the Partial Least Squares (PLS) technique (Ringle, Sarstedt, & Straub, 2012), which empirically assigns unequal weights to each indicator; i.e., they need not all be assumed equally important. With formative measurement, "golfing ability" could be assessed as a weighted index of "driving distance," "driving accuracy," "greens in regulation," "sand saves," "putts per green," etc., which capture proficiency at different aspects of the game.[11] In our figure's example, one could argue that long education necessarily reduces one's chances of also having long experience. Therefore, indices created via the formative logic should not be evaluated through factor analysis or Cronbach's alpha, and if reviewers ask you to do so, you'll need to gently educate them. This said, it is advisable to read the case against formative measurement before applying it (Edwards, 2011). Edwards is no dummy; he can only be ignored at peril.

Although the reflective measurement logic dominates and the formative approach is being increasingly understood and accepted, they are not the only shows in town. Recently, John Rossiter's C-OAR-SE has been launched as a frontal attack on the general superiority of multiple-item over single-item measurement in general and on the reflective paradigm (as commonly executed in business research) in particular (Rossiter, 2002, 2010, 2011). After some initial silence, his approach is gaining both critical and supportive interest (Diamantopoulos, 2005; Diamantopoulos et al., 2012; Finn & Kayande, 2005; Hadwich et al., 2010). It is well worth having a look at, but please do *not* use it for the purpose of defending some lame, single-item measure that has nothing to do with Rossiter's measurement theory. Reading Rossiter's argument should remind you of always considering the basics of content validity before worrying about anything else (see further below). Whether or not you should buy all of his arguments is another matter.

The approach to measurement also varies by discipline. Faced with different options, empirical economists may be more prone to try to tease out which is the "best indicator" or to include several separate indicators of the same construct in the same model rather than combining several indicators into an index. Similar examples occur in entrepreneurship (Carton & Hofer, 2006; Reynolds, Storey, & Westhead, 1994). Based on my experience from research on firm growth, I would recommend

[11] Top golfers would be better on average on all of these compared to medium golfers, and the indicators would therefore be correlated. However, among medium golfers, the correlations would be lower, and the compensatory nature of the different indicators would show more clearly. The measurement philosophy is not that some underlying overall ability causes skill at each part off the game; rather, it is skill at the various aspects that build up overall ability.

theorizing and testing hypotheses regarding different forms of growth separately, rather than combining growth indicators in a summated index (Davidsson & Wiklund, 2000; Delmar, Davidsson, & Gartner, 2003; Shepherd & Wiklund, 2009).

6.4 Taking Validity Seriously

It is still quite common in entrepreneurship research that the only evidence delivered in support of the quality of the operationalizations used are (a) references showing the measure has been used before and (b) the magical Cronbach's alpha coefficient (cf. Slavec & Drnovšek, 2012). Neither provides any direct evidence that the measure captures what it is supposed to measure. Prior use proves only that the measure has been used previously, not that it has been proven valid. Cronbach's alpha is an indicator of internal consistency of a multiple-item index, an indicator of reliability.[12] Along with the 5 % rule for statistical significance, the magical number 0.70 (or even 0.80) on the Cronbach's alpha test (Nunnally, 1967; Nunnally & Bernstein, 1994) is one of the great examples of blind following of rules in research. To see how limited information this test gives, consider the examples in Fig. 6.2.

In this figure, "X" denotes the theoretical construct, whereas "x" denotes what our operationalization actually captures. Panel (a) shows the ideal situation when we have a perfectly valid measure. This is what we want to achieve. Panel (b) shows what we are more likely to end up with most of the time (*X* here covers the circle; *x* the entire irregular shape that partially overlaps it). This is an operationalization, which captures most of what it is intended to measure but also contains considerable

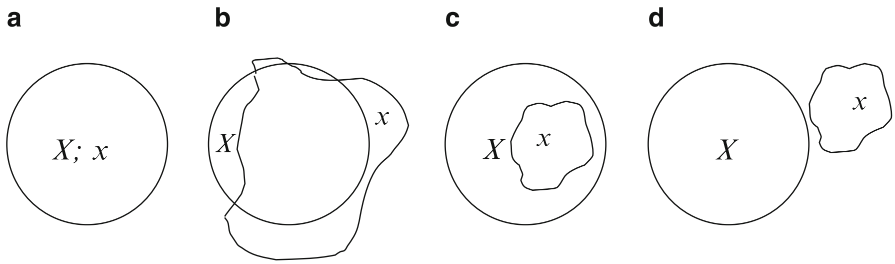


Fig. 6.2 Relationships between theoretical constructs and operationalizations

[12] A single alpha coefficient is obviously not indicative of test-retest reliability, which is the type of reliability most definitions imply. The logic underlying inter-item correlations as a measure of reliability is roughly this: Assume we have a pool of items supposed to reflect the theoretical construct. If we repeatedly draw and sum up n items at a time from this pool, will the score for an individual case, and therefore the cases' scores relative to one another, be essentially the same regardless of which subset of n items we happen to use? Something of that sort.

noise—systematic or random components that do not reflect the theoretical construct we are after.

Importantly, if we make internal consistency (Cronbach's alpha) the sovereign criterion, we are likely to end up with a situation like (c). This is particularly likely if we want to develop a measure with a small number of items, because the shortcut to a high alpha value with a small number of items is to have items that are very similar. They will then have high intercorrelations and therefore yield a high alpha value. However, *they may be far from exhausting the theoretical concept we are after; i.e., they lack or have incomplete content validity.* Concerns of that nature have contributed to the rise of formative measurement and the introduction of C-OAR-SE. The logic of the alpha criterion—as well as the fit indices used when we run SEM models—lures us to measure with high precision but a corner of the theoretical construct we're after. In short, a high alpha does not prove validity. In fact, it is entirely possible to have a high alpha value also in situation (d), i.e., when we measure with high precision something entirely different from the theoretical construct we were after.

Researchers who are serious about validity should do more than checking factor structure and Cronbach's alpha. These are but two out of the *13* criteria for evaluation of validity/reliability suggested by Robinson, Shaver & Wrightsman (1991; for an application of these useful criteria, see Brown, Davidsson, & Wiklund, 2001; cf. also Slavec & Drnovšek, 2012). However, rather than learned rules of what is "required" or what you can "get away with," I think the starting point should be—quite simply—*can I really trust my measures?* Do *I* honestly believe that application of these measures will lead us closer to the truth about the phenomenon I am investigating? If I don't, what would be the point of doing the research? And this applies to all kinds of data collection, including information from open-ended interviews and supposedly factual information in archival data. Using this mind-set, you would likely reinvent many of the validity notions and tests whether or not you knew their common labels.

Applying this type of mind-set to my very first piece of research, I supplemented interviews with small firm managers about their thoughts and feelings about growing the business with a conjoint analysis task (see Lohrke, Holloway, & Woolley, 2010) that systematically compared their preference for different, future growth scenarios—a *triangulation* effort if you will (Davidsson, 1986). In *Business Dynamics in Sweden* (Davidsson, Lindmark, & Olofsson, 1994a, 1994b, 1996, 1998), we did not simply accept archival data "as is." We discussed data quality issues with experts at Statistics Sweden, picked the best alternative indicator when several were available, examined the data for local firms about which we had some knowledge, and even cross-checked data by interviewing firms (see first edition of this book, Chap. 7, for details). If we could not trust the data, why would we do the project?

A short of perfect—and therefore realistic—example of validation seeking in survey research can be drawn from my dissertation study (Davidsson, 1989). I included a measure of *Need for Achievement* (*n*Ach) that I had developed myself, based on McClelland comprehensive description of this theoretical notion (McClelland, 1961). It was four-item measure with five-point response scales. In translation, the items read:

1. I have always wanted to succeed and accomplish something in my life.
2. I find it hard to understand people who always keep on striving toward new goals although they have already achieved all the success they could possibly have imagined (reversed).
3. To face new challenges and to manage to cope with them is important to me.
4. I am so satisfied with what I have achieved in my life that I think now I can confine myself to keeping what I already have (reversed).

My research concerned the extent to which *n*Ach influenced growth and other indicators of continued entrepreneurship in small firms. With Grand Old Man McClelland (influenced by Freudian psychology of the subconscious) claiming that *n*Ach cannot be measured in this direct manner and a Cronbach's alpha of just 0.55, I would seem to be on thin ice. However, neither McClelland nor Cronbach nor Nunnally (nor Davidsson) is an indisputable authority. So rather than just giving up, I first scrutinized "The Achieving Society" (McClelland, 1961) for claims unrelated to the contents of my items about high *n*Ach people, for which I also happened to have data in my study. I found the following:

– They are moderate risk-takers. They like to take some objective risks but are not attracted to games of chance.
– Profit is important to them as a measure of success and not for its own sake.
– Ownership control is not critical to them.
– They prefer experts over friends as workmates.

I then checked the empirical patterns in my data and found that these predictions about people with high achievement motivation were, by and large, borne out. As a result, I could conclude that "it seems that the *n*Ach-index measures a psychological difference between subjects and that labeling this difference 'need for Achievement' is reasonably well justified" (see Davidsson, 1989, pp. 164–165 for details). With this support for the validity of the operationalization, the substantive relationships I hypothesized and found between *n*Ach and entrepreneurial behavior could not as easily be disregarded as effects of method artifacts or some alternative theoretical variable. Moreover, the substantive analyses gave additional support for the validity of the measure, as I could show that *n*Ach—as the theory would predict—was positively related to growth aspirations for those managers who believed that growth would lead to increased profits, but not for those who did not hold such beliefs. This more than any other test convinced me that my simple, questionnaire measure actually captured *n*Ach to a reasonable degree (although revisiting it today, I must say item 4 is hopelessly age sensitive…).

Although admittedly being a far from perfect example about a far from perfect measure, I hope it has effectively illustrated the underutilized principle of seeking validation for a measure through testing its *construct validity* regarding substantial relationships postulated by theory. For reasons discussed above in relation to Fig. 6.2, it could well be argued that some evidence of this kind should always be presented; without it, the technical criteria like factor structure or Cronbach's alpha have little meaning.

6.5 Some Balancing Exercises

When choosing or developing operationalizations, there are several partly contradictory interests that have to be balanced. One is the choice between the "perfect" operationalization for a specific type of venture or context versus the most broadly applicable operationalization. This is one of the method consequences of the heterogeneity of the entrepreneurship phenomenon discussed in previous chapters. The best measure of firm size may be the number of vehicles for a taxi company, the number of seats for a restaurant operation, and the quantity of electricity delivered for a power station. However, how are we to compare the firms' size or performance across these different measures? Sales and number of employees are more generally applicable but may have other disadvantages (Davidsson & Wiklund, 2000; Shepherd & Wiklund, 2009).

We may want to solve that dilemma through narrow sampling and/or by using a measure that is suitable for the type of sample we have access to, be it broad or narrow. This brings in the question of using an existing measure vs. developing a purpose-built one. Consider, for example, the nine items commonly used to assess *entrepreneurial orientation* (EO). This EO measure can be accused of all sorts of shortcomings. For example, the items seem to be a mix of preferences, past behaviors, and beliefs. Further, it has been argued that some items do not fit the three strategic dimensions that the measure it intended to capture, namely innovation, risk-taking, and proactiveness (Lumpkin & Dess, 1996). For these reasons, it would be tempting for a researcher to develop a new and better measure for assessing firm-level propensity for entrepreneurial behavior. However, using the EO scale has huge advantages. First, the measure has theoretical backing (Miller & Friesen, 1978, 1982). Second, several studies have investigated the conceptual and empirical properties of the scale (Covin & Wales, 2012; George & Marino, 2011; Kreiser, Marino, & Weaver, 2002; Runyan et al., 2012). Third, lots of studies have applied the measure in different contexts and found meaningful results (Rauch, Wiklund, Lumpkin, & Frese, 2009; Rosenbusch, Rauch, & Bausch, 2013).

With a new measure all these advantages would be lost. My personal experience is also that developing useful new measures may be harder than many people realize. For example, for my *Culture and Entrepreneurship* studies (Davidsson, 1995; Davidsson & Wiklund, 1997), I tried a total of eleven different three- to five-item measures of "general values" intended to capture different dimensions of regional differences in mentality. Despite varying degrees of theoretical anchoring and pretesting efforts, it was only for two indices that I reached Cronbach's alpha values above 0.70, namely for those I borrowed from Lynn (1991). These were previously tested both in the English original and Swedish translation. Again, internal consistency is not all you should ask of a measure, but my relative failure to achieve very strong results in these studies was, in all likelihood, partly due to the questionable quality of some of the operationalizations. If measures of reasonable quality are available, you may be better off using them than trying—and likely failing—at developing the "perfect" operationalization.

Yet another balancing issue is the trade-off between maximizing content validity on the one hand and response quality, response rate, and coverage of possible sources of unmeasured heterogeneity on the other. If we look at a complex construct in isolation, we may find that a 20-item battery is the ideal way of measuring it. However, this may lead to respondent fatigue, so that the quality of their responses to individual items drops; they leave items unanswered, or they opt out of the survey altogether. Further, to include the long, "perfect" operationalization, we may have to sacrifice inclusion of other variables in order not to make total questionnaire length go out of hand. What to do? Well there are actually techniques you can apply that do not require that every respondent answer every item or every wave of data collection. With planned partial non-response, you get the correlations you need, and with modern imputation techniques, you may end up with better overall quality by trying to collect less data rather than more (see, e.g., Little & Rhemtulla, 2013).

Yet another issue to balance is the extent to which you customize the data capture (in primary data collection) based on what is already known about the case. We have noted repeatedly that entrepreneurship typically is a very heterogeneous phenomenon. As a result, all questions or particular phrasing of questions do not apply equally well to all cases. If you do not customize the questionnaire, you may end up with crappy data or respondents dropping out because they feel you obviously did not listen and take account of the information they already provided you with on prior questions. In CAUSEE, there must be a million different ways you can go through the interview in terms of skip patterns and wording variations based on the type of founders, type of venture, and its stage of development. As a result, there was considerable programming and testing costs, and due to all the variations the questionnaires ended up printing out to over 100 pages in later waves (see eprints.qut.edu.au/49327/). However, these adaptations also mean you show respect for the respondent and obtain higher response rates (85 % or above of all eligible cases in waves 2–5) and higher response quality. The below illustrates just a couple of examples of how team size, venture type, status of the venture and the founder, and other information provided earlier in the interview or in previous waves determine whether a question is asked and how it is worded.

ASK IF W1-Q43 NE 1 or MISSING

B22	**OPERATIONAL (A41 = 1 OR 2) SOLO TEXT:** Have you joined any Internet-based networks or communities for the purpose of helping the development of this business, will you do so in the future, or is this not relevant to this business?
	OPERATIONAL (A41 = 1 OR 2) PARTNER TEXT: Have you or your partner joined any Internet-based networks or communities for the purpose of helping the development of this business, will you do so in the future, or is this not relevant to this business?
	OPERATIONAL (A41 = 1 OR 2) TEAM TEXT: Have you or other owners joined any Internet-based networks or communities for the purpose of helping the development of this business, will you do so in the future, or is this not relevant to this business?
	TERMINATED (A41 = 3) SOLO TEXT: Did you ever join any Internet-based networks or communities for the purpose of helping the development of this business? [IF "NO" record "3"]
	TERMINATED (A41 = 3) PARTNER TEXT: Did you or your partner ever join any Internet-based networks or communities for the purpose of helping the development of this business? [IF "NO" record "3"]
	TERMINATED (A41 = 3) TEAM TEXT: Did you or other owners ever join any Internet-based networks or communities for the purpose of helping the development of this business? [IF "NO" record "3"]

1	Yes
2	No, not yet—will in the future
3	No, not relevant

ASK IF C20 = 2

C21 (W1-Sect1[GA1]-Q5a)	**ENGAGED (A40 = 1 OR 2)/PRODUCT TEXT:** Is the product that this new business will sell completely developed and ready for sale or delivery, has it been tested with customers as a prototype, is it being developed as a model, or is the product still in the idea stage?
	ENGAGED (A40 = 1 OR 2)/SERVICE TEXT: Is the service that this new business will sell completely developed and ready for sale or delivery, has it been tested with customers as a procedure, is it being developed as a procedure, or is the service still in the idea stage?
	REDUCED (A40 = 3)/PRODUCT TEXT: To the best of your knowledge, before or after your involvement ended, was the product that this new business would sell completely developed and ready for sale or delivery, had it been tested with customers as a prototype, was it being developed as a model, or was the product still in the idea stage?
	REDUCED (A40 = 3)/SERVICE TEXT: To the best of your knowledge, before or after your involvement ended, was the service that this new business would sell completely developed and ready for sale or delivery, had it been tested with customers as a procedure, was it being developed as a procedure, or was the service still in the idea stage?

1	Completed and ready for sale
2	Prototype/procedure tested with customers
3	Model/procedure is being developed
4	Still in the idea stage; no work done yet

6.6 Entrepreneurship-Specific Operationalization Challenges

This said, for many core constructs in entrepreneurship, no satisfactory measures may as yet be available. Therefore, there is still a need to develop validated operationalizations of such things as *entrepreneurial cognition*, *entrepreneurial strategy/action/behavior*, *opportunities* (if you insist) or *new venture ideas*, and *new venture outcomes.* I will discuss some such attempts below, although I will defer the discussion of conceptualization and operationalization of outcomes until the next chapter.

6.6.1 Operationalizing Effectuation and Bricolage

Despite its widespread popularity, we still do not have an agreed-upon operationalization of the notion of effectuation (Sarasvathy, 2001, 2008) and its subdimensions. The reason for this may in part be conceptual (Is it cognitive, behavioral, or something else? Do all the current dimensions [and no other] belong there? Do we have reason to expect them to be highly correlated? Would the latter apply to all people or only to expert entrepreneurs? Is effectuation really [always] the opposite of causation?). In part, it may just be that we have not yet found the right way to capture this intriguing construct. The most promising attempt to date is arguably the work by Gaylen Chandler and co-workers (Chandler, DeTienne, McKelvie, & Mumford, 2011; Chandler, DeTienne, & Mumford, 2007; Perry et al., 2012).

When we started the CAUSEE study, even less had been done toward operationalizing *entrepreneurial bricolage.* So we decided to try to do just that. Developing such a measure is a tough task as there is absolutely no guarantee that the complex set of behaviors theoretically described as bricolage can be meaningfully captured by a sufficiently small set of items in a questionnaire. Our work started with me approaching Ted Baker (as in Baker & Nelson, 2005) at the design stage of the project. Ted realized that at some point the emerging theory of entrepreneurial bricolage would have to be operationalized and tested in broadly based samples, so he was interested. I wanted some interesting, unique, and theory-based content in CAUSEE (relative to PSED I and II and other counterpart studies), and by involving Ted—the theory expert—I felt confident that at least *content validity* of our items would be secured.

Theoretically, the measure should capture aspects of "making do," using "resources at hand," and "recombination of resources for new purposes" (Baker & Nelson, 2005). Aiming to capture a unidimensional "gestalt" of bricolage behavior rather than three separate dimensions, Ted tried to include more than one aspect of

Table 6.2 The CAUSEE (Baker-Davidsson) operationalization of entrepreneurial bricolage

READ: The following statements are about how your business uses various kinds of resources to deal with new challenges. By resources, we mean things like materials, equipment, people, or anything else that can be used to get a job done. By challenges, we mean both new problems and new opportunities. When I say "we" or "our," I mean you personally or anybody else acting on behalf of the business. I want you to respond on a scale from 1 to 5, where 1 means "never" and 5 means "always." OK, does the following represent how you never, rarely, sometimes, often, or always go about doing things for your start-up? Firstly, … **READ STATEMENT**
1. We are confident of our ability to find workable solutions to new challenges by using our existing resources
2. We gladly take on a broader range of challenges than others with our resources would be able to
3. We use any existing resource that seems useful to responding to a new problem or opportunity
4. We deal with new challenges by applying a combination of our existing resources and other resources inexpensively available to us
5. When dealing with new problems or opportunities, we take action by assuming that we will find a workable solution
6. By combining our existing resources, we take on a surprising variety of new challenges
7. When we face new challenges, we put together workable solutions from our existing resources
8. We combine resources to accomplish new challenges that the resources weren't originally intended to accomplish
9. To deal with new challenges, we acquire resources at low or no cost and combine them with what we already have

bricolage in most items, in apparent but carefully considered breech of standard textbook advice not to "double barrel" questions.

After discussions among colleagues and pretesting on a sample of analyzable size, we settled for the nine items (about the maximum number respondents would accept without finding the repetition too tedious) in Table 6.2. One of my more important contributions was to suggest the response scale be *never-always* rather than *agree-disagree* to reflect the behavioral nature of the construct. For the same reason, we strove to make the main verb in each item clearly behavioral. In line with textbook advice, we considered reverse- or negatively worded items but ruled them out on the grounds that there is no theoretically or logically clear "opposite" of bricolage.[13] The pretesting had also indicated—in line with what I had experienced in prior research—that "resources" is a tricky concept for many respondents. We were working with a very diverse sample, so we could not remedy this by exemplifying much in the item wording. Our solution was to develop a rather detailed introduction to the section, to be read to the respondents.

With a notion like bricolage, it is easy to realize that a single item cannot possibly capture the construct. By using multiple items, we can capture the construct better while reducing random measurement error. I hope an inspection of the items will convince you that the measure has content validity. Further, internal non-response did not reach

[13] Besides, the merits of reverse items are not as clear as some textbooks would have it (Barnette, 2000; Locker, Jokovic, & Allison, 2007; Stewart & Frye, 2004; Wong, Rindfleisch, & Burroughs, 2003).

Table 6.3 Internal consistency of the CAUSEE bricolage measure ($N = 60–1405$)

Sample	# Factors (eigenvalue > 1)	Variance extracted, factor 1	Loading range	Cronbach's alpha
W1: All cases	1	45	0.57–0.72	0.82
W1: Nascent-Regular	1	45	0.62–0.71	0.82
W1: Nascent-High Potential	1	48	0.51–0.80	0.84
W1: Young Firm-Regular	1	48	0.57–0.73	0.83
W1: Young-High Potential	2	36	0.42–0.74	0.74
W2: All	1	47	0.64–0.73	0.84
W2: Nascent-Regular	1	46	0.59–0.74	0.83
W2: Nascent-High Potential	2	44	0.49–0.80	0.80
W2: Young-Regular	2	49	0.63–0.75	0.85
W2: Young-High Potential	2	46	0.57–0.73	0.82
W3: All	1	50	0.67–0.78	0.85
W3: Nascent-Regular	2	48	0.55–0.78	0.84
W3: Nascent-High Potential	2	48	0.56–0.83	0.83
W3: Young-Regular	1	52	0.66–0.80	0.87
W3: Young-High Potential	1	44	0.52–0.78	0.82
Examples of other breakdowns				
W2: All product firms	1	47	0.59–0.77	0.84
W2: All services firms	1	46	0.65–0.72	0.83
W2: All product-service mix	1	49	0.59–0.79	0.84
W2: All solo founder firms	1	47	0.62–0.74	0.84
W2: All team firms	1	46	0.60–0.72	0.83
W2: All university educ. resp.	1	42	0.60–0.73	0.80
W2: All no university educ.	1	50	0.62–0.78	0.85

During the publication process these analyses were rerun with slightly different results. The corrected analyses will appear in a forthcoming paper by Davidsson, Baker and Senyard

dramatic levels; the highest non-response occurs for item 8 (31 cases, or 2.2 %), indicating the respondents did not have much problem making sense of the statements.

We regard our bricolage measure as reflective and hence performed the analyses in Table 6.3 (and more). Because we had a large sample (1410 cases across four subsamples) and multiple waves of data collection (W1–W3), we could perform not just one but many different tests, greatly increasing the informational value of the results. The displayed results are based on an eight-item index, because after initial tests, we dropped item 9 due to somewhat poor correlations in some subsamples. We also identified a theoretical reason for this problem: our use of the verb “acquire” may have triggered the unintended interpretation “buy at full cost” rather than “obtain at little or no cost” (in later waves we therefore changed “acquire” to “access” with good results; however, the displayed analyses do not include this item).

In each wave, the analysis using all cases yielded the single factor we aimed for. The same holds for all subcategories analyzed (only W2 results displayed). In some analyses on smaller samples, a second factor is extracted; however, this second factor consistently has an eigenvalue just marginally above 1 (1.04–1.20). One analysis yields a result that is short of satisfactory, namely “W1: Young-High Potential.” Otherwise, the variance extracted by the first factor is essentially stable, around 45–50 % across all analyses. This may seem low-ish as does the (lower end of the) loading range, but these indicators are not unsatisfactory given that there are sub-components of bricolage which are not evenly represented across items. The Cronbach’s alphas exceed commonly accepted values, and there seems to be no real reason for concern that the measure works much better for some types of firms, stages, or founders, i.e., *uneven validity* (although there is a possible curiosity associated with university-educated founders). All in all the results are satisfactory; the only wart (W1: Young-High Potential) derives from a small sample and the same sample performs well in W3, so the W1 results may well be due to stochastic reasons.

So far, so good—but not all is necessarily cool and dandy. Figure 6.3 reveals that when we compute a summated index, the lower (left) half of the distribution closely corresponds to a normal curve. To the right, we see that close to 9 % of the respondents have the maximum score on all items in the index. This could indicate that variance within this group is not adequately captured by the index (right-truncation). It could also indicate that many in this group have given exaggerated and not very well-considered responses. It is an issue worthy of further examination. Underlying this distributional problem are high item means, between 3.37 and 4.45 on five-point scales in the combined sample in W1. Especially for those items where “always” was the modal response, the standard solution of making item wording or response alternatives more extreme is a standard solution that could be tried.

Thanks to having multiple waves we can also assess *test-retest reliability*. Although we would not expect founders to use the exact same level of bricolage year in and year out, finding very low correlations across years would be an indication the measure isn’t worth much. We find the W1–W2 and W2–W3 correlations to be 0.42 and 0.48, respectively, in the overall sample. The W1–W3 correlation is slightly lower (0.39) which is logical if there are real changes in behavior over time. Just to make sure these correlations are not driven by those responding “always” to everything, I ascertained that the same pattern holds up if we exclude them (0.39, 0.41, and 0.28). The pattern also stands up in various subsamples.

More importantly, does the measure behave as expected in relation to other variables? Although it is hard to hold many strong theoretical beliefs regarding this embryonic theory, one expectation I would have is that the use of bricolage is reduced over time. As firms become more established, their behavior is likely to

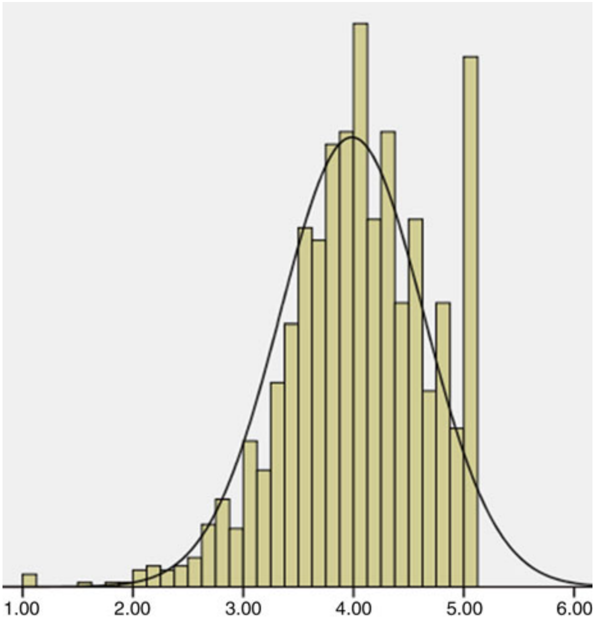


Fig. 6.3 Distribution of scores on the summated Bricolage index

become more mainstream (Baker & Nelson, 2005). This prediction is borne out. The mean on the bricolage index is 4.01, 3.86, and 3.79 among survivors in the overall sample, and although the annual decline may seem small, it is "highly significant" in statistical terms. The pattern again holds up in various subgroups although in smaller subsamples the difference is too small to be regarded statistically significant at conventional levels.

I would also expect founders with significant prior start-up experience to use more bricolage. This, too, holds up in all three waves in the overall sample. A more stringent analysis uses only identical, surviving cases with complete data. Such an analysis shows that in the nascent venture random sample, those with experience from at least two prior start-ups have bricolage scores of 4.19, 3.93, and 3.80 over the three waves, with novice founders scoring 3.88, 3.76, and 3.72. In the Young Firm random sample, the corresponding series are 4.14–3.88–0.3.84 vs. 3.87–3.77–3.58. In both cases, the group differences are statistically significant in the first two waves. Those with experience from one prior start-up tend to fall neatly in-between. By contrast, there are no differences by sex in the use of bricolage.

Yet another justified suspicion is that those using bricolage extensively would undertake more changes to their "business model" or "new venture idea," in the spirit of tinkering with limited resources. Correlating the bricolage score with another index capturing the numbers of reported changes in the concurrent and ensuing periods yields positive coefficients in the 0.11–0.18 range ($p<0.01$ in all cases). Positive sign holds up in seven of the eight subsample analyses I performed,

Table 6.4 Innovation items from the entrepreneurial orientation scale (Covin & Wales, 2012)

1. In general, the top managers of my firm favor…		
Strongly emphasize the marketing of the company's present products	1 2 3 4 5 6 7	Strongly emphasize R&D
2. How many new lines of products or services has your company introduced over the past 5 years (or since its establishment)? [reverse item]		
A lot of new products/services	1 2 3 4 5 6 7	No new products/services
3. Changes in product or service lines…		
Have been mostly of a minor nature	1 2 3 4 5 6 7	Have usually been quite dramatic

with five also being statistically significant at conventional levels (i.e., $p < 0.05$). Significant relationships with innovativeness and financial performance have been reported in Senyard, Baker, Steffens, and Davidsson (2014) and Senyard (2015).

Convinced yet? In all honesty, you shouldn't be completely convinced, but that also goes for most measures you see in the literature. Although I find the above encouraging, we would like to have more evidence before we are fully satisfied with a measure like this one. First, it would be nice to have additional examples of theoretically predicted relationships while not finding correlations where they should not occur. Second, it would be *very* nice to have evidence of *convergent validity* with some more close-up, comprehensive assessment of bricolage behavior for a sufficient number of cases. Julienne Senyard actually took a look into the issue when doing case studies among the surveyed firms. It was not a large-scale assessment, but at least the exercise did not become a cause of alarm. At this point, I would say we have *a* measure that seems good enough to be better than having *no* measure.

6.6.2 Operationalizing Novelty

In Chap. 1 I discussed "degrees of entrepreneurship," which suggests that "novelty" or "innovativeness" is one of the most important characteristics of new ventures and new venture ideas. Prior attempts at assessing novelty illustrate the range of alternatives that are available as well as their relative strengths and weaknesses. Consider first the innovation items in the EO scale (Table 6.4). This operationalization is supposed to be applicable across industries. This necessitates some vagueness and weak quantification ("a lot"; "minor"), and yet one might wonder whether an item like (2) can have satisfactory instrument equivalence (even validity) across firms in, e.g., construction in retailing. Despite its apparent breadth, the measure is restricted to product/service innovation in a corporate context. Quite a bit of adaptation would be needed before using it for nascent ventures (note, e.g., the expectation of years of history and multiple product/service lines, and reference to "R&D" which may not reflect the natural language of services start-ups even if they are highly innovative).

Toward the other extreme, Cliff, Jennings, and Greenwood (2006) developed a context-specific and comprehensive operationalization of *organizational* innovation among law firms. After familiarizing themselves with the industry through literature and qualitative work, they had six experts first describe the "standard" ways of organizing activities along 15 dimensions. They then had the same experts assess how much alternative practices deviated from this norm. With this information they could collect objective indicators of those 15 characteristics through face-to-face interviews. The deviations from norm were finally combined into a unidimensional index of the extent of organizational innovation displayed by each firm. Arguably, such a measure should have much higher validity and reliability—but suffers of course from not being directly transferrable to other contexts.

In CAUSEE we wanted to assess novelty of the new venture ideas in a broadly based sample. To make the challenge even greater, we wanted to include several types of novelty, not just product innovation. Inspired by Schumpeter's (1934) five types of "economic development" (but dropping "reorganizing an entire industry," of which we did not expect [m]any instances), we expanded on work by Dahlqvist and Wiklund (2012) in developing a measure adapted for service-dominated times and the reality of nascent and young ventures. We asked about novelty in four areas: (a) product/service, (b) production/sourcing, (c) marketing methods, and (d) target market selection. The exact wording was adapted to each domain of novelty and type of firm (nascent vs. young; product vs. services) but followed the same structure. The following example is for product novelty among nascent, product-producing firms:

Q1.	Will you offer a product which is entirely new for [respondent's industry inserted]? (If yes, tentative score = 2 and go to Q2; otherwise, tentative score = 0 and go to Q3)
Q2.	Will the product be entirely new to the world or entirely new just in the places where you are going to be active? (If yes, final score = 3; else final score = 2)
Q3.	If not entirely new, will the product be a substantial improvement compared to what other businesses have offered before? (If yes, final score = 1; else final score = 0)

The measures of the four types of novelty can be used separately. However, we also wanted to combine them to an index of "total novelty," which consequently can have scores between 0 and 12. This would be a *formative* index, because we do not assume (in the main) an underlying, overall urge to be novel to manifest itself in each of the items. Instead, each venture (founder) can choose to be innovative or imitative on each dimension independently. Hence, the traditional, psychometric evaluation of factor structure and intercorrelations (Cronbach's alpha) do not apply. However, there are still some things we can do to evaluate our measure:

Content validity. The dimensions we include would seem relevant, but are they exhaustive? Perhaps we should have included (more clearly than production/sourcing) also organizational innovation (cf. above)?

Test-retest reliability. Like with bricolage, we have repeated measures through three waves (W1–W3), but even with perfect reliability we should not expect perfect correlations due to real changes in the level of novelty. It turns out that correlations

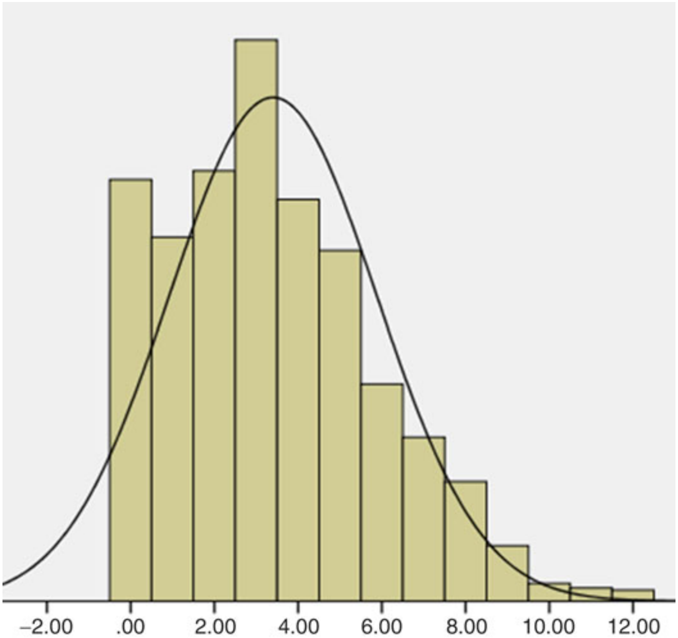


Fig. 6.4 Distribution of scores on the summated Novelty index

for the total index across waves vary between 0.62 and 0.72, i.e., pretty high, with the highest correlations occurring for W2–W3 comparisons, which is logical if there are fewer real changes to the new venture idea over time.

Predictive validity. Would you expect experienced entrepreneurs to work on more novel new venture ideas compared to novices? I would say yes. This is clearly borne out; with the same type of comparison we did for bricolage, the experienced founders come out 1.18, 1.14, and 1.54 scale steps higher than the novices in W1, W2, and W3, respectively. This is a big difference for this type of scale and statistically significant by a country mile or two. There is also a difference by sex, although smaller than the experienced-novice contrast (0.54, 0.73, 0.73). In a minute or so, you will have reason to ponder whether this reflects a difference in innovativeness or in honesty/realism. The novelty index is also consistent with other indicators of novelty such as agreeing that R&D is a main priority for the business (i.e., convergent validity if you will).

So far, things look pretty good. What about the distribution of the novelty index? Figure 6.4 gives the answer. Although the shape is not extreme, it looks like we here have the mirror "problem" of left-truncation, compared to the bricolage measure. However, we know most start-ups are imitative—suggesting zero-inflation—so there is little reason to expect a normal curve in this instance (Crawford et al., 2015). In fact, the normal-looking right tail should concern us more. Five cases have score 12, meaning they rate themselves as "new to the world"—on every dimension of novelty! A full 44 (of 1410) rate their novelty *on average* as "new to the industry" across the four dimensions (i.e., a total score of 9 or above). Get real! Clearly there is some exaggeration going on in these self-reports.

There is also evidence that the respondents learn over time and realize that their ideas aren't as unique as they first thought. The mean self-reported total novelty among survivors in the full sample drops from 3.27 in W1 to 2.99 in W2 to 2.73 in W3. But how do we know this is not because they really make their venture less novel because some initial novelty attempt didn't work? Well, there is some of that going on as well. However, we know there is also initial exaggeration, because we had computer programming rigged to stop and ask about the reason when novelty scores between waves deviated non-trivially. Cool, uh?

The most common adjustment over time is in the direction of less novelty, and the most common reason reported for this is admittance of initial overestimation rather than changes of the new venture idea. Does this tendency to exaggerate prove our novelty measure useless? I would say: Not at all. First, if all respondents exaggerate equally, it does not affect estimated relationships. This said, it would no doubt be risky to relate our novelty measure to other self-report, agreement-type questions, and the possibility of response artifacts need to be carefully considered. Further, I would be more comfortable using the W2 or W3 response as there is clear indication that respondents learn to better assess their venture's true level of novelty over time.

Without correction, the measure should of course not be used to compare groups that are known to have unequal propensity for bias. It is possible in this case to find out how factors like (a) venture stage of development, (b) respondent sex, and (c) prior entrepreneurial experience are intricately related to (1) reported level of overall novelty in W1, (2) propensity to adjust downward the self-reported novelty from one wave to the next, and (3) propensity to attribute such downward adjustments to initial overestimation. Based on such knowledge, it is possible to correct for or at least properly caution against possible biases.

Applying these precautions I think this operationalization can be quite useful, especially given the many positive signs in other aspects of our validity check. Recalling the discus throwing analogy and the crudeness of this type of measure, I would say that for the most part, the risk of under-estimation of real relationships is much greater. Besides, many of the other measures you and others are using are likely subject to similar problems or worse—it is just that nobody collects or reports the data that can show it.

6.6.3 Operationalizing Entrepreneurial Action

Bill Gartner (1988) made a strong, early argument that entrepreneurship researchers should focus on what entrepreneurs *do,* i.e., on entrepreneurial *action*, or behavior. Recently, other strong voices have made similar arguments (Dimov, 2011; Klein, 2008; McMullen & Dimov, 2013; McMullen & Shepherd, 2006; Venkataraman, Sarasvathy, Dew, & Forster, 2012). These calls have not remained unheeded. For example, the emerging theories of effectuation and bricolage, and associated operationalizations, can be said to capture types or patterns of action. To some extent perhaps even entrepreneurial orientation can be added here.

This said, one of the most important developments in this area in recent times has arguably been the capturing of a large number of "gestation activities" in the PSED and its international counterpart studies. Table 6.5 displays a rather long list of such activities.[14] These have turned out to be among the most useful and versatile questions included in these studies. It has proven especially valuable that each behavior is time-stamped, i.e., that each affirmative answer is followed by a question about in what year and month this behavior was initiated or completed.

Individually and collectively, these questions have been used (a) for defining the start and end point of the venture creation process and hence to define and trim the valid sample; (b) to describe different types of venture creation processes and activity patterns within them; (c) as a basis for reorganizing the data set from an interview-timeline to a venture-timeline logic, and as (d) independent, (e) control, or (f) dependent variables in exploratory and theory-driven analyses of relationships (e.g., Brush, Edelman, & Manolova, 2008; Delmar & Shane, 2004; Dimov, 2010; Gordon, 2012; Liao & Gartner, 2006; Liao, Welsch, & Tan, 2005; Lichtenstein, Carter, Dooley, & Gartner, 2007; Reynolds, 2007; Samuelsson & Davidsson, 2009).

However, the experiences to date have also revealed limitations and thus room for improvement (Davidsson & Gordon, 2012). This involves improvements that go beyond operationalization to include also aspects of conceptualization and design. For example:

1. Distinguishing between behaviors—that is, actions initiated by the focal actor—and the achieving of milestones that may be the fruit of previous action. For example, talking to customers is a behavior; reaching profitability is not. This should be a relatively easy distinction and some researchers have already employed it.
2. Dealing with retrospection. PSED is intended to follow processes in "real time"; however, on average about half of the relevant "gestation activities" have already been undertaken when the case is sampled. Further, with 12 months between interviews there is room for retrospection bias within waves as well. This can be remedied through even earlier "catch" and/or sample trimming, systematic analysis of biased spacing of activities that are reported prospectively vs. retrospectively, more frequent interview waves, and perhaps more intense following of (fewer) cases through experience sampling methodology (Uy, Foo, & Aguinis, 2010).
3. Conceptualization of types of "gestation activities." Improvement in this regard is perhaps the most important task for future research. Early attempts show that trying to include all manifest "gestation activities" as measured leads to more confusion than clarity (Liao & Welsch, 2008; Liao et al., 2005). Further, although the selection of gestation activities was in part inspired by Katz and Gartner

[14] I am getting a bit lazy towards the end of the chapter, so I am reusing the table from the first edition. See eprints.qut.edu.au/49327/ for an updated set with more Internet-related questions.

Table 6.5 Gestation behaviors in the Swedish PSED (cf. Davidsson & Honig, 2003)

Gestation activity	Question
1 Business plan (initiated)	Have you prepared a business plan?
1 Business plan (informally written)	What is the current form of the business plan? Is it unwritten (in your head), informally written, formally prepared, or something else?
1 Business plan (formally written)	
2 Development of product/service (initiated)	At what stage of development is the product or service that will be provided to the customers? (+ timestamp)
2 Stage of dev. of product/service	At what stage of development is the product or service that will be provided to the customers?—idea or concept
2 Stage of dev. of product/service	– Initial development
2 Stage of dev. of product/service	– Tested on customers
2 Stage of dev. of product/service	– Ready for sale or delivery
3 Marketing	Have you started any marketing or promotional efforts?
3 Marketing	Approximately how much time, in hours, have you and your partners, if any, devoted to marketing, i.e., finding and canvassing customers?
3 Marketing	Approximately how much money, own labor not counted, have you devoted to marketing, i.e., finding and canvassing customers?
4 Intellectual property	Have you applied for a patent, copyright, or trademark?
4 Intellectual property	Has the patent, copyright, or trademark been granted?
5 Obtained licenses (applied)	Has the new business obtained any business licenses or operating permits from any local, county, or state government agencies?
5 Obtained licenses (granted)	Has the new business obtained any business licenses or operating permits from any local, county, or state government agencies?
6 Physical resources	Have you purchased any raw materials, inventory, supplies, or components?
6 Physical resources	Have you purchased, leased, or rented any major items like equipment, facilities, or property?

Table 6.5 (continued)

Gestation activity	Question
7 Financial resources	Have you developed projected financial statements such as income and cash flow statements and break-even analysis?
7 Financial resources	Have you saved money in order to start this business?
7 Financial resources	Have you asked others or financial institutions for funds?
7 Financial resources	Has this activity been completed (successfully or not)?
7 Financial resources	Have you established credit with a supplier?
7 Financial resources	Have any of your customers helped financing the start-up?
8 Human resources—team (initiated)	Have you organized or helped to organize a start-up team?
8 Human resources—team (completed)	Is this in process or completed?
8 Human resources—workforce	Are you presently devoting full time to the business, 35 or more hours per week?
8 Human resources—workforce	Have you hired any employees or managers for pay, those that would not share ownership?
8 Human resources—workforce	How many employees are working full time for the new company?
8 Human resources—workforce	How many part time employees are working for the new company?
8 Human resources—knowledge	Have you taken any classes or workshops on starting a business?
8 Human resources—knowledge	How many classes or workshops have you taken part in?
8 Human resources—knowledge	Have you gathered any information to estimate potential sales or revenues, such as sales forecasts or information on competition, customers, and pricing?
8 Human resources—knowledge	Have you discussed the company's product or service with any potential customers yet?
8 Human resources—household help	Have you arranged childcare or household help to allow yourself time to work on the business?
9 Social capital/networking	In order to help get this business going, have you gotten involved in any business networks, such as a trade association, chamber of commerce, or a service club like the Lions or Rotary?
10 Contact channels	Does the company have its own phone number?
10 Contact channels	Does the company have its own mail address?

10 Contact channels	Does anyone in the team have a mobile mainly used for the business?
10 Contact channels	Does the company have its own visiting address?
10 Contact channels	Does the company have its own fax number?
10 Contact channels	Is there an e-mail or Internet address for this new business?
10 Contact channels	Has a Web page or homepage been established for this business?
11 Registration	Has the new business registered with PRV?
11 Registration	Has the company received a company tax certificate?
12 Subsidy (applied for)	Have you applied for an enterprise allowance ("starta-eget-bidrag")? (IF YES) Has the application been granted, is it under processing, or has it been refused?
12 Subsidy (granted)	Has the enterprise allowance been granted?
13 Exchange	Have you received any money, income, or fees from the sale of goods or services?
13 Exchange	Does the monthly revenue exceed the monthly expenses?
13 Exchange	(IF YES ABOVE) Do the monthly expenses include owner/manager salary in the computation of monthly expenses?

(1988), even those who use their categories *Intentionality*, *Resources*, *Boundaries*, and *Exchange* do not agree on which activity goes with which construct (Brush et al. 2008; Samuelsson, 2004). Further, others (including myself in Table 6.5) have suggested a range of conceptual labels for various groupings within essentially the same super-set of activities: *Personal Planning-Personal Preparation-Organizational/Financial Structure-Business Presence-Production Implementation* (Reynolds, 2007), *Assessing Market-Estimating Profits-Completing Groundwork-Structuring the Company-Setting up Operations* (Gatewood, Shaver, & Gartner, 1995), *Legitimation Activities-Generating Social Relationships-Transforming Resources* (Delmar & Shane, 2004), and a simple one that I have found promising: *Discovery-Exploitation* (Gordon, 2012; cf. Shane & Venkataraman, 2000). Focusing on activity patterns over time rather than on what particular actions were undertaken, Lichtenstein et al. (2007) used the categories *Rate-Timing-Concentration* (cf. Gordon, 2012; Hopp & Sonderegger, 2014). It would be handy if at some point we could reach some agreement on a suitable set of constructs and operationalizations of these!

4. Determining the relative importance of different actions as indicator of a theoretical construct. The list in Table 6.5 seems to include items that vary quite a bit in relative importance to the success of a start-up venture. However, researchers have had little objective basis for discriminating among them and have resorted to giving them equal weight or to subjectively pick only those appearing most important. PLS analysis offers a way to empirically assess the relative importance of alternative indicators (Ringle et al., 2012).
5. Distinguishing between true "gestation activities" undertaken specifically for *the current venture* and other preceding actions that at the time were not undertaken for the purpose of the focal venture but which may nonetheless have effects on its success chances. Actions that benefit the current venture may have been done for other purposes at an earlier date, which may confuse the apparent timeline of the focal venture's development. A recent dialogue in *Journal of Business Venturing Insights* unveiled the potential importance of this distinction (Davidsson, 2015; Delmar, 2015; Honig & Samuelsson, 2014).
6. Dealing with the fact that not all activities apply to all ventures and thus that completion of a given number of activities does not mean the same rate or amount of progress across all ventures. The PSED questions actually include a "not relevant" response alternative, but this information has hardly ever been used by analysts.
7. Improved operationalization of the timing and amount of activity within each conceptual category. A major drawback with the PSED approach to capturing gestation activities is that it assesses their initiation and/or completion as one-off events to be "ticked off." There is no assessment of how much time and effort the founders direct to the conceptual categories under point (3) above in each time period. One early attempt actually did this (Gatewood et al., 1995) and some more sophisticated version of their approach is worth trying in future research.

6.7 Operationalization Issues on Aggregate Levels

Although so far I gave up the first edition's organization of this chapter by level of analysis, I would like to conclude with a brief discussion of operationalization issues on the aggregate level, i.e., industry, region, or nation. I'd like to focus the discussion here on operationalization of "entrepreneurship" (entrepreneurial activity) itself and on the two related issues in regard to this: the perils of *uneven validity* or lack of *measurement invariance* (Runyan et al., 2012; Schaffer & Riordan, 2003) and the development of richer indicators of entrepreneurship.

If there is uneven validity, we do not pick up the same phenomenon even if we try to apply exactly the same instrument across contexts. A telling example of this was when the early editions of the Global Entrepreneurship Monitor (GEM) applied just one global indicator of entrepreneurial activity (the proportions of adults involved in a business start-up or young firm) and countries like Uganda and Peru came out as the most entrepreneurial in the world (Ács, Desai, & Hessels, 2008). Even the wording within individual items may have differential applicability. I came across one example of this in the *Culture and Entrepreneurship* study, where my committing the Deadly Sin of double loading an item only became evident after trying to apply the same instrument in another country. This was one of the Autonomy items, which read "I have probably found it harder than others to let authorities like parents, teachers and superiors decide for me." The Swedish respondents had no problem with this item, but in Estonia, respondents protested against my lumping parents together with other authorities! Apparently, in the latter country obeying parents and obeying bosses have very different implications or connotations, whereas in Sweden (my home country at the time) they can all represent the same theoretical category "authorities." This shows how nonobvious cultural sensitivity to operationalizations can be. Remember that this also applies to the meaning of items across industries within the same country and language.

GEM is also a good example of heeding Zahra and Wrights (2011) call for richer indicators of entrepreneurship research. More recent editions not only consider the distinction between (perceived) opportunity and necessity as drivers, but also distinguish between more and less growth-oriented start-ups as well as their level of technological sophistication. Further, the monitor now considers the fact that entrepreneurship can manifest itself differently by including measures of employee entrepreneurship alongside data on independent start-ups. Further, comparisons are made among groups of countries at a similar level of economic development (Álvarez, Urbano, & Amorós, 2014; Amorós, Bosma, & Levie, 2013; Xavier et al., 2013). Others have added even more extreme indicators, notably the prevalence of "billionaire entrepreneurs" in order to distinguish "entrepreneurship" from "mere self-employment" (Henrekson & Sanandaji, 2014). I definitely agree that the stock of self-employment is not a suitable indicator of national levels of entrepreneurship. However, as per my reasoning in Chap. 1, I think all entrants carry out the entrepreneurial function to some degree, and as sole indicator of entrepreneurship the prevalence of self-made billionaires is equally problematic: it is subject to considerable time lag; captures value appropriation rather than value creation, and also reflects things like institutional arrangements and aggregate demand in the home market. As a supplement it is a welcome addition, though.

A venture is a venture is a venture is…not true. One of my main frustrations from working with *Business Dynamics in Sweden* and *Culture and Entrepreneurship* was that I got a strong sense that some regions with similar business start-up rates may have start-ups of very different average quality. But we did not have a measure of quality. A very interesting development in this regard is Guzman and Stern's work on the *Entrepreneurship Quality Index* and the *Regional Entrepreneurship Cohort Potential Index*, an ingenious effort that appears so successful that one wonders whether perhaps it is somewhat possible, after all, to "pick winners" at an early stage with a meaningful level of precision (Guzman & Stern, 2015a, 2015b). Their exact methodology cannot be used elsewhere and the approach may not work quite as well in regions with less representation of the top layer of entrepreneurial endeavors, but even so the general approach can likely inspire reasonably successful efforts in other types of environments as well.

6.8 Summary and Conclusion

I have argued in this chapter that we should take operationalization issues—and validity in particular—seriously. It is not simply about adhering to rules and trying to get away with what we have according to some standard criteria. Our preference for one measure over another should definitely *not* be that "it works," i.e., it supports the hypotheses you want to test. Making that the main criterion is just bad science, driven by confirmation bias. We sure have enough of that already.

Instead, we have to seriously think through what basis we have for believing that our measures capture the theoretical constructs of interest and to strengthen that basis by careful development work and analysis of the measure's properties and relationships with other construct. In this work, we need to balance the interest of "perfection" against prior use and validation. That balance often comes out in favor of the latter; trying to develop a perfect measure may lead to a sub-standard alternative in the end. It is harder than we might think. This said, entrepreneurship is still a young field and this means that the arsenal of existing operationalizations is not always sufficient.

We also need to balance context-specific versus more globally applicable operationalizations. The former, applied to relatively homogenous samples, can often allow a much richer and seemingly more valid operationalization, but this comes at the cost of reduced generalizability and removing possibilities of direct comparison across contexts. Then again, when we compare something across countries or industries using the same measure across these contexts, we need to consider the possibility that observed differences reflect method artifacts—lack of measurement invariance—rather than substantive differences. No space is entirely safe ground when doing research.

I have also emphasized that the "reflective index" measurement paradigm, with associated factor structure and Cronbach's alpha evaluations, is not the only show in town. There are other approaches to operationalizing tricky constructs, such as formative indices, C-OAR-SE, and conceptual disaggregation in order to achieve stronger links between constructs and their operationalizations. My examples of recent, entrepreneurship-specific operationalizations included both formative and reflective indices and demonstrated validation efforts which yielded both support and some cause for concern (or inspiration for further improvements). This is always the case—nothing is ever perfected.

Lastly, if you remember which golf tournament I used as a basis for discussing validity issues or remember the names of the top players or the countries they come from, I think I may have a valid indicator of you concentrating on the wrong things…

References

Ács, Z. J., Desai, S., & Hessels, J. (2008). Entrepreneurship, economic development and institutions. *Small Business Economics, 31*(3), 219–234.

Álvarez, C., Urbano, D., & Amorós, J. E. (2014). GEM research: Achievements and challenges. *Small Business Economics, 42*(3), 445–465.

Amorós, J. E., Bosma, N., & Levie, J. (2013). Ten years of global entrepreneurship monitor: Accomplishments and prospects. *International Journal of Entrepreneurial Venturing, 5*(2), 120–152.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329–366.

Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*(3), 361–370.

Bouckenooghe, D., De Clercq, D., Willem, A., & Buelens, M. (2007). An assessment of validity in entrepreneurship research. *Journal of Entrepreneurship, 16*(2), 147–171.

Brown, T., Davidsson, P., & Wiklund, J. (2001). An operationalization of Stevenson's conceptualization of entrepreneurship as opportunity-based firm behavior. *Strategic Management Journal, 22*(10), 953–968.

Brush, C. G., Edelman, L. F., & Manolova, T. S. (2008). Properties of emerging organizations: An empirical test. *Journal of Business Venturing, 23*, 547–566.

Carton, R. B., & Hofer, C. W. (2006). *Measuring organizational performance: Metrics for entrepreneurship and strategic management research*. Cheltenham, UK: Edward Elgar Publishing.

Chandler, G. N., DeTienne, D. R., & Mumford, T. V. (2007). Causation and effectuation: Measurement development and validation. *Frontiers of Entrepreneurship Research, 27*(13), 3.

Chandler, G. N., DeTienne, D. R., McKelvie, A., & Mumford, T. V. (2011). Causation and effectuation processes: A validation study. *Journal of Business Venturing, 26*(3), 375–390.

Cliff, J. E., Jennings, D. P., & Greenwood, R. (2006). New to the game and questioning the rules: The experiences and beliefs of founders who start imitative versus innovative firms. *Journal of Business Venturing, 21*, 633–663.

Covin, J. G., & Wales, W. J. (2012). The measurement of entrepreneurial orientation. *Entrepreneurship: Theory and Practice, 36*(4), 677–702.

Crawford, G. C., Aguinis, H., Lichtenstein, B., Davidsson, P., & McKelvey, B. (2015). Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.001.

Crook, T. R., Shook, C. L., Morris, M. L., & Madden, T. M. (2010). Are we there yet? An assessment of research design and construct measurement practices in entrepreneurship research. *Organizational Research Methods, 13*(1), 192.

Dahlqvist, J., & Wiklund, J. (2012). Measuring the market newness of new ventures. *Journal of Business Venturing, 27*(2), 185–196.

Davidsson, P. (1986). *Tillväxt i små företag: En pilotstudie om tillväxtvilja och tillväxtförutsättningar i små företag* (Small Firm Growth: A Pilot Study on Growth Willingness and Opportunity for Growth in Small Firms) (Studies in Economic Psychology No. 120). Stockholm: Stockholm School of Economics.

Davidsson, P. (1989). *Continued entrepreneurship and small firm growth.* Doctoral dissertation, Stockholm School of Economics, Stockholm.

Davidsson, P. (1995). Culture, structure and regional levels of entrepreneurship. *Entrepreneurship & Regional Development, 7*, 41–62.

Davidsson, P. (2015). Data replication and extension: A commentary. *Journal of Business Venturing Insights, 3*, 12–15.

Davidsson, P., & Gordon, S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics, 39*(4), 853–876.

Davidsson, P., & Honig, B. (2003). The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing, 18*(3), 301–331.

Davidsson, P., Lindmark, L., & Olofsson, C. (1994a). *Dynamiken i svenskt näringsliv (Business Dynamics in Sweden).* Lund, Sweden: Studentlitteratur.

Davidsson, P., Lindmark, L., & Olofsson, C. (1994b). New firm formation and regional development in Sweden. *Regional Studies, 28*, 395–410.

Davidsson, P., Lindmark, L., & Olofsson, C. (1996). *Näringslivsdynamik under 90-talet (Business Dynamics in the 90s).* Stockholm: Nutek.

Davidsson, P., Lindmark, L., & Olofsson, C. (1998). The extent of overestimation of small firm job creation: An empirical examination of the 'regression bias'. *Small Business Economics, 10*, 87–100.

Davidsson, P., & Wiklund, J. (1997). Values, beliefs and regional variations in new firm formation rates. *Journal of Economic Psychology, 18*, 179–199.

Davidsson, P., & Wiklund, J. (2000). Conceptual and empirical challenges in the study of firm growth. In D. Sexton & H. Landström (Eds.), *The Blackwell handbook of entrepreneurship* (pp. 26–44). Oxford, MA: Blackwell Business.

Delmar, F. (2015). A response to Honig and Samuelsson (2014). *Journal of Business Venturing Insights, 3*, 1–4.

Delmar, F., Davidsson, P., & Gartner, W. B. (2003). Arriving at the high-growth firm. *Journal of Business Venturing, 18*(2), 189–216.

Delmar, F., & Shane, S. A. (2004). Legitimating first: Organizing activities and the survival of new ventures. *Journal of Business Venturing, 19*, 385–410.

Diamantopoulos, A. (2005). The C-OAR-SE procedure for scale development in marketing: A comment. *International Journal of Research in Marketing, 22*(1), 1–9.

Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research, 61*(12), 1203–1218.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Marketing Science, 40*(3), 434–449.

Dimov, D. (2010). Nascent entrepreneurs and venture emergence: Opportunity confidence, human capital, and early planning. *Journal of Management Studies, 47*(6), 1123–1153.

Dimov, D. (2011). Grappling with the unbearable elusiveness of entrepreneurial opportunities. *Entrepreneurship: Theory and Practice, 35*(1), 57–81.

Edwards, J. R. (2011). The fallacy of formative measurement. *Organizational Research Methods, 14*(2), 370–388.

Finn, A., & Kayande, U. (2005). How fine is C-OAR-SE? A generalizability theory perspective on Rossiter's procedure. *International Journal of Research in Marketing, 22*(1), 11–21.

Gartner, W. B. (1988). "Who is an Entrepreneur?" is the wrong question. *American Small Business Journal, 12*(4), 11–31.

Gatewood, R. D., Shaver, K. G., & Gartner, W. B. (1995). A longitudinal study of cognitive factors influencing start-up behaviors and success at new venture creation. *Journal of Business Venturing, 10*, 371–391.

George, B. A., & Marino, L. (2011). The epistemology of entrepreneurial orientation: Conceptual formation, modeling, and operationalization. *Entrepreneurship: Theory and Practice, 35*(5), 989–1024.

Gladwell, M. (2007). *Blink: The power of thinking without thinking*. Boston, MA: Back Bay Books.

Gordon, S. R. (2012). *Dimensions of the venture creation process: Amount, dynamics, and sequences of action in nascent entrepreneurship*. Doctoral dissertation, Queensland University of Technology, Brisbane.

Grégoire, D. A., Shepherd, D. A., & Schurer Lambert, L. (2010). Measuring opportunity-recognition beliefs. *Organizational Research Methods, 13*(1), 114–145.

Guzman, J., & Stern, S. (2015a). *Nowcasting and placecasting: Entrepreneurial quality and performance. Working Paper 20954*. National Bureau of Economic Research.

Guzman, J., & Stern, S. (2015b). Where is Silicon Valley? *Science, 347*(6222), 606–609.

Hadwich, K., Georgi, D., Tuzovic, S., Büttner, J., & Bruhn, M. (2010). Perceived quality of e-health services: A conceptual scale development of e-health service quality based on the C-OAR-SE approach. *International Journal of Pharmaceutical and Healthcare Marketing, 4*(2), 112–136.

Hauser, J. R., Toubia, O., Evgeniou, T., Befurt, R., & Dzyabura, D. (2010). Disjunctions of conjunctions, cognitive simplicity, and consideration sets. *Journal of Marketing Research, 47*(3), 485–496.

Henrekson, M., & Sanandaji, T. (2014). Small business activity does not measure entrepreneurship. *Proceedings of the National Academy of Sciences, 111*(5), 1760–1765.

Honig, B., & Samuelsson, M. (2014). Data replication and extension: A study of business planning and venture-level performance. *Journal of Business Venturing Insights, 1*, 18–25.

Hopp, C., & Sonderegger, R. (2014). Understanding the dynamics of nascent entrepreneurship—Pre-startup experience, intentions, and entrepreneurial success. *Journal of Small Business Management*. doi:10.1111/jsbm.12107.

Katz, J., & Gartner, W. B. (1988). Properties of emerging organizations. *Academy of Management Review, 13*(3), 429–441.

Klein, P. G. (2008). Opportunity discovery, entrepreneurial action, and economic organization. *Strategic Entrepreneurship Journal, 2*(3), 175–190.

Kreiser, P. M., Marino, L. D., & Weaver, K. M. (2002). Assessing the psychometric properties of the entrepreneurial orientation scale: A multi-country analysis. *Entrepreneurship: Theory and Practice, 26*(4), 71–94.

Liao, J., & Gartner, W. B. (2006). The effects of pre-venture plan timing and perceived environmental uncertainty on the persistence of emerging firms. *Small Business Economics, 27*(1), 23–40.

Liao, J., & Welsch, H. (2008). Patterns of venture gestation process: Exploring the differences between tech and non-tech nascent entrepreneurs. *Journal of High Technology Management Research, 19*(2), 103–113.

Liao, J., Welsch, H., & Tan, W. L. (2005). Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns. *Journal of High Technology Management Research, 16*(1), 1–22.

Lichtenstein, B. B., Carter, N. M., Dooley, K. J., & Gartner, W. B. (2007). Complexity dynamics of nascent entrepreneurship. *Journal of Business Venturing, 22*(2), 236–261.

Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives, 7*(4), 199–204.

Locker, D., Jokovic, A., & Allison, P. (2007). Direction of wording and responses to items in oral health-related quality of life questionnaires for children and their parents. *Community Dentistry and Oral Epidemiology, 35*(4), 255–262.

Lohrke, F. T., Holloway, B. B., & Woolley, T. W. (2010). Conjoint analysis in entrepreneurship research: A review and research agenda. *Organizational Research Methods, 13*(1), 16–30.

Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. *Academy of Management Review, 21*(1), 135–172.

Lynn, R. (1991). *The secret of the miracle economy. Different national attitudes to competitiveness and money*. London: The Social Affairs Unit.

McClelland, D. C. (1961). *The achieving society*. Princeton, NJ: D. Van Nostrand and Co., Inc.

McMullen, J. S., & Dimov, D. (2013). Time and the entrepreneurial journey: The problems and promise of studying entrepreneurship as a process. *Journal of Management Studies, 50*(8), 1481–1512.

McMullen, J. S., & Shepherd, D. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review, 31*(1), 132–152.

Miller, D., & Friesen, P. H. (1978). Archetypes of strategy formulation. *Management Science, 24*(9), 921–933.

Miller, D., & Friesen, P. H. (1982). Innovation in conservative and entrepreneurial firms: Two models of strategic momentum. *Strategic Management Journal, 3*, 1–25.

Nunnally, J. C. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Perry, J. T., Chandler, G. N., & Markova, G. (2012). Entrepreneurial effectuation: A review and suggestions for future research. *Entrepreneurship: Theory and Practice, 36*(4), 837–861.

Rauch, A., Wiklund, J. L., Lumpkin, G. T., & Frese, M. (2009). Entrepreneurial orientation and business performance: An assessment of past research and suggestions for the future. *Entrepreneurship: Theory and Practice, 33*(3), 761–787.

Reynolds, P. D. (2007). New firm creation in the United States: A PSED I overview. *Foundations and Trends in Entrepreneurship, 3*(1), 1–150.

Reynolds, P. D., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies, 28*(4), 443–456.

Ringle, C. M., Sarstedt, M., & Straub, D. (2012). A critical look at the use of PLS-SEM in MIS Quarterly. *MIS Quarterly, 36*(1), iii–xiv.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 1–16). San Diego, CA: Academic Press.

Rosenbusch, N., Rauch, A., & Bausch, A. (2013). The mediating role of Entrepreneurial Orientation in the task environment–performance relationship: A meta-analysis. *Journal of Management, 39*(3), 633–659.

Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19*(4), 305–335.

Rossiter, J. R. (2010). *Measurement for the social sciences: The C-OAR-SE method and why it must replace psychometrics*. New York, NY: Springer.

Rossiter, J. R. (2011). Marketing measurement revolution: The C-OAR-SE method and why it must replace psychometrics. *European Journal of Marketing, 45*(11/12), 1561–1588.

Runyan, R. C., Ge, B., Dong, B., & Swinney, J. L. (2012). Entrepreneurial orientation in cross-cultural research: Assessing measurement invariance in the construct. *Entrepreneurship: Theory and Practice, 36*(4), 819–836.

Samuelsson, M. (2004). *Creating new ventures: A longitudinal investigation of the Nascent venturing process*. Doctoral dissertation, Jönköping International Business School, Jönköping

Samuelsson, M., & Davidsson, P. (2009). Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics, 33*(2), 229–255.

Sarasvathy, S. D. (2001). Causation and effectuation: Towards a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–288.

Sarasvathy, S. D. (2008). *Effectuation: Elements of entrepreneurial expertise*. Cheltenham, UK: Edward Elgar Publishing.

Schaffer, B. S., & Riordan, C. M. (2003). A review of cross-cultural methodologies for organizational research: A best-practices approach. *Organizational Research Methods, 6*, 169–215.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, UK: MA: Harvard University Press.

Senyard, J. M. (2015). *Bricolage and early stage firm performance*. Doctoral dissertation, Queensland University of Technology, Brisbane.

Senyard, J., Baker, T., Steffens, P., & Davidsson, P. (2014). Bricolage as a path to innovativeness for resource-constrained new firms. *Journal of Product Innovation Management, 31*(2), 211–230.

Shane, S. A., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review, 25*(1), 217–226.

Shepherd, D. A., & Wiklund, J. (2009). Are we comparing apples with apples or apples with oranges? Appropriateness of knowledge accumulation across growth studies. *Entrepreneurship: Theory and Practice, 33*(1), 105–123.

Slavec, A., & Drnovšek, M. (2012). A perspective on scale development in entrepreneurship research. *Economic and Business Review, 14*(1), 39–62.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103*, 667–680.

Stewart, T. J., & Frye, A. W. (2004). Investigating the use of negatively phrased survey items in medical education settings: Common wisdom or common mistake? *Academic Medicine, 79*(10), s18–s20.

Suppes, P., & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1, pp. 3–76). New York, NY: John Wiley & Sons, Inc.

Uy, M. A., Foo, M. D., & Aguinis, H. (2010). Using experience sampling methodology to advance entrepreneurship theory and research. *Organizational Research Methods, 13*(1), 31.

Venkataraman, S., Sarasvathy, S. D., Dew, N., & Forster, W. R. (2012). Reflections on the 2010 AMR Decade Award: Whither the promise? Moving forward with entrepreneurship as a science of the artificial. *Academy of Management Review, 37*(1), 21–33.

Wong, N., Rindfleisch, A., & Burroughs, J. E. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of the material values scale. *Journal of Consumer Research, 30*(1), 72–91.

Xavier, S., Kelley, D., Kew, J., Herrington, M., & Vorderwülbecke, A. (2013). *Global Entrepreneurship Monitor (GEM) 2012 global report*. London, UK: Global Entrepreneurship Research Association.

Zahra, S. A., & Wright, M. (2011). Entrepreneurship's next act. *Academy of Management Perspectives, 25*(4), 67–83.

The Dependent Variable

7

Abstract

What is it that we are trying to explain? Nothing could be more important in research than clarity about the explanandum, the "dependent variable." This chapter discusses explananda at different stages of the entrepreneurial process and at different levels of analysis. The spectrum stretches from an individual's *intention* to engage in entrepreneurship to *success* at establishing a new venture in the market to the societal level *impact* of entrepreneurial endeavors. An important argument is that intermediate explananda like *continuation* (rather than termination) of the process or evidence of making *progress* in it should be regarded as theoretical constructs in their own right rather than as proxy operationalizations of *success*.

7.1 Levels and Stages

Arguably, nothing is more important to a field of research than clarity about the conceptualizations and operationalizations of its explananda, i.e., the phenomena that we are trying to explain, often expressed as our dependent variables (DVs). And yet, we err a lot on precisely that issue, leading to confused and impeded knowledge accumulation. As usual, I am not without guilt myself. Let's try to do better in the future.

In strategy, there is arguably *one*, main, dependent variable: firm performance. Yet, not even there do you find anything near acceptable evidence of conceptual clarity and validity of the operationalizations (Butler, Martin, Perryman, & Upson, 2012; C. Miller, Washburn, & Glick, 2013). Actually, it is quite messy—so much so that Miller et al. bluntly suggest that "current practices must be stopped" (2013, p. 959). In entrepreneurship, we have many DV candidates, implying an even greater challenge.

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_7

Table 7.1 Some key dependent variables in entrepreneurship research

Level → / ↓ *Time*	Individual	Venture	Aggregate (industry; region, cluster, nation)
	Intention		Intention rate
	NVI identification		
	Engagement	Initiation	Initiation rate
	Persistence	Continuation	Conversion rate
		Progress	
		Duration	Duration
	Success	Emergence success	Entry rate
		Business performance	
	Impact	Impact	Impact

The mistake researchers frequently make is to take some not very carefully thought through DV that is easily available; that has been used in published research before, or which has shown "promising" correlations in a shotgun, initial exploration of the data at hand. None of this demonstrates that the measure is a valid indicator of the relevant theoretical construct. Moreover, we too easily interpret a high value on the scale as "good" and a low value as "bad/less good." For example, in research on firm growth—a borderland between strategy and entrepreneurship (Davidsson, 2005; Davidsson et al., 2002)—growth is often uncritically accepted as "success" although there is reason to be vary about such a generalization (Davidsson, Steffens, & Fitzsimmons, 2007, 2009; Delmar, McKelvie, & Wennberg, 2013). Further, researchers frequently lump together or select rather carelessly among indicators of different forms of growth, which are likely to have different drivers and consequences (Davidsson, Achtenhagen, & Naldi, 2010; McKelvie & Wiklund, 2010; Shepherd & Wiklund, 2009). Similarly, research on nascent entrepreneurship has too often applied mere continuation of the process as success indicator. For example, virtually all published results indicating a positive effect of business planning uses this type of DV. When stricter performance criteria are employed, no positive effects are demonstrated (Davidsson & Gordon, 2012; Garonne, 2014). We ought to do better.

I have argued earlier in this book that the creation of new economic activities, alternatively phrased as the emergence of new ventures, constitutes the core of entrepreneurship. I have further argued that this is a process that has outcomes on multiple levels and which can also be studied on several levels of analysis. I have further said that comparing current "entrepreneurs" (founder-owners) with others confounds the tendencies to engage, persist, and succeed in entrepreneurial endeavors. Where does this all leave us in terms of conceptualization and operationalization of important dependent variables for entrepreneurship research? In Table 7.1 and below, I share some ideas on these important issues.

The first thing to realize about the entries in Table 7.1 is that none of them are interchangeable alternatives that can be thrown in as DV in an analysis to test a pre-developed theoretical argument. Each of the entries captures a *conceptually distinct* aspect of entrepreneurship and is thus likely to have in part its own unique set of antecedents and consequences. Hence, the theory-testing researcher's task is to select their target aspect of entrepreneurship to explain, argue why this aspect is important to understand, and theorize its particular antecedents. Rather than seeking antecedents or effects of a vague and undifferentiated notion of "entrepreneurship," we need to carefully choose a more precise concept and a matching operationalization, putting the hypothesized relationships to an adequate test. Doing this for all the entries in Table 7.1 is clearly a task much bigger than this chapter and this author. Below I will briefly elaborate on the conceptual meaning of the suggested DV alternatives and point to operationalizations of at least some of them. I will treat the independent, commercial start-up as the norm; however, minor variations in definitions and operationalizations would make the suggestions work for internal and social ventures as well.

7.2 Dependent Variables on the Individual Level

Intention. Entrepreneurial intention refers to an individual's felt and/or stated desire or plan to engage in the creation of new economic activities. Intention and similar notions like aspirations and career interest have been the focus of quite a number of entrepreneurship studies. I have done it myself in what I boast as "my best-cited, unpublished study" (Davidsson, 1995; it was never submitted to a journal; long story, different times). However, intention is by no means a current favorite with editors and reviewers of high-tier journals. Far too often, the authors fail to explain why we ought to find out more about intentions rather than something further down the track toward creating new ventures. In short, it is a DV that is often used due to ease of access rather than theoretical or practical relevance. You can't fool us on that one!

This said, there has been some progress in this stream in recent years in terms of generally improved quality (e.g., L. F., Lee, P. K. Wong, Foo, & Leung, 2011), validation of measures (Liñán & Chen, 2009; Thompson, 2009), and increasing precision regarding the time frame and the type of entrepreneurial activity the intention concerns, including internal venturing (e.g., Douglas, 2013; Douglas & Fitzsimmons, 2013). Future research ought to build on this in taking important next steps to be discussed further below. It may also be worth pointing out that current founder-owners should *not* routinely be excluded from the analysis as they may well intend to start additional ventures.

Intention as discussed here is different from the individual's ability to identify new venture ideas in particular situations or their stated willingness to act on particular ideas for new ventures (Dimov, 2007). The latter two come to the fore if the individual-venture dyad is used as the level of analysis. Such combinations of a particular individual and a particular, emerging venture could have been given its own treatment in Table 7.1 and in this elaboration. However, readers are probably able to work out for themselves many of the adaptations needed when individual- and venture-level perspectives are combined in this way. Further, this dyadic perspective is central to Shane and Venkataraman's (2000) notion of the *individual-opportunity nexus*, which we will deal with at length in the next chapter.

I discussed the notion of new venture idea (NVI) in Chap. 2, and we will return to it in the next chapter, so I will try to keep it brief here. In a particular process *new venture idea identification* (or "opportunity recognition" if you must) may precede, coincide with, or follow after an intention is formed. As an individual-level DV—an enduring characteristic of the individual—NVI identification is about an individual's *ability* to creatively imagine and articulate new potential ventures in response to informational cues. To avoid confounding with the individuals' prior knowledge or interest in specific domains, a test of this ability needs to be repeated across a range of contexts. Further, I would strongly argue that operationalizations of NVI identification should *not* be mixed up with the same individuals' intention to act on those ideas, which brings in a whole new set of antecedents. If it is the *ability* to identify NVIs that is of theoretical interest, the operationalization should capture just that. Thus, the operationalization should be directed at what McMullen and Shepherd (2006) call "3rd person opportunities."

Engagement (opposite: nonengagement) denotes an individual's active, behavioral involvement as a founder-owner in a new venture start-up attempt. The nascent entrepreneur sampling mechanism discussed in Chap. 5 is arguably the best developed, existing operationalization of *current* engagement. With slight variations, the definition and operationalization can be adapted to include, or focus exclusively on, variants like internal venturing or socially oriented start-ups. When aggregating across an individual's career, a simple count of the number of start-up attempts engaged in—possibly normalized by age or measured as deviation (e.g., regression residual) from a predicted norm value—might be a suitable measure. A biasing complication that makes the count not quite so simple is that the individual can choose to launch new economic activities as new, independent start-ups or within firms that they are already running (cf. Chap. 5, Fig. 5.3).

Engagement is clearly not the same as success, nor is it simply the natural extension of intentions. It appears quite common that individuals "drift" into a business start-up attempt without this being preceded by a strong intention of doing so having developed well in advance (Bhave, 1994; Sarasvathy, 2001). Further, from Chap. 3, we learn that *perceived behavioral control* is likely to play a role. To demonstrate the role and relevance of intentions, future research should do more to explain variation in the strength of the relationship between intention and engagement.

Persistence (opposite: disengagement) is an individual's continued active, behavioral involvement as a founder-owner in a new venture start-up attempt. Here, we see a first important difference between the individual and venture levels: a team start-up can continue while the focal individual jumps ship. With this in mind, assessing persistence should be rather straightforward, although I would strongly recommend basing it on specific, researcher-defined minimum criteria (e.g., hours invested during the last X months) rather than relying on the individual's own perception of continued involvement. The ease of assessment also makes it a deceptively convenient variable to pick. The problem is rather with its theoretical relevance and meaning. Conceptually, persistence is rather far removed from any meaningful notion of success and does not even guarantee a lot of progress (see further below). Thus, persistence should not be used as a proxy for either of these. Further, persistence is theoretically complex as it might reflect either the supposedly positive quality of tenacity (Baum & Locke, 2004) or unsound escalation of commitment to a venture attempt that should be abandoned (DeTienne, Shepherd, & De Castro, 2008) or a bit of both. This makes it hard to theorize the antecedents and effects of persistence unless its different forms can somehow be distinguished both theoretically and empirically.

It should be noted that a different notion of persistence has recently been applied in research taking a career perspective. However, what Patel and Thatcher's (2014) interesting study captures is persistence in *self-employment*, which is not a good indicator of entrepreneurship (Henrekson & Sanandaji, 2014). An individual's career persistence in *entrepreneurship* would rather be reflected in the tendency to repeatedly engage in business start-ups, making them *habitual entrepreneurs* (Ucbasaran, Westhead, & Wright, 2006). A complication here is that repeated attempts are often triggered by prior failure. This makes it difficult to interpret persistence as unambiguously positive because the tendency to reach a high persistence score is sometimes contingent on the prior propensity to fail. For this reason, research on "serial entrepreneurs" may be "contaminated" by a subgroup who belong to that category not because of entrepreneurial expertise but because they fail once and tend to continue to do so (cf. Jenkins, 2012).

Success. The notion of individual-level success demonstrates even more clearly the crucial difference between individual and venture levels of analysis. First, a serious interest in the individual level suggests that (relative) financial performance is not necessarily the key criterion and that subjective outcomes like goal achievement and satisfaction are highly relevant (van Gelderen, Van der Sluis, & Jansen, 2005; Venkataraman, 1997). Second, if a founder manages to make the greater fool (Fearless & Clueless Investors, Inc.) pay millions or billions for an unproven venture (Superhype.com) that eventually fails, then the individual has enjoyed quite spectacular financial success. Hence, there is every reason to keep apart the *venture's* ability to generate a surplus from its operations on the one hand and the founders' economic gain from the venture on the other. A further reason for distinguishing the levels is the career perspective. The fact that an individual makes a painful loss on a single venture does not preclude the possibility that over their career they are involved in a large number or high proportion of successful venture start-ups and/or manage to amass a significant fortune from their entrepreneurial activities[1] (cf. Sarasvathy, 2004; Sarasvathy, Menon, & Kuechle, 2013). For such

[1] Note that the sentence you just read suggested three possible operationalizations of individual-level success from a career perspective: number of successful start-ups, proportion of such, and accumulated personal wealth from these start-ups.

reasons, I stated in the first edition of this book that "If we are to successfully explain or predict entrepreneurship with (distal) variables on the individual level, then entrepreneurship has to be broadly operationalized and/or assessed over a longer period of time." We can replace "entrepreneurship" in that statement with several of the individual-level DVs discussed above.

Impact. My inclusion of the cross-level notion of societal impact as DV reflects the observation in Chap. 1 that there is an important distinction between micro- and aggregate level outcomes. This is arguably a relatively more important issue for other levels included in Table 7.1, so suffice it here to say that societal impact is always relevant and important—for me it is the reason why I have an interest in entrepreneurship in the first place. The external impact of individual ventures is also notoriously difficult to trace. This does not mean we should not try as best we can. A question on the individual (career) level, though, is whether we can approximate an individual's entrepreneurial contribution with the fortune they have amassed (Henrekson & Sanandaji, 2014)? Perhaps. However, as per Fig. 1.3, I cannot easily shake off an uneasy discomfort with effectively equating value appropriation with value creation.

7.3 Dependent Variables on the Venture Level

Initiation of a new venture start-up attempt means that the first concrete action has been taken toward the realization of a new venture idea. Initiation thus largely coincides with an individual's (or the first team member's) *engagement* and can be assessed in the same manner. That is, the case should meet the minimum eligibility criteria for being a nascent venture, which include that some concrete action toward realization must have been taken within the past 12 months. This said, a venture-level perspective may require greater temporal precision in order to make venture creation processes comparable. This is something researchers have struggled with. In PSED and similar studies, there is timing information for a large number of "gestation activities" (cf. Chap. 6). However, it turns out that simply taking the first of these does not really work, because founders may undertake one gestation activity and then do nothing for 5 or 10 years, only to come back and eventually pursue the start-up with greater intensity. A practical solution that has been suggested is to identify the first 12-month period within which at least two activities were undertaken and use the first of these as the initiation marker (Schoonhoven, Burton, & Reynolds, 2009; cf. also Chap. 5). Depending on the exact nature of the activity data, the time period and activity level required could be modified. What we know does *not* work is to pick one particular activity as the marker of initiation. Start-up processes proceed in all manner of sequence, and what we might think should happen early and late may turn out to do the exact opposite (Gordon, 2012; Liao, Welsch, & Tan, 2005).

To increase theoretical precision in the assessment of initiation, researchers may want to consider combining evidence of activity with evidence of intentionality (Katz & Gartner, 1988). As noted above, individuals may sometimes do things—e.g., develop skills and buy equipment necessary to repair violins in Bhave's (1994) example—although at the time this was not with an intention to start a business. This may also explain some of the long lags referred to above. I have argued elsewhere that a venture creation process requires the existence of a new venture idea (Davidsson, 2015). If so, initiation does not occur until gestation activities are undertaken with the explicit purpose of trying to realize an imagined new venture.

Initiation is in some cases—but far from always—triggered by a clearly identifiable External Enabler (or objective, pre-existing, actor-independent "opportunity" if you must) such as a new technology or a regulatory change. We will have more to say about NVIs and External Enablers in the next chapter (cf. Davidsson, 2015).

Continuation (opposite: discontinuation, termination) on the venture level is similar to individual-level persistence although it suffices that one member on a team persists in order for the venture to continue. In fact, it is conceivable—albeit highly unlikely at this early stage—that the entire team is shifted while the venture continues. Further, there is no venture level equivalent to persistence as an individual career notion across ventures. For these reasons, it is important to distinguish between individual persistence and venture continuation.

This said, the two have some pros and cons in common: ease of assessment combined with theoretical difficulty of interpretation. Is continuation a good or a bad thing? It is also worth reminding here that continuation is often used in terms of its opposite—*termination*—and that we are referring to continuation of a start-up *attempt*. Continuation after emergence success (see below) is a different matter. The reasons for continuing a start-up attempt may be different from the determinants of survival of established firms.

Progress is the completion of necessary and facilitating actions aimed at reducing the gap between the venture's current state on the one hand and emergence success (see below) on the other. It may consist of and be assessed as the attainment of separate, necessary, or near-necessary milestones or gestation activities, like completing a marketable product; fulfilling regulatory requirements, and having first sales. Alternatively, it may be conceived and conceptualized as the accumulation of necessary and supplementary steps such as creating a website, retaining an accountant, and applying for funding and/or intellectual property protection. The latter makes it a (more) continuous variable and also allows for calculating a *rate* of progress per time unit (Lichtenstein, Carter, Dooley, & Gartner, 2007). Researchers have been using indicators of progress to a considerable extent, and some try both types of operationalization (Tornikoski & Newbert, 2007). However, too often indicators of progress are used to test theorized drivers of success or interpreted as such. I would argue that if progress is the operationalization, then it should also be what the theoretical argument tries to explain. It is possible for a while to rapidly progress a start-up that never becomes an operational business.

Duration is conceptually easy to define as the time from initiation to an outcome criterion, most likely either emergence success (see below) or termination (i.e., non-continuation) or both. Both? Well, some theoretical predictions may concern *time to resolution* rather than the likelihood of the outcome being either "positive" or "negative." For example, I have come across—but not seen verified—the argument that business planning should help *either* reaching an operational stage sooner *or* realizing more quickly that the start-up attempt is a lost cause and should be terminated.

Empirically, duration may not be all that simple because some venture creation processes seem to go on forever (Reynolds, 2007). Moreover, duration is rarely the researcher's main theoretical interest. Rather, it enters the picture via various "back doors." One of these is the application of event history models (Blossfeld, Hamerle, & Mayer, 2014), which are inherently time based. The most important role for duration may be as a second, "shadow" DV to facilitate interpretation of results pertaining to more final outcomes. Specifically, more ambitious and more innovative ventures may take longer to realize (Samuelsson & Davidsson, 2009). Assessment of continuation and success *at a given point in time* may therefore confound duration with the ultimate outcome for the venture. This said, there is a growing theoretical interest in issues of time and timing and as a case in point I have just recently been involved in work using duration (i.e., time to termination or progression decisions) as the dependent variable (Bakker & Shepherd, 2015; Bakker, Shepherd, & Davidsson, 2014).

Emergence success denotes the transition from being merely a continuing and progressing nascent venture to establishing it as a continuous and sustainable market participant. This is arguably the most important venture-level DV and one of the most important for entrepreneurship in general. This is what entrepreneurship research can uniquely contribute, which other lines of research have not done very well: explanations for new economic activities coming into being.

Emergence success is also one of the harder DVs to precisely define and operationalize. The original PSED project relied on the respondent's perception that the venture was now "up and running," which is hardly satisfactory. In PSED II and CAUSEE, we therefore tried to apply researcher-controlled criteria. This is the key question sequence:

A22. Has this business [NAME] received any sales revenue, income, or fees for more than 6 of the past 12 months? Y/N

A24. (If A22 = Yes) Has your monthly revenue been more than monthly expenses for more than 6 of the past 12 months? Y/N

A26. (If A24 = Yes) Are salaries for the owners who were active in managing the business included in the computation of monthly expenses? Y/N

The originally intended PSED II criterion for having become an operational firm—what I call *emergence success*—is a "yes" to all three questions. Essentially, this captures sustained positive cash flow. However, the researcher can choose to only require "yes" to A22 or A22 + A24. The reasons for considering a more lenient criterion are (1) the respondents' ability to answer accurately and consistently—i.e., measurement validity—is increasingly in doubt as we move from A22 to A26, and

(2) it takes longer for a substantial share of the sample to fulfill the stricter criteria; the positive cash flow criterion would have excluded even a giant like Amazon for a very long time (Spector, 2002). Whether we like to label it "success" or not, the A22 criterion indicates that the venture is now a regular market participant, which can be justifiably argued as being qualitatively different from being a nascent venture.

On the other hand, relying on too lenient a criterion is risky. For example, Diochon, Menzies, and Gasse (2007) found that many ventures which self-reported as "up and running" did not confirm that status in the next round of interviews. Further, not even positive cash flow guarantees the label "success" is warranted, especially if the initial investments were high. Conceptually, I would argue that "emergence success" on the venture level has been demonstrated when the venture has established itself relatively solidly in the market and generated a surplus from its operations, which is large enough to cover all start-up costs. Based on this notion, we tried this little operationalization invention in the CAUSEE study (asked as additional criteria only if A24 = Yes):

L17. Let's assume I posed the following question to your accountant or some other person with good insights into the history and financials of this business. *Question*: "As of today, would it be possible to sell or walk away from this business, and it would have covered all the costs incurred for developing it to what it currently is?" What do you think that knowledgeable person would answer—"Yes, definitely," "Yes, probably," "No, probably not," or "No, definitely not"?

L18. (L17 = No) After you started trading in the market on a regular basis, has there been any point in time when that knowledgeable person would have said it would be possible to sell or walk away from the business and it would have covered all the costs incurred for developing it up until that point in time—"Yes, definitely," "Yes, probably," "No, probably not," or "No, definitely not"?

L19. (Follow up to L17 or L18) Would that "Yes" include reasonable remuneration to the owners for their work, at least similar to the salary they could have earned doing some work as an employee?

The success of these indicators is still to be determined. One issue is that we did not think of including these questions until the third wave of data collection. Further, as the alert reader would already have picked up, when I designed these questions, I had not yet fully worked out the distinction between individual and venture-level success criteria. I would now have phrased the question in terms of revenue from operations. In its current form, the question may better capture success for the individual-venture dyad.

Another attempt at operationalizing emergence success is the work I did with Magnus Klofsten around his notion of a *business platform* (Davidsson & Klofsten, 2003). This multidimensional assessment of the status of the idea, the product, the market, the team, the emerging organization, and the venture's relational capital may better distinguish Gartner's (1988) idea of creation of *new organizations* from successful cases of establishing oneself as a self-employed economic actor who is never going to take on employees. On the downside, there will always be issues about which criteria should be included, how to combine them, and exactly where to put the bar for having achieved emergence success. It should also be noted that Klofsten's instrument was conceived and developed for technology-based ventures. Adaptations of criteria and their assessment may be needed for applications in other areas.

I would argue that *business performance* is outside of the immediate domain of entrepreneurship research. The performance of ventures after they have been successfully established is not in the main a matter of *creation* of new economic activities, and its determinants will increasingly be found among circumstances arising after emergence success has been achieved. I therefore gladly leave it to other fields—strategy in particular—to worry about the definition and operationalization of business performance. However, I acknowledge that some entrepreneurship researchers may find it unsatisfactory to stop at emergence success. First, many entrepreneurship researchers reside in strategy departments (Baker & Pollock, 2007) and thus prefer to operate in both domains. Second, there is definitely an important difference between the emergence success of Twitter and Instagram compared to that of a new, local hairdresser, and part of that difference can certainly be traced to circumstances occurring before emergence success. Third, we are still at the *venture* level rather than the firm level, suggesting that the fields of strategy, management, and organization may not be able to provide the right tools, especially if the venture is but a small part of an existing firm (instead, innovation research might provide tools for that level). Finally, venture-level performance indicators like levels of sales, profit, and employment may serve as crude indicators of societal level impact.

Impact. It is no secret by now that I find new ventures' impact on the economy or society at large to be of utmost interest and importance (cf. Chap. 2) and difficult to research. Indicators of the size and growth of the venture imply the creation of employment and tax revenue but without consideration of what the venture re- or displaces in the process. Studies of spawning (Gompers, Lerner, & Scharfstein, 2005) directly address effects outside the venture's own immediate boundaries. Comprehensive case studies would otherwise seem to be one of the few ways one can get the broader impact of individual ventures (e.g., Müller, 2013). Although well-executed case studies are valuable for theory development, creative ideas of how else to assess the extent to and mechanisms by which new ventures affect their industry or locality are welcome! I can guarantee I am not the only reviewer-editor who would be positively inclined toward such submissions. It is perhaps worth reiterating here that although we may be provided with analyses of the broader impact of well-known success ventures, there may also be important insights to be gained from resurrecting from oblivion some unsung heroes in the category of "catalyst ventures" (see Fig. 1.3; Chap. 1) and trace their broader impact.

7.4 Dependent Variables on Aggregate Levels of Analysis

Before we go into specific, aggregate DVs, let's first note that the aggregates we usually think of are industries or spatial units (region, country) or a combination of these (cluster). However, other aggregates are possible, such as cohorts of people of the same age (e.g., generations) or sharing the "same" origin (e.g., university spin-offs vs. corporate spin-offs). Further, aggregate DVs are often expressed as *rates* (Aldrich & Wiedenmayer, 1993) so as to make absolute numbers comparable across unequally sized entities. In computing such rates, the researcher often has a choice between two denominators: the size of the relevant human population (e.g., number of working age inhabitants) and the size of the relevant business population (e.g., number of firms in industry X).

Although there may be exceptions, my vote would normally go to the human population. This is because using the size of the business population may lead to strange effects in our domain. For example, imagine two regions, each with 100,000 working age population. Region A currently has 100 firms while region B has 1000, i.e., a more small firm-oriented structure. Now imagine region A sees 100 start-ups in a year while region B has twice as many, 200. Which industry/region has the higher level of entrepreneurial activity? Using the size of the business population as denominator, region A is far more entrepreneurial, having one start-up for every existing firm, whereas B only has one for every five existing firms. For most conceivable purposes, I find this conclusion absurd, because clearly *people* in region B are twice as prone to start businesses.[2] For the same reason, the younger me was obnoxious enough to refuse to deliver data portrayed in this way for the harmonized analysis in Reynolds, Storey, and Westhead (1994).[3]

With that, let's turn to aggregate level DV alternatives. Country level data on *intention rate* are available from some sources. For example, Roy Thurik and collaborators have analyzed such measures (among others) from the Flash Eurobarometer in a series of papers (Grilo & Thurik, 2005, 2008; van der Zwan, Verheul, & Thurik, 2012). I collected intention data myself on the regional level in the *Culture and Entrepreneurship* studies. However, unless the interest is specifically in the current situation, in conversion rates (see below), or in explicating the vagaries of relatively small statistical effects (Chap. 9), most researchers would see little reason to use attitudinal data from samples when population data on actual start-ups are available.

The main interest would instead be in *entry rate* (Geroski, 1995). Statistics of this nature may be available from sources like the OECD, World Bank, Eurostat,

---

[2] The same example can be run for industries, although in that case using the business population rather than the size of the workforce could be defended for some purposes.

[3] I should also mention that in the first edition of this book, I had two entire chapters devoted to aggregate level analysis based on large, secondary data sets. Because there has been considerable development of such data sets and their analysis in recent years while I have not been doing much work of that kind lately, I decided to drop them and only include snippets here. Despite being somewhat dated and sometimes overly Davidsson particular, interested readers may find value in those chapters as well, and I am happy to provide them on request.

and national statistical agencies. What are the issues? First, do the entry statistics really reflect the entry of new economic activities? “New registrations” often include a substantial share of reclassification of ongoing activities due to changes in ownership, location, or main industry classification. This said, establishment data are typically less subject to this type of unwanted volatility than are firm data, and establishment data may well be the best approximation for “new activity.” There is also undercoverage. In some countries, this is mainly due to the existence of a large “informal sector.” In highly developed economies which also have good business statistics (which is not true for all countries), a major source of undercoverage is that new economic activities by existing firms (“corporate entrepreneurship”) are not captured unless a new legal entity is also formed. Second, are the data comparable? Here, comparisons across industries and regions within a country stand on safer ground than cross-country comparisons, which despite improvements remain problematic due to various differences in the procedure and quality of initial data capture. This said, data have historically been of higher quality in some industries (traditionally manufacturing). Researchers may also encounter the “problem” of data quality improving over time, leading to risk of statistical artifacts.

The overall advice would be: know your data! If you can get contacts inside the statistics provider, use them! Read whatever background information there is about the data. Learn about their limitations and potentials. Perhaps data from different sources can be combined and particularly questionable parts of the data trimmed? As someone who has worked deeply with large data sets, I like to see evidence that the researchers understand their data in detail and that they have done all they could to avoid being misled by artifacts of the data (Pe’er & Vertinsky, 2008, is a good example). I have found it very helpful to be able to link firm and establishment levels to correct and/or better understand the data. Linked employer-employee data sets (see Amaral, Baptista, & Lima, 2011; Campbell, 2005; Sørensen, 2007) open up additional possibilities to remove artifacts from the data as well as to use new types of aggregates for which to calculate rates.

The limitations of archival data are one of the reasons why seemingly more subjective and uncertain (due to possible sampling error) data, sometimes referring to incomplete cases of new venture creation, can actually be preferable for many purposes. In fact, the cross-country noncomparability of extant entry data was one of the main reasons why the Global Entrepreneurship Monitor (GEM) was created. Of course, we soon learned that we were not easily going to get the one, perfectly harmonized indicator of entrepreneurial activity from GEM, either. Many factors stood in the way: cultural- and language-based differences in the interpretation of the questions; unavoidable differences in the data collection procedures; the statistical uncertainty of rare events data from small samples, and real differences in why people go into self-employment as well how that relates to notions of “entrepreneurship.” Consequently, the results raised suspicion and bewilderment when developing countries were ranked as the most entrepreneurial (Reynolds, Bygrave, & Autio, 2004).

However, researchers quickly learned to make sense of the data (Wennekers, Stel, Thurik, & Reynolds, 2005), and GEM has responded well to the challenges.

With larger samples, more countries, more years, a richer set of entrepreneurship indicators, improved bases for comparison, and tighter governance, GEM now provides the knowledgeable user with a unique and high-quality data set (Álvarez, Urbano, & Amorós, 2014; Amorós, Bosma, & Levie, 2013; Bergmann, Mueller, & Schrettle, 2014). This said, the "modest majority" issue discussed in Chap. 5 must always be kept in mind, as well as the fact that data points for individual years are usually based on relatively small samples and therefore statistically uncertain.

GEM provides *initiation rates* through its estimates of the prevalence of "nascent entrepreneurs." This criterion is near identical to the *engagement* measure discussed on the individual level. GEM also provides a proxy for *entry rate* through its assessment of the prevalence of young, operational firms. Recently, GEM has added an indicator of employee engagement in entrepreneurial activities, providing a new basis for comparing corporate entrepreneurship across countries (Steffens, 2013).

Moreover, the existence of intention, initiation, and entry data for the same spatial entities (within a data set or by combining them) allows researchers to pursue questions concerning *conversion rates*. These convey differences in the propensity to make the transitions from intention to initiation and onward to entry (Bergmann & Stephan, 2013; Grilo & Thurik, 2005). This may lead to the identification of country-specific hurdles in the venture creation process and entry barriers vs. survival barriers across industries (cf. Geroski, 1995). Data on average process *duration* on the aggregate level could provide similar insights but are usually not available because the data does not capture the emergence process of individual cases over time.

A keen interest in *impact* seems to come more naturally to researchers addressing aggregate level issues, and although proving causality remains challenging, linking entrepreneurial activity to societal impact is somewhat more doable when both are assessed on the aggregate level. Insight and inspiration can be gained from previous work like van Praag and Versloot (2007) and Wennekers and Thurik (1999). Reviews of the evidence typically assign nuanced but predominantly positive effects of entrepreneurial activity on indicators like job creation, innovation, productivity, and economic growth. However, we still have considerable work to do in distinguishing between what Baumol (1990) calls productive, unproductive, and destructive expressions of entrepreneurship. Relatedly, the increased interest in sustainability calls for new types of evidence of the role of new ventures (Hall, Daneke, & Lenox, 2010).

7.5 Research on Job Creation

The clearest evidence on new ventures' impact on the economy at large concerns job creation, an area which I have dabbled in quite a bit myself (e.g., Davidsson & Delmar, 2006; Davidsson, Lindmark, & Olofsson, 1998). Comprehensive recent research confirms the conclusion we drew in the 1990s, namely that that the entry and early growth of new firms is a major force providing the economy with new jobs, whereas established SMEs as a group have negative net job creation (Criscuolo, Gal, & Menon, 2014; Haltiwanger, Jarmin, & Miranda, 2013). This does not mean that all is in perfect order in job creation research. In the first edition of this book, I devoted an entire chapter to issues in such research, and although this stream has improved, it may be worth reiterating some of the observations and advice from the conclusions of that chapter. Except for the bracketed comments, the remainder of this section is taken essentially verbatim from the first edition (then Chap. 8).

Regardless of the more or less noble reason for engaging in job creation research, I believe we should try to do a good job. This entails getting some of our basic bearings right. For example, policy-makers and researchers alike ought to realize that from the entrepreneur's perspective, the creation of new jobs is normally not a goal but perhaps a hesitantly accepted consequence of realizing ones real goals. Employees are costly, and reluctance to add personnel is therefore the norm also for dynamic and expansive entrepreneurs. As remarked on stage by a very colorful entrepreneur from the notoriously underemployed northern parts of Sweden, "Every time I have a new project the local politicians ask 'How many jobs is this going to create?' and I answer as usual 'As few as possible!'" Running a business is no charity.

What's a new job, anyway? It could mean a range of things. When an individual terminates her job with one employer and assumes a new post with another employer, that person feels she has a "new job." But this does not mean that her previous job ceases to exist or that her "new job" did not exist before, albeit performed by another individual. So from the firms' perspectives, no job losses or gains are necessarily involved in such a change. As we shall see, in job creation research, the "gross" number of jobs created or lost on some level of analysis (industry, region, economy at large) is usually the firm or establishment-level net change in the number of people employed. Such an analysis sees no difference between a firm that keeps up the same employment numbers by having exactly the same people do exactly what they did last year and another one that changes all its people and/or their work tasks but happens to end up with exactly the same number of people. This also means that if researchers aren't careful, there may be a far cry between the theoretical concept of "job" they use on the one hand and its operationalization on the other.

And perhaps it is a little narrow sighted to just count the numbers of jobs, assuming that more is always better in this regard. I mean, if more work is what we want, we can follow the old example of destroying the terrible machines that make us redundant and get rid of the division of labor that has made us efficient. We would certainly be kept busy if we had to start looking for some iron ore every time we needed a safety pin. Yeah, one should perhaps think a little about the quality of jobs, too, and the value of leisure.

Before we close the books on job creation research, there are a few more issues that deserve at least brief mention:

Choice of microlevel unit. Establishment vs. company vs. company group. In some data sets, it is not possible to identify firms (legal units), only establishments (plants, workplaces). This is a severe limitation and the implication is clear: inferences about job creation in different firm size classes (or other categories of firm) should not be made. So when a theory about firms and categories of firms guides the research, establishment-level data yield distorted results. More specifically, such data exaggerate gross job volatility relative to company level data because within-firm shift of employment between establishments is included in the establishment but not the company level analysis. This may, however, be a desirable feature when the theory is about industries or regions because when establishment rather than company data are aggregated to these levels, relatively more of the total gross changes are captured and assigned to the right entity (industry or region code).

Artificial changes. Ownership changes (including spin-offs and mergers), geographic moves, and change of legal form of the firm frequently lead to the registration of firm (and sometimes establishment) births and deaths when no genuinely new business activity has been set up or no old one has really ceased to exist. The effect is a general overestimation of gross job volatility in the economy, compared with a definition that assumes some "genuine" change. This is a problem that is more pronounced among small units compared to large ones, leading one to suspect general overestimation of job volatility among young and small units (and therefore within industries or regions dominated by such). However, it may well be those fewer cases where some unusual form of restructuring leads to the appearance of entirely new large units that generate the greatest—albeit detectable—errors in data for individual regions or industries in particular years. The occurrence of artificial changes may also be unevenly distributed across industries—consider, for example, the special characteristics of workplaces in the construction industry or the rather frequent ownership transfers that signify certain other industries. In addition, the occurrence of artificial changes may change over time, making the researcher curse the increasing quality of their data!

Length of analysis period. The length of the analysis period is critical and must be considered when, e.g., comparing results from different studies. If 2-year rather than 1-year periods are chosen, some firms may enter and exit within that period and therefore never be recorded. At least for firm-level analysis this will lead to underestimation of job volatility among young and small firms, because these short-lived entrants are young by definition and very rarely of large size. Possibly (but not necessarily), however, the use of longer analysis periods leads to relatively more regression fallacy problems [see first edition, Chap. 8]. The length of the analysis period also affects what part of job gains is attributed to births versus expansion, as a new firm's second-year expansion will be counted as "birth jobs" in one analysis and "expansion jobs" in the other.

Assigning size class to new entrants. If base-year size definition is used (i.e., the entities are assigned to the size class they belong to at the beginning of the analysis period), new entrants cause a special problem. These firms *do not have* a base-year size. For this as well as other reasons, job changes associated with births and deaths [entry and exit] should be separated from those associated with expansion and contraction among continuing firms. Again, new entrants are typically small. Mixing them with the smallest size class of established firms may lead to the image that "small firms grow a lot" when the truth is that established firms in that category have an aggregate decline in employment. In fact, failure to make this distinction should perhaps carry the blame for much policy and research interest being directed to small firms, when it really should have been directed to the creation of new ones, i.e., to the importance of entrepreneurship [cf. Audretsch, Grilo, & Thurik, 2007; Greene, Mole, & Storey, 2007; Lundström & Stevenson, 2005].

Do all the methodological problems discussed above mean that any research effort to gain insight into the mechanisms of job creation is inherently futile? I would say no; in research, we will always have to live with limitations and make assumptions and simplifications. If the alternative is to make unfounded guesses, less-than-ideal research always has a place. The important thing is that the researcher is aware of the limitations, does her best to handle them, and communicates her result with a level of confidence that accords with the quality of the data.

Do we need more research on job creation? In order to establish that small (and—in particular—new) firms have been overrepresented as job creators during the last few decades, we need no more research. Few findings in the domain of social science have as solid empirical support as that [well, maybe we needed more evidence to convert some diehards, confirm it still holds, and reinforce the distinction between smallness and newness. If so, we now have that evidence; see Criscuolo et al., 2014; Haltiwanger et al., 2013]. To follow the development over time should in the future rather be a task for statistical organizations [this has happened; a lot more longitudinal data sets exist]. There are, however, many other reasons why an entrepreneurship researcher should show an interest in job creation. When we do that in the future, the issues discussed above suggest that we should:

1. Make sure the data are longitudinal and of high enough quality to make the effort worthwhile in the first place.
2. Clarify to ourselves and readers what "new job" and job losses actually mean on the basis of the data at hand and that we use theory and make comparisons with other research in accordance with this notion of "new job" [in this type of research, it tends to mean net addition to the number of employees in an entity, not numbers of new positions or recruitments made and not necessarily employment for previously unemployed people].
3. Separate in the analysis the job changes that are attributable to births, deaths [of firms/ventures!], expansions, and contractions, respectively.
4. Likewise separate job changes attributable to organic changes from those resulting from mergers, acquisitions, and splits [just about to happen now in some countries].
5. Apply a size definition that comes as close as possible to momentary size [see first edition or de Wit and de Kok (2013) about this innovation of mine, now usually called "dynamic sizing"]. Base-year size is a defensible alternative as long as relatively short analysis periods are used and the number of size classes is relatively small. If the data permit, a correction like Davidsson et al.'s (1998) should supplement the analysis.
6. Express the results on job changes as gross and net absolute figures and shares and relate these to corresponding figures for the employment base. Rates are a more dubious matter when analyzing total job creation in the economy over a

number of 1-year analysis periods. Rates are something we want to compute for cohorts of firms, while categories like size classes, industries, and regions are genuinely moving targets in the sense that they continuously change their members. However, studies of growth rates for stable cohorts of firms in different initial size classes would certainly complement the other type of study and therefore add to our understanding.

7. Consider not only the numbers but also the quality of new jobs and use other outcome measures alongside with job creation. We shouldn't forget that we do not live to work and that when we do work it is nice if the work comes with both intrinsic and extrinsic rewards!

Doing perfect research is not possible, so it is unlikely that we can deal fully satisfactorily with all the above points at once. But thanks to forerunners and their shortcomings, it should now be possible to deal more satisfactorily with them than has been the case in earlier job creation research. If we can't, I seriously think we should not conduct the research at all. It would just be unnecessary work. More work does not always equal more well-being, neither economic nor other and neither on the individual nor the societal level. It applies to unnecessary research work as well.

7.6 Summary and Conclusion

This chapter has been a call on entrepreneurship researchers to take greater care in their conceptualization and operationalization or their explanandum, the dependent variable that the research is assumed to explain. Arguably, nothing could be more important. The nature of entrepreneurship as a process that can be addressed at different levels of analysis means that many alternative DVs are on offer. It is imperative that the chosen DV is properly theorized and that the empirical DV accurately matches the theoretical construct. For many purposes, microlevel *success* and aggregate level *impact* are ultimately the most important explananda. This said, in order to develop a correct understanding of these, we may have to dig deeper also into issues pertaining to *engagement/initiation*, *persistence/continuation*, *progress*, and *duration*. These are theoretical concepts in their own right. Their unique antecedents and consequences should be theorized and tested rather than using these intermediate variables as poor operationalizations of success and impact.

Job creation is an explanandum whose popularity waxes and wanes with the swings of the business cycle. Toward the end of the chapter, I pointed to new evidence of the importance of start-ups for job creation and reiterated some advice about doing work on that topic. The most important advice is to be clear about what a "new job" actually means in this type of research and to be aware of how various design choices will affect the attribution of new jobs to different categories of potential job creators.

References

Aldrich, H. E., & Wiedenmayer, G. (1993). From traits to rates: An ecological perspective on organizational foundings. In J. Katz & R. Brockhaus (Eds.), *Advances in entrepreneurship, firm emergence, and growth* (1st ed., pp. 145–196). Greenwich, CT: JAI Press.

Álvarez, C., Urbano, D., & Amorós, J. E. (2014). GEM research: Achievements and challenges. *Small Business Economics, 42*(3), 445–465.

Amaral, A. M., Baptista, R., & Lima, F. (2011). Serial entrepreneurship: Impact of human capital on time to re-entry. *Small Business Economics, 37*(1), 1–21.

Amorós, J. E., Bosma, N., & Levie, J. (2013). Ten years of global entrepreneurship monitor: Accomplishments and prospects. *International Journal of Entrepreneurial Venturing, 5*(2), 120–152.

Audretsch, D. B., Grilo, I., & Thurik, A. R. (2007). *Handbook of research on entrepreneurship policy*. Cheltenham, UK: Edward Elgar Publishing.

Baker, T., & Pollock, T. G. (2007). Making the marriage work: The benefits of strategy's takeover of entrepreneurship for strategic organization. *Strategic Organization, 5*(3), 297–312.

Bakker, R. M., Shepherd, D. A., & Davidsson, P. (2014). When to pull the plug and when to take the plunge: Timing strategic decisions about new ventures. In Humphreys, J. (Ed.) *Best Paper Proceedings: The Power of Words*. Academy of Management, PA.

Bakker, R. M., & Shepherd, D. A. (2015). Pull the plug or take the plunge: Multiple opportunities and the speed of venturing decisions in the Australian mining industry. Paper accepted for publication in *Academy of Management Journal*.

Baum, J. R., & Locke, E. A. (2004). The relationship of entrepreneurial traits, skill, and motivation to subsequent venture growth. *Journal of Applied Psychology, 89*(4), 587–598.

Baumol, W. J. (1990). Entrepreneurship: Productive, unproductive and destructive. *Journal of Political Economy, 98*(5), 893–921.

Bergmann, H., Mueller, S., & Schrettle, T. (2014). The use of Global Entrepreneurship Monitor data in academic research: A critical inventory and future potentials. *International Journal of Entrepreneurial Venturing, 6*(3), 242–276.

Bergmann, H., & Stephan, U. (2013). Moving on from nascent entrepreneurship: Measuring cross-national differences in the transition to new business ownership. *Small Business Economics, 41*(4), 945–959.

Bhave, M. P. (1994). A process model of entrepreneurial venture creation. *Journal of Business Venturing, 9*, 223–242.

Blossfeld, H. P., Hamerle, A., & Mayer, K. U. (2014). *Event history analysis: statistical theory and application in the social sciences*. East Sussex, UK: Psychology Press.

Butler, F. C., Martin, J. A., Perryman, A. A., & Upson, J. W. (2012). Examining the dimensionality, reliability, and construct validity of firm financial performance. *Strategic Management Review, 6*(1), 57–74.

Campbell, B. A. (2005). *Using linked employer-employee data to study entrepreneurship issues* (Handbook of Entrepreneurship Research, pp. 143–166). Berlin: Springer.

Criscuolo, C., Gal, P. N., & Menon, C. (2014). *The dynamics of employment growth: New evidence from 18 countries*. Paris: OECD Publishing.

Davidsson, P. (1995). *Determinants of entrepreneurial intentions working paper 1995:1*. Jönköping: Jönköping International Business School. http://eprints.qut.edu.au/2076/1/RENT_IX.pdf.

Davidsson, P. (2005). Entrepreneurial growth. In M. A. Hitt & R. D. Ireland (Eds.), *Entrepreneurship* (2nd ed., Vol. III, pp. 80–82). Maden, MA: Basil Blackwell & Mott, Ltd.

Davidsson, P. (2015). Entrepreneurial opportunities and the entrepreneurship nexus: A reconceptualization. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.002.

Davidsson, P., Achtenhagen, L., & Naldi, L. (2010). Small firm growth. *Foundations and Trends in Entrepreneurship, 6*(2), 69–166.

Davidsson, P., & Delmar, F. (2006). High-growth firms and their contribution to employment: The case of Sweden 1987-96. In P. Davidsson, F. Delmar, & J. Wiklund (Eds.), *Entrepreneurship and the growth of firms*. Cheltenham, UK: Edward Elgar Publishing.

Davidsson, P., Delmar, F., & Wiklund, J. (2002). Entrepreneurship as growth; growth as entrepreneurship. In M. A. Hitt, R. D. Ireland, S. M. Camp, & D. L. Sexton (Eds.), *Strategic entrepreneurship: Creating a new mindset* (pp. 328–342). Oxford, UK: Basil Blackwell & Mott, Ltd.

Davidsson, P., & Gordon, S. R. (2012). Panel studies of new venture creation: A methods-focused review and suggestions for future research. *Small Business Economics, 39*(4), 853–876.

Davidsson, P., Kirchhoff, B., Hatemi-J, A., & Gustavsson, H. (2002). Empirical analysis of growth factors using Swedish data. *Journal of Small Business Management, 40*(4), 332–349.

Davidsson, P., & Klofsten, M. (2003). The business platform: Developing an instrument to gauge and assist the development of young firms. *Journal of Small Business Management, 41*(1), 1–26.

Davidsson, P., Lindmark, L., & Olofsson, C. (1998). The extent of overestimation of small firm job creation: An empirical examination of the 'regression bias'. *Small Business Economics, 10*, 87–100.

Davidsson, P., Steffens, P., & Fitzsimmons, J. (2007). *Performance assessment in entrepreneurship research: Is there a pro-growth bias*. Retrieved from http://eprints.qut.edu.au/archive/00012040.

Davidsson, P., Steffens, P., & Fitzsimmons, J. (2009). Growing profitable or growing from profits: Putting the horse in front of the cart? *Journal of Business Venturing, 24*(4), 388–406.

de Wit, G., & de Kok, J. (2013). Do small businesses create more jobs? New evidence for Europe. *Small Business Economics,* 1–13.

Delmar, F., McKelvie, A., & Wennberg, K. (2013). Untangling the relationships among growth, profitability and survival in new firms. *Technovation, 33*(8), 276–291.

DeTienne, D. R., Shepherd, D. A., & De Castro, J. O. (2008). The fallacy of "only the strong survive": The effects of extrinsic motivation on the persistence decisions for under-performing firms. *Journal of Business Venturing, 23*(5), 528–546.

Dimov, D. (2007). From opportunity insight to opportunity intention: The importance of person-situation learning match. *Entrepreneurship: Theory and Practice, 31*(4), 561–583.

Diochon, M., Menzies, T. V., & Gasse, Y. (2007). From becoming to being: Measuring firm creation. *Journal of Enterprising Culture, 15*, 21–42.

Douglas, E. J. (2013). Reconstructing entrepreneurial intentions to identify predisposition for growth. *Journal of Business Venturing, 28*, 633–651.

Douglas, E. J., & Fitzsimmons, J. R. (2013). Intrapreneurial intentions versus entrepreneurial intentions: Distinct constructs with different antecedents. *Small Business Economics, 41*(1), 115–132.

Garonne, C. (2014). *Business planning in emerging firm: Uses and effects.* Doctoral dissertation, Queensland University of Technology, Brisbane.

Gartner, W. B. (1988). "Who is an Entrepreneur?" is the wrong question. *American Small Business Journal, 12*(4), 11–31.

Geroski, P. A. (1995). What do we know about entry? *International Journal of Industrial Organization, 13*, 421–440.

Gompers, P., Lerner, J., & Scharfstein, D. (2005). Entrepreneurial spawning: Public corporations and the genesis of new ventures, 1986 to 1999. *Journal of Finance, 60*(2), 577–614.

Gordon, S. R. (2012). *Dimensions of the venture creation process: Amount, dynamics, and sequences of action in nascent entrepreneurship*. Doctoral dissertation, Queensland University of Technology, Brisbane.

Greene, F. J., Mole, K., & Storey, D. (2007). *Three decades of enterprise culture? Entrepreneurship, economic regeneration and public policy*. Basingstoke: Palgrave MacMillan.

Grilo, I., & Thurik, R. (2005). Latent and actual entrepreneurship in Europe and the US: Some recent developments. *The International Entrepreneurship and Management Journal, 1*(4), 441–459.

Grilo, I., & Thurik, R. (2008). Determinants of entrepreneurial engagement levels in Europe and the US. *Industrial and Corporate Change, 17*(6), 1113–1145.

Hall, J. K., Daneke, G. A., & Lenox, M. J. (2010). Sustainable development and entrepreneurship: Past contributions and future directions. *Journal of Business Venturing, 25*(5), 439–448.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who creates jobs? Small versus large versus young. *Review of Economics and Statistics, 95*(2), 347–361.

Henrekson, M., & Sanandaji, T. (2014). Small business activity does not measure entrepreneurship. *Proceedings of the National Academy of Sciences, 111*(5), 1760–1765.

Jenkins, A. (2012). *After firm failure: Emotions, grief, and re-entry*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Katz, J., & Gartner, W. B. (1988). Properties of emerging organizations. *Academy of Management Review, 13*(3), 429–441.

Lee, L., Wong, P. K., Foo, M. D., & Leung, A. (2011). Entrepreneurial intentions: The influence of organizational and individual factors. *Journal of Business Venturing, 26*(1), 124–136.

Liao, J., Welsch, H., & Tan, W. L. (2005). Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns. *Journal of High Technology Management Research, 16*(1), 1–22.

Lichtenstein, B. B., Carter, N. M., Dooley, K. J., & Gartner, W. B. (2007). Complexity dynamics of nascent entrepreneurship. *Journal of Business Venturing, 22*(2), 236–261.

Liñán, F., & Chen, Y. W. (2009). Development and cross-cultural application of a specific instrument to measure entrepreneurial intentions. *Entrepreneurship: Theory and Practice, 33*(3), 593–617.

Lundström, A., & Stevenson, L. (2005). *Entrepreneurship policy: Theory and practice*. Berlin: Springer.

McKelvie, A., & Wiklund, J. (2010). Advancing firm growth research: A focus on growth mode instead of growth rate. *Entrepreneurship: Theory and Practice, 34*(2), 261–288.

McMullen, J. S., & Shepherd, D. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review, 31*(1), 132–152.

Miller, C. C., Washburn, N. T., & Glick, W. H. (2013). The myth of firm performance. *Organization Science, 24*(3), 948–964.

Müller, S. (2013). *Entrepreneurship and regional development: On the interplay between agency and context.* Doctoral dissertation, Department of Business Administration, Aarhus University.

Patel, P. C., & Thatcher, S. M. (2014). Sticking it out: Individual attributes and persistence in self-employment. *Journal of Management, 40*(7), 1932–1979.

Pe'er, A., & Vertinsky, I. (2008). Firm exits as a determinant of new entry: Is there evidence of local creative destruction? *Journal of Business Venturing, 23*(3), 280–306.

Reynolds, P. D. (2007). New firm creation in the United States: A PSED I overview. *Foundations and Trends in Entrepreneurship, 3*(1), 1–150.

Reynolds, P. D., Bygrave, W. D., & Autio, E. (2004). *Global Entrepreneurship Monitor 2003. Executive report.* Babson College/Ewing Marion Kauffman Foundation, London Business School, London, UK, and Kansas, MO.

Reynolds, P. D., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies, 28*(4), 443–456.

Samuelsson, M., & Davidsson, P. (2009). Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics, 33*(2), 229–255.

Sarasvathy, S. D. (2001). Causation and effectuation: Towards a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–288.

Sarasvathy, S. D. (2004). The questions we ask and the questions we care about: Reformulating some problems in entrepreneurship research. *Journal of Business Venturing, 19*(5), 707–720.

Sarasvathy, S. D., Menon, A. R., & Kuechle, G. (2013). Failing firms and successful entrepreneurs: Serial entrepreneurship as a temporal portfolio. *Small Business Economics, 40*(2), 417–434.

Schoonhoven, C. B., Burton, M. D., & Reynolds, P. D. (2009). Reconceiving the gestation window: The consequences of competing definitions of firm conception and birth. In P. D. Reynolds

& R. T. Curtin (Eds.), *New firm creation in the United States* (pp. 219–237). New York, NY: Springer.

Shane, S. A., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review, 25*(1), 217–226.

Shepherd, D. A., & Wiklund, J. (2009). Are we comparing apples with apples or apples with oranges? Appropriateness of knowledge accumulation across growth studies. *Entrepreneurship: Theory and Practice, 33*(1), 105–123.

Sørensen, J. B. (2007). Bureaucracy and entrepreneurship: Workplace effects on entrepreneurial entry. *Administrative Science Quarterly, 52*(3), 387–412.

Spector, R. (2002). *Amazon.com: Get big fast*. New York, NY: HarperCollins.

Steffens, P. R. (2013). Culture as a driver of entrepreneurship: Contrasting independent entrepreneurship versus employee entrepreneurship. *Paper presented at the ACERE Conference*, Brisbane, Feb 5–8. Retrieved from http://eprints.qut.edu.au/59472/.

Thompson, E. R. (2009). Individual entrepreneurial intent: Construct clarification and development of an internationally reliable metric. *Entrepreneurship: Theory and Practice, 33*(3), 669–694.

Tornikoski, E. T., & Newbert, S. L. (2007). Exploring the determinants of organizational emergence: A legitimacy perspective. *Journal of Business Venturing, 22*(2), 311–335.

Ucbasaran, D., Westhead, P., & Wright, M. (2006). *Habitual entrepreneurs*. Cheltenham, UK: Edward Elgar Publishing.

van der Zwan, P., Verheul, I., & Thurik, A. R. (2012). The entrepreneurial ladder, gender, and regional development. *Small Business Economics, 39*(3), 627–643.

van Gelderen, M., Van der Sluis, L., & Jansen, P. (2005). Learning opportunities and learning behaviors of small business starters: Relations with goal achievement, skill development, and satisfaction. *Small Business Economics, 25*, 97–108.

van Praag, C. M., & Versloot, P. H. (2007). The economic benefits and costs of entrepreneurship: A review of the research. *Foundations and Trends in Entrepreneurship, 4*(2), 65–154.

Venkataraman, S. (1997). The distinctive domain of entrepreneurship research: An editor's perspective. In J. Katz & J. Brockhaus (Eds.), *Advances in entrepreneurship, firm emergence, and growth* (Vol. 3, pp. 119–138). Greenwich, CT: JAI Press.

Wennekers, S., Stel, A., Thurik, A. R., & Reynolds, P. D. (2005). Nascent entrepreneurship and the level of economic development. *Small Business Economics, 24*, 293–309.

Wennekers, S., & Thurik, R. (1999). Linking entrepreneurship and economic growth. *Small Business Economics, 13*(1), 27–55.

8 The Entrepreneurship Nexus

"[T]here is remarkable consensus on the definition of an opportunity"

(Alvarez, Barney, & Anderson, 2013, p. 302)

Abstract

What drives the progress and success of start-up processes? The "entrepreneurship nexus" perspective holds that entrepreneurial processes and their outcomes are shaped by the interplay of individuals and the "opportunities" they pursue. This chapter argues that progress within this perspective has been hampered by inescapable complexities inherent in any notion of "opportunity." To make further progress, the chapter argues that what prior research has been trying to capture in the "opportunity" construct needs to be studied under three separate theoretical notions: external enablers, new venture ideas, and individuals' opportunity confidence.

8.1 Killing a Darling

In their seminal "Promise" article, Shane and Venkataraman (2000) introduced the idea of the individual-opportunity nexus. This framework delineates the domain of entrepreneurship research as one focusing on entrepreneurial opportunities, the individuals who pursue them, and the interplay between these entities. A microlevel focus is evident from the statement that individuals and opportunities are the first-order forces explaining entrepreneurship; environmental forces cannot by themselves explain entrepreneurship because entrepreneurship requires agency (cf. Shane, 2003, p. 214). The authors also emphasize a process perspective; entrepreneurial undertakings evolve over time.

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_8

Opportunities are conceived of as situations offering profit potential. They objectively exist prior to their discovery and exploitation. However, their recognition is a subjective process; entrepreneurs act on conjectures (cf. Eckhardt & Shane, 2013, p. 220). Opportunities are entrepreneurial when they involve "new means-ends frameworks" rather than optimizing within existing frameworks. Rather than affording primacy to the individual entrepreneur, the nexus framework makes the opportunity an equal partner in the equation. Hence, we should not just ask "Why do individuals become entrepreneurs?" but also "Where do entrepreneurial opportunities come from?". In seeking microlevel explanations for entrepreneurial action and outcomes, we should not only look for the qualities of the individuals involved but also at the characteristics of the opportunities they pursue and the fit between opportunity and individual (Shane, 2000). In principle, the framework invites us to alternatively start at the opportunity end, hypothesizing action and outcome effects on the basis of generic "characteristics of opportunities" and only secondarily include characteristics of the people who discover, conceive, or create them as additional explanations and moderators of the opportunity effects.

First outlined by Venkataraman (1997), the framework has been elaborated, explained, defended, and partly reinterpreted by, e.g., Eckhardt and Shane (2003, 2010, 2013); Shane (2003, 2012) and Shane and Eckhardt (2003), and cited and commented on by hundreds of other scholars. In Chap. 2, I reiterated an extensive list of the virtues of this framework compared to what preceded it and showed further appreciation by incorporating some of it as core elements in my own delineation of the scholarly domain of entrepreneurship research. Arguably, one of the most important virtues is that Shane and Venkat's ideas have reduced the exaggerated focus on the individual (cf. Ross, 1977) that signified early entrepreneurship research (Gartner, 1988). Further—and perhaps even more importantly—it has helped to conceptually liberate entrepreneurship from its strong attachment to the particular context of small and/or owner-managed businesses. It has helped turning the focus instead toward covering the earliest stages of new venture development, from nonexistence to existence, across organizational contexts. Arguably, this focus sets up entrepreneurship research to make more valuable theoretical contributions to various disciplines and specializations within management and organizational research (cf. Low, 2001; Sorenson & Stuart, 2008; Wiklund, Davidsson, Audretsch, & Karlsson, 2011).

As a result, Shane and Venkataraman's framework—especially its focus on entrepreneurial opportunities[1]—has become widely accepted and attributed a dominant position in entrepreneurship research (Korsgaard, 2013, p. 3; Plummer, Haynie, & Godesiabois, 2007, p. 354; Shane, 2012, pp. 16–18; Venkataraman, Sarasvathy,

[1] Although "opportunities" had been discussed before in entrepreneurship (e.g., Gaglio, 1997) and management (e.g., Jackson & Dutton, 1988), most of the work has appeared after the publication of Shane and Venkataraman's paper. Short, Ketchen, Shook, and Ireland (2010) reviewed 68 papers in 16 leading journal, which used "opportunity" in relevant ways in title, abstract, or keywords. Only eight of these were published before 2000. A whopping 150 papers fulfilling the criteria were published in the same journals 2010–2014 (Davidsson, 2015).

Dew, & Forster, 2012). Their framework has later been recast as "Discovery Theory" and contrasted with "Creation Theory," which assumes that opportunities are not objective and pre-existing but contingent on entrepreneurs' perception and a process of social construction, thus maintaining a central position for the opportunity concept (Alvarez & Barney, 2007, 2010, 2013; Alvarez et al., 2013). This centrality of "opportunities" in entrepreneurship research is also manifested in Shane and Venkataraman (2000) receiving the AMR Decade Award and strongly influencing the revised Domain Statement of the Entrepreneurship Division of the Academy of Management (cf. http://aom.org/DIG/# and Mitchell, 2011).

I believe that the idea of an entrepreneurship nexus is a very good one and that realization of the research program it implies would represent significant progress for economic and organizational research as a whole, filling remaining theoretical gaps between nonexistence and existence of economic activities and organizations. Although we sometimes portray entrepreneurship as the journey from "nothing" to "something" (Baker & Nelson, 2005), it is clearly the case that entrepreneurs do not create new economic activities out of thin air—there has to be *something* there for them to act upon. Shane and Venkat's point is that the presence and quality of that *something* is as important as the presence and quality of entrepreneurs. Not even Steve Jobs was consistently successful with every venture he attempted.

However, there is a little problem: the fact is that despite mushrooming research on "entrepreneurial opportunities," little conceptual and empirical progress has been made on core aspects of the nexus idea, namely the delineation and effects of salient "characteristics of opportunities" and their interaction with characteristics of the entrepreneurial agent.[2] I argue that the main reason for this limited progress is the very notion of "opportunity" itself. This is not just because of Scott Shane's insistence on "opportunity" being objective, pre-existing, and actor independent (Eckhardt & Shane, 2013; Shane, 2012; Shane & Venkataraman, 2000). I will argue that even if we were to accept the existence and operability of such entities, *they would still be the wrong nexus partner in further development and testing of the entrepreneurship nexus framework*. I will also argue that *any* conceivable conceptualization of "entrepreneurial opportunity" will fail at that mission and that in "Creation Theory" there is no need for the "opportunity" construct at all. My deep

[2] This is not to deny the progress that has been made on topics such as the sources of opportunities (e.g., Eckhardt & Shane, 2003; Holcombe, 2003; Plummer et al., 2007), different types of opportunities (e.g., Eckhardt & Shane, 2003; Companys & McMullen, 2007; Sarasvathy, Dew, Velamuri, & Venkataraman 2003), and (to a lesser extent) operationalizations of their characteristics (Dahlqvist & Wiklund, 2012; Grégoire, Shepherd, & Schurer Lambert, 2010). The greatest conceptual and empirical progress has arguably been made on drivers of the perception of opportunity (e.g., R. A. Baron, 2006; R. A. Baron & Ensley, 2006; Dimov, 2007; Grégoire & Shepherd, 2012; Grégoire, Barr, & Shepherd, 2010; Shane, 2000; Shepherd & DeTienne, 2005), including McMullen and Shepherd's (2006) important distinction between (perception of) first-person and third-person opportunities. However, progress on core issues pertaining to the nexus framework has been limited (Davidsson, 2015).

dive into the literature on entrepreneurial opportunities[3] has convinced me that this construct is essentially like phlogiston[4] for entrepreneurship research. Although it has been useful for a while, clinging to it will severely harm our future development.

8.2 The Allure of "Entrepreneurial Opportunity"

At this point, some readers are probably intrigued by the bold suggestions above and curiously waiting to read my full argument, whereas others are starting to have serious concerns about the mental health of the author. *Of course* there are opportunities to create new businesses that entrepreneurs identify (or create) and act upon! How can you even *talk* about entrepreneurship without referring to opportunities? In fact, it *is* difficult to sustain a conversation on entrepreneurship for more than a minute or two without feeling a strong need for using the o-word or some close synonym. It seems to be a notion we need and have little trouble with in lay conversation. It also appears to be needed and useful in the entrepreneurship classroom, and despite all what I say in this chapter, it may be the case that we can fruitfully continue to use it in that context. But this does not prove it to be conceptually useful for research purposes. In the development of prospective, microlevel theory of entrepreneurial processes, I argue that the notion of opportunity makes us confound one or more things that should be conceptually separated: independent vs. dependent variables, external conditions and subjective perceptions, entrepreneurial agents and the entities they act upon, and the contents vs. the favorability of the latter entity (Davidsson, 2015). That looks like quite a bit of potential confusion.

This said, I freely admit that in some senses the notion of opportunity—even pre-existing and actor-independent ones—is both simple and appealing. In fact, all that is needed in order for this idea to make sense is an assumption of disequilibrium (which we have made throughout this book). The economy is not in full equilibrium; ergo, there is potential for improvements of the economic system. Instances of such potential improvements, which involve changes to what is being offered in the market (cf. Chap. 1), are what we call entrepreneurial opportunities. Duh! This thinking is also in line with our historical experience. In each time period, a number of individuals (or other entrepreneurial agents) try to launch new economic activities, and some of them are successful. This indicates that opportunities were present; the prevailing conditions were such that with more or less creativity it was

[3] Important ideas, an adapted figure, ditto table, and some small snippets of text in this chapter first appeared in Davidsson (2015).

[4] The theorized substance of phlogiston was a leap forward over the ancient, four-element theory of (proto-) chemistry, and in its heyday it helped develop human understanding of phenomena like combustion, metabolism, and corrosion as well as how they are interconnected. However, over time its assumed characteristics turned out to be anomalous and with the new discovery/theory of oxygen, phlogiston was proven nonexistent and the associated theory inferior to the oxygen- and energy-based explanations that superseded it.

possible for these agents to successfully undertake actions that converted their ideas (or hunches or conjectures) into tangible, viable, and goal-achieving ventures.

This indeed makes it hard to think or talk about entrepreneurial processes without referring to opportunities, maybe even to their pre-existing and actor-independent variety. To see exactly *how* difficult, consider the following excerpt from some of the most prominent critics of Discovery Theory and proponents of the Creation Theory notion that opportunities do not exist independently of the agent's perception and are instead socially constructed by the agent in a co-creation journey:

> As the opportunity creation process begins, actors engage in activities consistent with prior beliefs about the nature of the opportunities they might face, together with their understanding of the resources and abilities they have to exploit these opportunities. (Alvarez et al., 2013, p. 308)

If these authors do not believe in pre-existing, agent-independent opportunities, it sure looks like the actors within their theory do so. That is, the actors (agents) in Creation Theory embrace Discovery Theory! If they didn't, they would hardly have reason to start the process by pondering what opportunities they might encounter or whether they have the means to profit from these apparently pre-existing entities.

The fact that some of the most prominent contributors to the opportunity literature stumble like this indicates that "opportunity" is a very difficult concept to define and apply in a consistent manner. In short, there is severe lack of *construct clarity* (Suddaby, 2010) in the literature on entrepreneurial opportunity, which means that by clinging to it we risk engaging in a confusing cacophony rather than scholarly progress. Let's take a closer look at this problem.

8.3 What's Not So Merry About "Opportunities"

In Chap. 3 we noted that well-defined, theoretical concepts are incredibly useful abstractions that allow us to exchange ideas about things outside of our immediate, shared perceptions and sensations. We also noted that well-defined concepts are a necessary foundation of good theory. Is "entrepreneurial opportunity" a well-defined concept that facilitates our scholarly conversation about the phenomena it is supposed to capture and hence a strong foundation for useful theory? A few years ago, Hansen, Shrader, and Monllor (2011) reviewed the "entrepreneurial opportunities" literature—and found a conceptual mess. They noted that even many conceptual papers on opportunities did not define this central concept at all, thus failing to meet what Suddaby (2010) calls the "bare minimal standard of construct clarity." In the minority of empirical papers that provided a definition, the operationalization often did not correspond to the definition given.

Hansen et al. were able to group the various definitions in the literature into six distinct views on what an "entrepreneurial opportunity" is. My own comprehensive review of 210 papers published in leading journals (Davidsson, 2015) confirmed the problematic conceptual state of this literature. A large majority never define the concept and those who do so find it difficult to stay true to their own definition.[5] Implicit and explicit conceptualizations vary markedly in the essential properties assigned to “entrepreneurial opportunities.” They are thus variously portrayed as a confluence of current external circumstances, imagined future ventures, future action paths, or imagined future states. See Fig. 8.1 for some examples.

Scope conditions vary markedly regarding who is the supposed agent, the organizational/market/industry contexts, spatial boundaries, and what degree of novelty and profit intent are required. Table 8.1 enriches the picture by exemplifying the extraordinarily varied language that surrounds the notion of “entrepreneurial opportunity.” In short, this literature is about as close to conceptual clarity and agreement as we are to the North Star; i.e., we are talking light years—and many of them! Remarkable, indeed.

No one owns a concept and the fact that different authors use “opportunity” differently is not a flaw in itself, although it makes communication and stock taking of our collective knowledge more challenging. The greater problem is rather that with or without definition, the meaning of “opportunity” is rarely clear *within* works. This is not because authors in this stream are stupid or lazy but because “opportunity” is inherently and inescapably an elusive concept.

What are the consequences of working with unclear concepts? There are at least two: theorists will fail to use the concept in a consistent manner, and empiricists will apply operationalizations which we don’t know quite what they measure. For example, Shane and Venkataraman (2000) define opportunities as external situations (p. 220), yet the entrepreneurs can choose to sell their opportunity (external situation) to another actor (p. 224). Opportunities are a chance “to introduce” something (p. 220), yet they remain opportunities for as long as they generate a profit (p. 221), which can logically only happen after introduction. Eckhardt and Shane (2010) repeat this contradiction (pp. 49, 54) and also introduce a level inconsistency by sometimes discussing new technologies as opportunities in and of themselves (p. 49), whereas at other times they reserve that term for specific applications of said technologies (p. 61; see also Shane, 2003, pp. 34 vs. 24). Drifting into the empirical, we may note that Shane (2000) defines opportunities as profitable (p. 451), yet four of his eight cases so labeled did not turn a profit, and at least one was not even technologically feasible (p. 455).

Turning to the “creationists,” we have already noted the paradoxical embracing of Discovery Theory on the part of the actors in Creation Theory. There are numerous other instances of confusion in the works by Alvarez and co-workers regarding Suddaby’s (2010) scope conditions time and space (level). As to the latter, it is often unclear whether the descriptions of opportunity creation concern the evolution of a

[5] The definition in the top right corner of Fig. 8.1 has a level of recurring use (with or without adaptation; see Davidsson, 2015) within and across authors. This is the definition originally suggested by Shane and Venkataraman (2000; see also Shane, 2000) with reference to Casson (1982). It is not Casson’s wording, though; Casson (1982) does not offer a definition of “opportunity.” Apart from its creative grammar (sell organizing processes?), note that this definition implies the scope conditions *for-profit*, *innovation*, and a *user pays revenue model*.

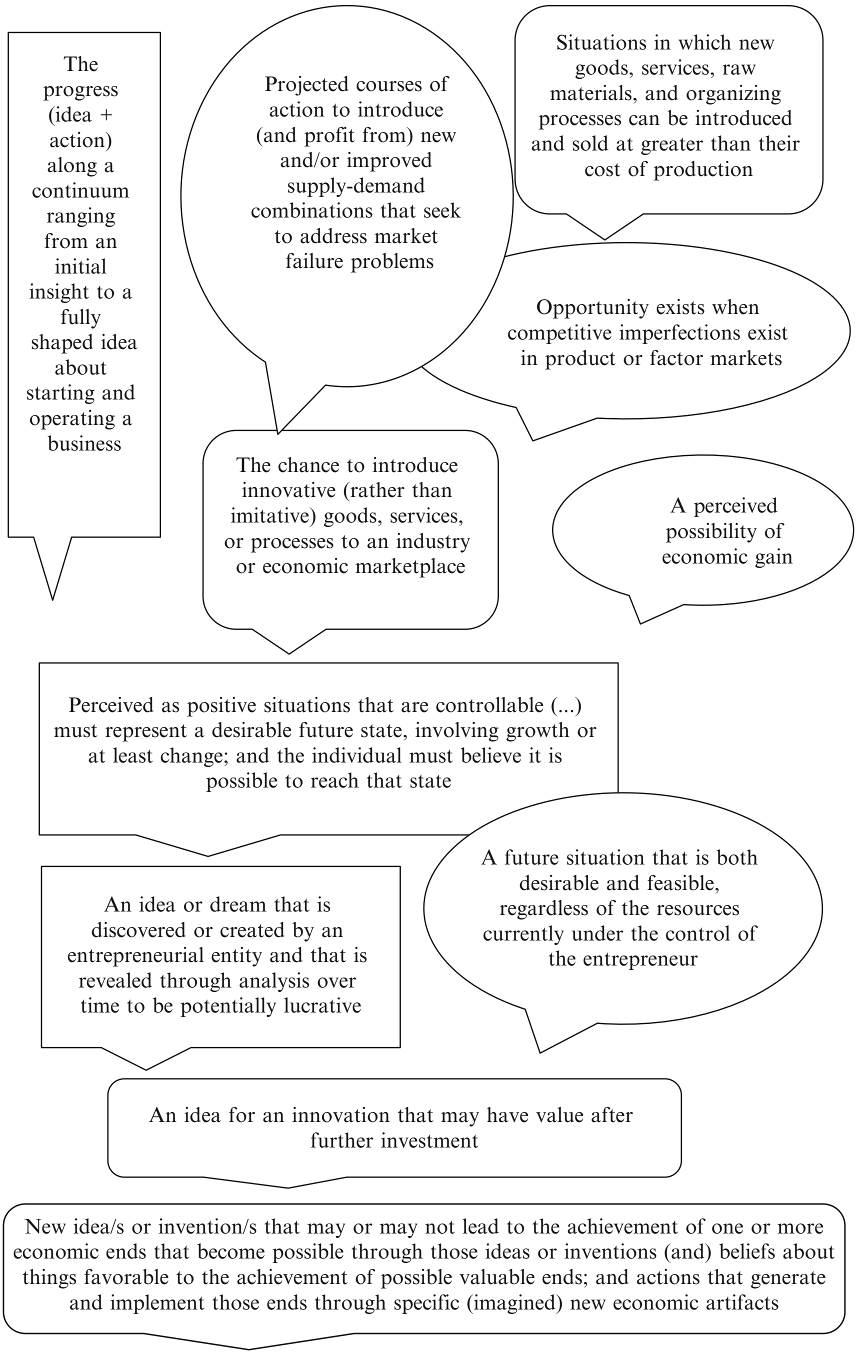


Fig. 8.1 Definitions of (entrepreneurial) opportunity (cf. Davidsson, 2015)

Table 8.1 Examples of what opportunities "are" or "can be" in the "entrepreneurial opportunities" literature

Existence/nature		Perception/search	Evaluation/action	
Accessible	Market based	Attended to	Abandoned	Selected
Advantageous	Motivated	Believed	Acted on	Seized
Afforded	Novel	Compared	Adapted	Shaped
Apparent	Objective vs. Subjective	Conceived	Addressed	Subject to:
Appealing		Defined	Analyzed	– Attitude
Appropriate	Occurring	Detected	Assessed	– Continuation
Arbitrage vs. Innovative	Offered	Discovered	Capitalized on	– Due diligence
	Opened up	Envisioned	Captured	– Judgment
Arising	Present	Faced	Chosen	– Persistence
Attractive	Presented	Focused	Considered	Taken advantage of
Available	Productive	Found	Constructed	
Commercial vs. Technological	Profitable	Framed	Created	Transformed
	Provided	Groped for	Declined	Transitioned
Complex	Real	Identified	Destroyed	Utilized
Demand side vs. Supply side	Risky	Imagined	Developed	Weighed
	Subject to:	Interpreted	Disliked	
Desirable	– Sets	Learned	Enacted	
Emerging	– Space	Missed	Estimated	
Enabled	– Structure	Noticed	Evaluated	
Existing	True vs. False	Observed	Executed	
Exploitable	Uncertain	Perceived	Exhausted	
Feasible	Unexploited	Recognized	Exploited	
First vs. Third person	Unfolding	Searched for	Expressed	
	Unforeseen	Seen	Formed	
Fleeting	Unfruitful	Sensed	Generated	
Fruitful	Unsubstantiated	Sought	Imitated	
High potential	Untapped	Spotted	Implemented	
Indicated	Urgent	Subject to:	Instantiated	
Innovative vs. Imitative	Viable	– Acquaintance	Intended	
		– Alertness	Interpreted	
Internal vs. External		– Attention	Legitimized	
		– Exposure	Made	
Immediate		– Ideas	Maximized	
Heterogeneous		– Insight	Objectified	
Kirznerian vs. Schumpeterian		– Intuition	Passed (on)	
		– Knowledge	Pursued	
Latent		– Vision	Realized	
Legitimate vs. Illegitimate		Tracked	Refined	
		Unknown	Rejected	
Lucrative		Unperceived	Responded to	

successful venture (only) or the opening up of an entire new market (niche) or product category, in which other actors can also play. As regards time, consider the following quote:

> Creation opportunities are social constructions that do not exist independent of entrepreneur's perceptions [references]. However when entrepreneurs act to exploit these socially constructed opportunities, they interact with an environment—the market—that tests the veracity of their perceptions. (Alvarez & Barney, 2007, p. 15; Alvarez & Barney, 2013, p. 155)

Remember here that "opportunity" is defined in a rather objectivist manner as a market imperfection, although this market imperfection is not assumed to pre-exist the entrepreneur's perception and the process of social construction that their action triggers. It is very hard to tease out from the above quote where in the process we are at different places in the statement, and when and on what basis the label "opportunity" is afforded to the entity acted upon. Is the social construction already there, ready to be exploited? Isn't it the process of attempted exploitation that creates the social construction? Attempted exploitation of what? And when the opportunity actually is created—when whatever was acted upon has actually become a market imperfection—why does its veracity need to be tested? Wasn't that what happened much earlier in the process (and throughout it) before whatever was acted upon actually became an "an opportunity"—a truly existing market imperfection that can be exploited for profit? One may also wonder whether the market imperfection seizes to exist because the entrepreneur dies/no longer perceives it. One way of making sense of the statement is to assume that by "do not exist" they mean "will not come into existence" and that they use the same label—opportunity—for the early, unproven perception and for the manifest (albeit socially constructed) market imperfection that has been proven through successful exploitation at the end of the process. But this is not clear.

The following, slightly edited example from an "opportunity recognition" task provides further illustration of the perils of empirical work based on an unclear notion of "opportunity":

> After reading the below description of a new technology, take a few minutes to list any potential business opportunities that you can think of, based on this technology. The ideas can—but do not have to—be related to your current business. [Description of technology]. Please use the space below to list any ideas for new products, services, or business opportunities based on the above technology.

Within an "opportunity" frame, there is in this example no telling what the participants report or are supposed to report. Any ideas for new ventures regardless of their quality? Only those that they consider to have a minimum level of commercial viability for somebody? Or only those they consider to have that level of commercial viability if they themselves would try to exploit it? If we are not clear about what an "opportunity" is or about what the measure captures, we cannot effectively theorize about the drivers and consequences of scoring high or low on this task, and we cannot know whether a high score is somehow "good" or not.

These are examples from highly accomplished scholars and papers published in highly ranked journals. Although there are examples of scholars who evidently think long and hard about their use of the o-word, I would say the average paper in this research stream does worse rather than better than these examples in terms of conceptual clarity. As a result, we have a very confused and confusing conversation, and far less progress than what would be ideal.

8.4 The Merits and Impossibilities of "Objective Opportunities"

8.4.1 Objective Opportunity as a Theoretical Construct and Assumption

After flirting with notions that allow more creation and subjectivity (Eckhardt & Shane, 2003; Shane, 2003), the prime proponents of Discovery Theory have recently returned to the original definition of opportunity as objective, pre-existing, and actor independent (Eckhardt & Shane, 2013; Shane, 2012). On one level, it is easy to sympathize with this view. As noted previously, while some entrepreneurs do fantastic and unexpected things, they are not magicians or gods and hence cannot create ventures, markets, or products through imagination and willpower alone. There has to be *something* out there that allows them to come up with and successfully exploit ideas for new ventures, even if it is likely to require more effort and creativity than Kirzner's (1973, p. 47) ten-dollar bill. Even a social construction requires some preparedness by others to participate in its construction. And as Shane (2012, p. 12) remarks, "I do not know of any entrepreneurship scholar who would argue that scientific advance, political and regulatory changes, and demographic and social shifts do *not* make it possible to introduce new and potential[ly] profitable resource combinations." It would seem ill advised for entrepreneurship research to ignore such things in our effort to understand those phenomena that have been addressed under the rubric of "entrepreneurial opportunities." Shane (2012, p. 15) further explains why he thinks it is prudent and important to view opportunities as objective and actor independent:

> If opportunities are completely subjective and are created by entrepreneurs regardless of the objective conditions surrounding them, then Da Vinci should have been able to found an airline (…) The idea that opportunities—situations in which people have the potential to make a profit—are objective is not a semantic point. It is a necessary concept to preserve the ideas that entrepreneurship can be unsuccessful and that entrepreneurship depends on the nexus of people and opportunities (…) Viewing entrepreneurial opportunities as subjective also clashes with the idea that entrepreneurship involves the nexus of individuals and opportunities. If opportunities are formed in the minds of entrepreneurs, as the subjectivists argue, then the opportunity side of the individual-opportunity nexus is a function of the individual. And if both opportunities and individuals are a function of individuals, then no nexus exists. Instead, all aspects of entrepreneurship are a function of the individual, and the person-centric perspective on entrepreneurship must be correct (…) Therefore, I maintain that objective opportunities must be a central part of the explanation of the opportunity-based perspective on entrepreneurship that researchers have been developing over the past decade.

Noting that we will have reason to come back to some of these assertions, it is worth pointing out also that theorists are entitled to come up with whatever theoretical constructions they wish. Hence, the theorist may postulate that in their theoretical world there exist a finite number of well-defined, ready-to-use entrepreneurial opportunities that the theory's agents may or may not discover and exploit. Perhaps we do not truly believe that opportunities in the real world exist in a form similar to apples that can simply be picked from a tree ready to be eaten, but we may still deem it conceivable that a theory based on such an assumption could lead to interesting insights that may be usefully applied in a real-world setting. Further, computer simulation is a tool that also allows this type of theorizing and which may prove to further augment the insights gained from this theory of objective opportunities. Thus, on an abstract, aggregate level, the theoretical tool of objective, pre-existing, and actor-independent opportunities might work quite well, at least as long as the key interest is not linked to change and evolution of "opportunities" or their role in failure (P. Klein, 2008).

It is when we climb down to the microlevel of individual agents and ventures and get serious about time and process that the conceptual problems with objective opportunities come to the fore. These problems are arguably what have made many of our colleagues raise both ontological and epistemological objections against the idea of opportunities as presented in Discovery Theory and hence refrained from trying to contribute to the realization of the research program implied by the nexus idea. The root—or at least one root—of this problem is the distinction between entrepreneurship as societal phenomenon and as research domain that I developed in Chaps. 1 and 2. Shane and Venkataraman (2000) took an important step toward making this distinction, but not the full stride, and this feeds into the problem of objective opportunities and the original formulation of the entrepreneurship nexus.

Shane and Venkataraman (2000, p. 218) assert that "entrepreneurship involves the nexus of two phenomena: the presence of lucrative opportunities and the presence of enterprising individuals (Venkataraman, 1997)." I argue that in saying this, they are really talking about the societal phenomenon. They are implicitly standing at the end of one or more successful entrepreneurial journeys, looking back. From this vantage point, they can see that the successful case at hand was contingent on certain qualities and actions of the individuals involved. However, they can also identify external circumstances that facilitated the journey. In a thought experiment, they can change or remove this or that circumstance and rather safely conclude that without these conditions present, the end result could not have been (as) successful (or at least that the venture would have had to be different in important ways in order to reach the same level of success). This indicates that somewhere earlier in the journey a confluence of favorable circumstances—an "opportunity"—was present and available for the entrepreneur's exploitation attempt.

By contrast, when we place ourselves at the beginning of the process and embrace entrepreneurship as a research domain (Chap. 2), things look different. On the agent's side of the nexus, we need to start with "potential agents" rather than "enterprising (in the future) individuals," because an important task of the theory would be to help us understand which potential agents will actually take action (and how successfully). On the "opportunity" side, two types of objections are triggered. First, it appears unrealistic to assume that the "opportunity" is fully developed at this early stage. The agent may have to take some action for the "proto-opportunity" to become a fully fledged opportunity, e.g., completing a technological invention, lobbying for a regulatory change, or making a sufficient number of potential stakeholders change their beliefs and/or wants. This is why there is Creation Theory.

Second, and relatedly, in the absence of a known, positive outcome, the dual nature of "opportunity" as consisting of both *contents* and *favorability* reveals itself. We will have reason to return to this important issue shortly. When we theoretically postulate the objective existence of "opportunities," are we then referring to their contents (substance, the particular set of external circumstances that constitute them) or are we also implying that their favorability is an objective fact? Objectively favorable for what? For whom? Says who? Many more of us are prepared to accept the idea of objective existence of the *contents* of what someone wishes to call on "opportunity" than on the objective favorability of that entity. It would thus seem to be the favorability part that is the ontologically more questionable, as it postulates foresight whereas agreement on existence of the contents does not require any knowledge of the future.

8.4.2 Objective Opportunity as an Object of Empirical Research

The conceptual problems of objective, pre-existing, and actor-independent opportunities are small in comparison to the empirical challenges associated with this concept. In short, it doesn't work. At all. Why? Because not even on the aggregate level, and in retrospect, can we infer from, e.g., the frequency of successful start-ups, anything about spatial or temporal differences in the prevalence of "entrepreneurial opportunities" as conceived in Discovery Theory. This is because it is not possible to effectively account for differences in, e.g., the stock of human capital and the quality of other action alternatives (such as paid employment, study, retirement, crime, or just bumming around, i.e., opportunity costs—sorry about mixing in this other notion of "opportunity"). Even if we were to accept their objective existence (including favorability), we can never, ever know the universe of objectively existing "opportunities" that nobody acts upon. As a consequence, we cannot know how much of that universe is or ever will be acted upon. Neither can we sample randomly or probabilistically from the population of objective opportunities.

True, in Chap. 5 I go on about how theoretical relevance is more important than statistical representativeness. Unfortunately, that does not save us here. Having not-acted-upon opportunities in the sample would be *very* theoretically and also practically relevant, because this is the basis for understanding what type of opportunities tend to remain undiscovered. Wouldn't we want to be able to tell students and policy-makers a thing or two about that? Therefore, on the microlevel, any kind of unbiased assessment of the effects of (the characteristics of) actor-independent "opportunities" on action and outcomes would require us to identify and sample "opportunities" independent of them being acted upon. To illustrate the impossibility of this task, consider the following example.

In the mid-1990s, in country B, an entrepreneurial team came up with the then novel idea to provide commuters with a tabloid size newspaper for free, letting advertising pay all costs and profit. The team launched the *RideRead* newspaper in the capital with considerable success.[6] Based on deep contextual knowledge and the luxury of later comparison, one can identify the intricate collection of external conditions that enabled *RideRead*'*s* initial success as including[7]:

(a) High incidence of habitually reading morning papers
(b) A culture of not chatting with strangers while commuting
(c) High minimum wages leading to high distribution costs for traditional newspapers
(d) A well-developed and not overcrowded public transportation system (thus allowing reading on board) governed by a monopoly provider, i.e., a one-stop shop to negotiate (exclusive) access to deliver the product
(e) A strong non-vandalism and non-littering culture, allowing distribution via unattended racks
(f) Traditionalism among incumbents dictating that morning papers be broadsheet (less suitable for reading in commute)
(g) The availability of new, labor-saving print media technologies in combination with strong trade unions blocking their adoption among incumbents
(h) The *non*existence of today's electronic devices and online contents to compete for commuters' attention and advertisers' dollars

Objective pre-existence of the "entrepreneurial opportunity" in this case must refer to the confluence of (a, b…h, and more) at the time the *RideRead* team initiated their start-up. This is the type of entity we would have to identify—in large numbers—*without* anyone having acted upon them. Those who can identify this type of entity a priori tend to be super-entrepreneurs rather than entrepreneurship researchers, and even so I do not think the world has yet seen an individual or team that would be able to identify a large enough sample of these entities to suffice for quantitative work. I would say that most of the time, not even entrepreneurs themselves fully or correctly identify the set of external circumstances that facilitate their success. In the *RideRead* case, if the founders ever became aware of all of (a, b…h) it was probably in arrears when their paper bombed in some other cities around the world, where conditions turned out to be much less favorable.

---

[6] This is, of course, about the freesheet Metro and its original launch in Stockholm in 1995. I anonymized the case only to make it a little bit harder to identify the author in earlier submissions of what became Davidsson (2015). I cut it from the final version and now that I use it here, I really wanted to keep the *RideRead* name as well. Clever, ay?

[7] As a citizen of country B—that is, born a B-countrarian—I was equipped with this contextual knowledge and hence could make sense of *RideRead*'*s* initial success, at least in arrears. I claim no general ability to spot "objective opportunities" with this kind of detail and clarity, and especially not a priori.

The international expansion of *RideRead* also illustrates the fundamental difference between subjective ideas for new ventures (which are also often called "opportunities" in the literature) and objective opportunity as per Discovery Theory. The idea underlying *RideRead* can be summarized as "to provide commuters with a free daily newspaper of decent journalistic quality by saving on distribution and production costs, covering costs and generating profits solely from advertising." This idea remained essentially the same wherever they launched around the world (as I recall it, Boston, London, Paris, Santiago, and Zurich were among the early ones) whereas the list of salient circumstances would differ greatly across sites.

Hence, sampling and measuring objective opportunities is an impossible task even for those who accept their existence or their potential theoretical usefulness.[8] No wonder, then, that much of the (limited) research that has explicitly addressed issues pertaining to the entrepreneurship nexus has been experimental rather than observational (Davidsson, 2015). Although the experimental stream has made good progress, it also has limitations. It is mostly confined to the earliest stages of the entrepreneurial process, and it has a tendency to portray things like "70 % chance of success" as an objective characteristic of an opportunity whereas this assessment in real life—if occurring at all—would arguably be the outcome of substantive characteristics of the individual, the evaluated entity, and the fit between the two.

8.4.3 Why "Opportunity" Is the Wrong Nexus Partner, Anyway

I interpret—with some justification, I hope—Shane and Venkataraman's (2000) individual-opportunity nexus idea as saying that the first-order forces explaining entrepreneurial action and outcomes (see Shane & Venkataraman, 2001) are (a) the entrepreneurial agent, (b) the entity upon which the entrepreneurial agent acts, and (c) the fit and interplay between these two entities. This is depicted in Fig. 8.2. This stylized depiction may look like a mediated moderation model and a linear view of the process. Under that interpretation, it is important to point out that the merged arrow to the left denotes both independent and interactive effects. However, the figure is meant to cover also the case where the arrows only function as markers of overall temporal order, allowing for any type of research that focuses on these entities and their interplay in entrepreneurial processes.

Now imagine that we place ourselves in this figure not at the end of the process looking back, but at its beginning, applying the view of the entrepreneurship research domain developed in Chap. 2. We would then start from an assumption of uncertainty: we don't know who is or isn't going to act, what action path they are

[8] This said, I acknowledge that instances where the focal entity essentially consists of a *site* and its characteristics, such as well-defined geographical markets (Barreto, 2012), real estate development sites (Fiet, 2007), or potential mining sites (Bakker & Shepherd, 2015), offer some potential for a reasonable approximation of "objective opportunity" for some types of research designs relying on observational data. There would still be question marks for "objective favorability," though, as well as ambiguity regarding what are the constituent parts of the "opportunity."

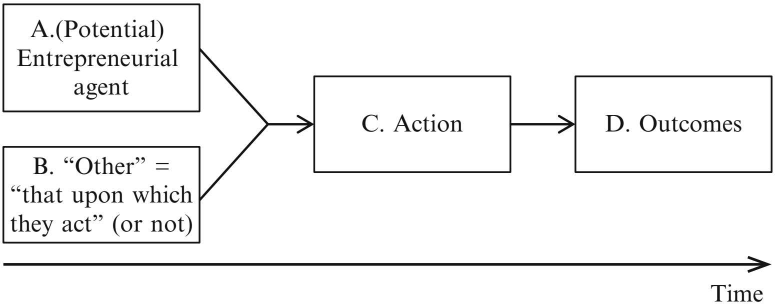


Fig. 8.2 A depiction of the entrepreneurship nexus

going to take, or what outcomes they are going to achieve. So we would start to mold research questions like "What is it about A and B and their interaction that induces (a particular type of) entrepreneurial action?" and "What is it about A and B and their interaction that can explain the type and quality of outcomes reached?". Now imagine that in the B box, we put objective opportunity—"an economic circumstance where if the correct good or service were to be properly organized and offered for sale that the result would be profitable" (Eckhardt & Shane, 2010, p. 48).[9]

The consequence of this is that the nexus idea breaks down. If the B entity is defined as objectively favorable, then failure to act, and a nonprofitable outcome, can only be explained in this model with reference to the agent: it was the wrong agent, or the agent took the wrong action. The opportunity itself cannot contribute to negative outcomes. Putting objectively favorable opportunities in the B box is like reserving the A box for "enterprising individuals," defined as individuals who are particularly prone to engage in entrepreneurial endeavors and bringing them to a successful conclusion. If we did that, negative outcomes could only be due to deficiencies in the opportunity. But the opportunity is favorable by definition, so there would be no negative outcomes.[10] Everyone's a winner! Eh…what was it that I said in Chap. 4 about only studying winners?

The notion of opportunity has an inescapable connotation of favorability, and if the favorability is objective (as per Eckhardt and Shane's definition), it means that the concept is confounded with a range of potential dependent variables, namely

[9] Note that under this definition, there is no uncertainty about the opportunity itself; the only uncertainty concerns whether agents are going to correctly discover and exploit it. Shane (2012, p. 15) tries to get away from this by saying that the definition of opportunity in Discovery Theory only requires the opportunity to have a probability of a profitable outcome which exceeds zero. This, of course, is inconsistent with Eckhardt and Shane's statement and effectively excludes nothing. Since flight was not against the laws of nature even in medieval times, I would grant Da Vinci a success chance of at least 0.00000000000000000001—a number exceeding zero—had he set his mind to creating LEO Air. As we shall soon see, excluding no sets of external circumstances may actually be a good thing, but the more important lesson is that the entire exercise of finding the criterion that draws the line a priori between opportunities and non-opportunities is a vain and unnecessary effort.

[10] Admittedly, the scheme could actually work if we were *only* interested in issues of fit. But admitting this in the main text wouldn't allow as powerful and ending to the paragraph.

outcomes. This is not a good thing. Further, in real life, B entities ("others") are encountered or conceived by all kinds of A entities (individuals and other potential entrepreneurial agents such as teams and organizations), ranging from those who have very low propensity to ponder/act upon/bring to success such B entities to those who are almost pathologically inclined to take entrepreneurial action or exceedingly skillful at achieving entrepreneurial success. Obviously, a theory based on the nexus idea should allow for this variability if it is to be useful for explaining the full range of action, inaction, and outcomes.

By the same token, B entities ("others") can themselves potentially be of any level of inherent quality. If we accept that there exist economic circumstances that allow profitable action (provided the correct good or service were to be properly organized and offered for sale), then we must also accept that there exist economic circumstances that do *not* qualify as opportunities. If we do not do that, the nexus model again breaks down because everything is an opportunity, and variation in action and outcome can only be explained by deficiencies on the part of the agent. Further, in real life agents in all likelihood sometimes evaluate and act on situations that would not qualify as "opportunities" as per Discovery Theory. Again, to capture inaction and failure, the theory should allow for such entities.[11] Objectively favorable opportunity is simply the wrong B entity for this theory's own good.[12] There may be reason to focus on objective, external situations—especially if we can find ways to deal with such a focus empirically—but to artificially restrict the valid cases to those that are favorable by definition leads to all kinds of trouble, from philosophical and logical objections to internal inconsistencies in the theoretical argumentation to overlap with an important aspect of what the theory should be able to explain, namely variance in the outcomes of entrepreneurial processes.

8.4.4 Why Other Notions of "Opportunity" Do Not Quite Cut It, Either

What about other notions of "opportunity"? Can we replace objective, external, actor-independent opportunity with some other conceptualization that also uses that label? No, not really. Why? Well, for starters, *any* conceivable notion of

[11] Current formulations of Discovery Theory both do and do not allow for this. It is sometimes emphasized that opportunities are uncertain and/or that they may have expired by the time the entrepreneur is ready to launch. It is also acknowledged that agents act on subjective conjectures along with an implication that these may be wrong. However, there is no concept for such non-opportunities that agents may unwisely try to exploit or a concept that includes both opportunities and such non-opportunities ("conjecture" denotes positive evaluation and not the substance of the [non-] opportunity; see Eckhardt & Shane, 2003, p. 339). Individual and opportunity remain the elements of the nexus.

[12] One way of making sense of how this came to be is to note that "objective opportunity" as defined by Shane and Venkataraman (2000) may actually suit their first central research question ([1] why, when, and how opportunities for the creation of goods and services come into existence?) but not the remaining two questions ([2] why, when, and how some people and not others discover and exploit these opportunities?, and [3] why, when, and how different modes of action are used to exploit entrepreneurial opportunities?).

"opportunity" comes with a connotation of favorability, and this is always problematic for our current purposes. The Latin root of "opportunity" is *ob portum veniens*, which denotes the wind that allows a ship to make it to harbor.[13] Favorable. Lexical definitions always include favorability of some sort in the definition, often by using that very word. Create a definition stripped of every last smidgen of favorability connotation, and you will immediately ask why on Earth the label "opportunity" should be used for the type of entity you just defined.

We have already seen that objective favorability leads to conceptual overlap with outcomes as well as to range restriction of the non-agent side of the entrepreneurship nexus. We have also seen that the complex nature of "opportunity" as a bundle of contents and favorability triggers internal inconsistencies in the theoretical argumentation more generally and not just when the opportunity is assumed objective and pre-existing. It is not surprising that such lapses happen considering the challenging context: an uncertain process of emergence that evolves nonlinearly over time, meaning that the environment may vary and people may change their minds.

Figure 8.1 shows that many suggested definitions view opportunity as a subjective perception. Shouldn't we give that perspective a go? Why not just accept the agent's viewpoint: if the (potential) entrepreneur thinks it is an opportunity, then let's use that label? After all, we can probably all agree that it makes sense to suggest that potential entrepreneurs choose to act on whatever it is that they see or imagine because *they* believe it is an opportunity for them, right?

Well, there are a few problems (cf. P. Klein, 2008). We would have to grant the o-word to clearly delusional ideas pursued by lunatics. "Opportunity evaluation" becomes a tricky notion, because when someone awards the label "opportunity" to an entity, it has already been evaluated. Otherwise it would not have been given that label. Further, what should be our name for the evaluated entity before the agent decides to act on it, thus dubbing it an "opportunity"? And then we have the important distinction between "third-person" and "first-person" opportunities (McMullen & Shepherd, 2006). What if the focal agent thinks this is not an opportunity for him/her but possibly for someone else? Should the researcher therefore adopt the o-label? More importantly, what if the agent decides not to act or gives up the start-up attempt midway into the process after losing faith? What are those deselected and abandoned entities? As researchers, we cannot leave it to the real-world agents to decide what we mean by our theoretical constructs. The favorability of "opportunity" creates all kinds of problems that we can easily avoid by using a concept that does not imply favorability of particular "B entities" (Fig. 8.2).

A number of thoughtful contributions suggest what we can call the Evolving Idiosyncrasy View of entrepreneurial processes and opportunities. This view is most clearly expressed by Dimov (2011, pp. 64–66) and Sarason, Dean, and Dillard (2006) but appears elsewhere as well. Under this view, "opportunity" is predominantly used to denote a subjective and unproven idea but can also include elements of the unique fit with the agent's person and resources, which facilitates a successful outcome. The idea exists early in the process, but can change considerably during

[13] www.vocabulary.com/dictionary/opportunity

its course, and take on increasing "objectification" over time (Wood & McKinley, 2010), reminding of Sarasvathy's (2001) description of effectual processes. The evolving idiosyncrasy view emphasizes the interplay between agent and the "opportunity" in line with the nexus idea. However, the perspective also emphasizes uniqueness of each "opportunity" and insists on the inseparability of the "opportunity" from the entrepreneur. This is in direct violation of viewing the agent and the "opportunity" as separate entities whose characteristics may have both direct and interactive effects. Thus, this perspective does not invite theorizing about—or empirical assessment of—abstracted characteristics of the B entities in Fig. 8.2, which I would argue is central to nexus theorizing.

Alvarez and Barney's Creation Theory was not developed with the aim to further the nexus framework, so it is not surprising that it can't help us out, either. The definition used is actually similar to the objectivist view—an opportunity is an imperfection in product or factor markets (Alvarez et al., 2013, p. 302; cf. Alvarez & Barney, 2010, p. 559). However, in this theory, this is the successful end result of the entrepreneur's creative journey. Since the opportunity *is* the outcome, it cannot be one of the explanations of it and hence not help us solve our problem. Further, one might ask if "opportunity" even is a necessary concept in this theory. If we are talking about the end result of a creative entrepreneurial journey, then perhaps we could use more established and less elusive terms, like "successful venture/firm start-up" if we are referring to the microlevel or "new [product-] market [niche]" if we are talking about a more aggregate entity that may become populated also by other agents. In the latter case, it would seem that one agent's creation opportunity is other agents' discovery opportunity. Be that as it may; the theorizing is not designed to help us out with our current problem.

8.5 Instead

8.5.1 Requirements for Conceptual Clarity

So what can we do instead? Give up and devote our lives to studying something else, like the sex life of *Taenia saginata*, for example? No offense implied on those who may do just that, but I think the entrepreneurship research community should continue to explore the nexus idea and to study those phenomena we have previously studied under the "opportunity" label. But in order to do so more effectively, we need better, more workable concepts. Further, considering the complexity and variability of these things we have been calling "opportunities," we will need more than one construct to capture the phenomena with satisfactory precision and clarity. I have developed such a scheme in some detail in Davidsson (2015) and encourage the reader to consult that source. Below I repeat some of it and add some ideas not presented in that work.

If we wish to develop a theory about "that, upon which (potential) entrepreneurs (may) act" and how the characteristics of such entities influence entrepreneurial action and outcomes (directly and in interaction with the agent's characteristics), then we need to clearly separate some things that have been blurred in previous conceptualizations:

1. We need to avoid conceptual overlap between the nexus elements (A and B in Fig. 8.2) on the one hand, and that, which they are assumed to affect or explain on the other (C and D in the figure).
2. We need to avoid conceptual overlap between the two nexus elements (A and B in Fig. 8.2). That is, we need to be clear about what attributes are associated with the agent and which belong to the "other."
3. We need to separate the favorability of B from its "substance" or "contents." In the case of objective favorability, this is a special case of (1); if the favorability is the agent's perception, it is a special case of (2).
4. We need to clearly distinguish between external conditions on the one hand and subjective perceptions on the other. I would also hold that to the best of our ability, we should acknowledge that both play important roles in entrepreneurial processes.

Apart from these domain-specific issues, we may benefit from trying to apply Suddaby's (2010) general advice on how to achieve construct clarity: define the constructs, explain their essential properties, and spell out important scope conditions, such as where the concept is situated in time and as regards level of analysis. When I engaged in such an exercise (based on points 1–4 and Suddaby's criteria), I arrived at three constructs: *external enabler*, *new venture idea*, and *opportunity confidence*. The former two are candidates for playing the part of the agent's nexus partner, i.e., the B entity in Fig. 8.2. Opportunity confidence is a supplementary construct which partials out the assessment of favorability from the contents of external enablers and new venture ideas. The essence of these constructs is summarized in Table 8.2 with some further explanation below. A more elaborate presentation and argument is provided by Davidsson (2015).

8.5.2 External Enablers

External enabler (EE) stands for a distinct, external circumstance, which—by affecting supply, demand, costs, prices, or payoff structures—can play an essential role in eliciting and/or enabling a variety of venture development attempts by several entrepreneurial agents. Prior research has demonstrated that focusing on distinct changes in, e.g., technologies, regulations, or sociocultural conditions can enrich our understanding of how new economic activities emerge (e.g., Barreto, 2012; Grégoire, & Shepherd, 2012; Hiatt, Sine, & Tolbert, 2009; Navis & Glynn, 2010; Sine & B. Lee, 2009; Shane, 2000, 2004).

Table 8.2 Concepts suggested to replace "entrepreneurial opportunity"

External enabler	New venture idea	Opportunity confidence
Definition		
A distinct external circumstance, which has the potential of playing an essential role in eliciting and/or enabling a variety of entrepreneurial endeavors by several (potential) agents	An "imagined future venture," i.e., an imaginary combination of product/service offering, markets, and means of bringing the offering into existence	The result of an agent's evaluation of a stimulus (external enabler or new venture idea) as a basis for the creation of new economic activity
Examples of what it is		
Changes (or else hitherto unused potential) in technology, demography, culture, human needs, and wants; institutional framework conditions, macroeconomic conditions, and the natural environment	The *contents* of "imagined future ventures"	An assessment that:
	Any of the eight conceived applications of 3DP™ technology described by Shane (2000)	"EE *x* makes this a good/bad time for people to introduce new ventures"
		"I/someone could (not) profit from using technology X in manner Z to serve market Y"
Examples of what it is not		
The complete set of external circumstances that influence the fate of a particular venture	A manifest venture or business model	Identification of new venture ideas
Necessarily perceived or acted upon	Ideas that increase efficiency of existing operations	The contents of external enablers or new venture ideas
Generally favorable for the economy at large or for all types of ventures and agents	Necessarily innovative, complete, or well articulated; acted upon; successful, or perceived by anyone to be "an opportunity"	Entrepreneurial self-efficacy
		Necessarily even remotely correct
		Necessarily referring to self

Favorability		
Assumed on aggregate level for some activities (unknown which) but not necessarily overall. Favorability for particular applications is not assumed	Not built into construct—revealed through empirical analysis	Degree of favorability—from low to high—is the essence of the construct
	Variable across ideas, agents, and contexts	Subjective and ipsative
		Of varying magnitude and uncertainty
Level		
Aggregate, pertaining to multiple potential activities and agents	Venture (through one or more cognizing agents)	The evaluator: individual
		The evaluated: micro to macro
Time		
Some EEs are always present in a disequilibrium economy; particular EEs are temporary	Existing when cognized, in operation for the duration of the venture creation process	Momentary
		Periodically reassessed

External enabler is an aggregate-level construct. "Enabler" sounds peculiarly favorable given what has been argued above, but this favorability simply reflects the theoretical assumption (and historical experience) that the economy is always in disequilibrium. It is a (rather realistic) theoretical assumption that external enablers give room for some new economic activity; there is absolutely no suggestion that an EE can be known a priori to be favorable for a particular start-up attempt. Further, there is no assumption that EEs are favorable for the economy overall. Hence, tragedies like 9/11, the Chernobyl nuclear disaster, and the Boxing Day tsunami were no doubt EEs for some ventures.

External conditions, and especially changes in them, are unquestionably a trigger of new economic activities. By using EEs as nexus partner, we acknowledge their importance. Grégoire and Shepherd (2012) provide a wonderful and ingenious role model for how this can be done. In an experimental setting, they provide participants with combinations of real technologies and real, unsatisfied needs—i.e., EEs is our current vocabulary—and investigate the propensity to identify potential for new venture attempts based on these EEs. Further, they do what we need to do to make real progress along this track: they theorize the characteristics of EEs, in their case in terms of structural vs. superficial alignment. Future work should follow this example of going beyond merely categorizing EEs as "new technology," "regulatory change," etc. If we do that, we may gain insights into stronger communalities across certain manifestations of different types of EEs and stop implicitly assuming that all cases of one class of EEs (e.g., demographic shifts) are similar to each other and different from other types of EEs.

Longitudinal data collected before and after a "natural experiment" can also be used. Shane's (2003) work on the effects of the Bayh-Dole Act is one example. However, real-time identification of EEs would be highly challenging, although perhaps not as impossible as is early recognition of complete and not-yet-acted upon "objective opportunities." In many cases—notably imitative start-ups and corporate "me-too" entry in mature industries—no clearly identifiable external enabler has a major role in triggering entrepreneurial processes or influencing their degree of success. Therefore, EE as the nexus partner has limited application in research using observational data. While EEs can usefully serve as context for such work—e.g., by sampling cases triggered by the same EE—my main candidate as non-agent nexus element is the *new venture idea* (NVI).[14]

[14] Three arguments have been raised against using subjective ideas in a nexus approach. As we noted earlier in this chapter, Shane (2012, p. 16) argues that without actor-independent "opportunities," there is no meaningful nexus. Second, the evolving idiosyncrasy view holds that the "opportunity" (this usually refers to a subjective idea) is so intertwined with the agent that they cannot be meaningfully separated (Dimov, 2011; Sarason et al., 2006). Third, there is the suggestion that the solution to the elusiveness of "opportunities" is not to put subjective ideas in their place but to increase the focus on actions in the entrepreneurial process (e.g., Dimov, 2011; P. Klein, 2008). See Davidsson (2015) for rebuttals of all three.

8.5.3 New Venture Ideas

I sneaked in a preview of the new venture idea concept in Chap. 2 already. Now the time has come to take a deeper look. NVIs are "imagined future ventures" (cf. Cornelissen & Clarke, 2010; P. Klein, 2008), i.e., imaginary combinations of product/service offerings, potential markets or users, and means of bringing these offerings into existence. They are the *contents* of what others may have called "opportunity recognition/identification/discovery" but *not the favorability*, and they can theoretically be of any quality including ideas that no one would ever want to act on. They may or may not reflect agents' interpretation of EEs; sometimes there is no obvious connection to an identifiable external condition. Although NVIs are cognitions, they are not inseparable from a particular individual and are best represented as a construct on the level of the (potentially emerging) venture. A well-articulated NVI can be shared within a team, transferred between successive champions, or formulated by a researcher and communicated to participants in an experiment. This implies that characteristics of NVIs can be meaningfully conceptualized separately from a particular agent (cf. Katz & Gartner, 1988). This indicates that a major mission for future research is to identify, conceptualize, and operationalize the salient characteristics of new venture ideas. Psychology has the "Big Five" personality characteristics (John & Srivastava, 1999), and Diffusion Theory likewise has five attributes of innovations which affect their rates of adoption (Rogers, 1995). Our work to come up with something as powerful for NVIs does not start from scratch (see Davidsson, 2015), but we sure have a long way to go. When we have fixed that little problem, we can move on to the next theoretical challenge, namely working out the theoretical mechanisms by which characteristics of NVIs operate, and put these to empirical testing.

The fact that an NVI can be of any quality reminds us that *identification* of NVIs does not require positive *evaluation* of them. It is therefore essential to separate the two (see *Opportunity Confidence* below). It is also worth noting that unlike "(perceived/subjective) opportunity," the suitability of the NVI label is not contingent on agent, time, or outcome. Important for the general applicability of NVI as nexus element is that it is inconceivable to have a venture creation process without an NVI. Consequently, the (rather frequent) type of process that Bhave (1994) calls "internally stimulated" does not become a *venture creation process* until the hobbyist-cum-entrepreneur starts to think of the solution as the basis for a business venture, because that is when the solution transforms into an NVI (cf. Katz & Gartner, 1988, on "intentionality").

As regards Shane's (2012) concern that without "objective opportunities" there is no nexus because everything is a function of the individual, note that NVIs presented to participants in an experiment are not generated by those participants. Further, agents are evidently capable of coming up with *and pursuing* ideas of varying quality. This is what habitual entrepreneurs do; in parallel and/or over time, they try to implement a range of NVIs. Sometimes they bomb and sometimes they enjoy great success. The agent remains constant so there would seem to be some potential in investigating the role of the characteristics of the NVI and its fit with the agent.

However, Shane's concern is not completely invalid. In addition, NVIs are undoubtedly both conceptually and empirically challenging because they are often rudimentary at early stages and may undergo considerable change and elaboration as the process unfolds. Then again, as I say repeatedly in this book, research is supposed to be challenging—that is part of the fun! Shane's concern seems somewhat exaggerated, and as we have seen above, it is actually (and ironically) the insistence on using "objective opportunity" as nexus element that makes the nexus idea break down.

8.5.4 Opportunity Confidence

Opportunity confidence (OC) refers solely to an individual's *evaluation*—along the whole range from maximally negative to maximally positive—of a stimulus (EE or NIV) as a basis for the creation of new economic activity. We have shunned favorability above; in the case of OC, the degree of favorability is what the concept is all about. I retain the o-word in this label because here it is justified; agents take action (or make recommendations to others) or not depending on the degree to which *they* view the stimulus at hand as opportunity. The OC construct is not entirely new—it was introduced by Dimov (2010) and others have used other concepts for similar ideas. My definition and elaboration contains new elements, though.

By supplementing EE and NVI with OC, we can remove all connotation of microlevel favorability from the former concepts. Unlike Dimov (2010), I also make a clear distinction between OC on the one hand and *entrepreneurial self-efficacy* on the other. These two refer to nexus elements B and A, respectively, in Fig. 8.2. As the venture creation process progresses, new information may make you more or less excited about the prospects of the NVI you are pursuing, but the journey may also make you adjust your perception of your own aptitude for entrepreneurship. The distinction is important, e.g., for what we can expect an individual to do after they decide to terminate a start-up attempt (Jenkins, 2012).

As pointed out by Dimov (2010), use of the OC concept also reminds us that the level of confidence is something that varies over time. Therefore, it needs to be reassessed periodically in longitudinal research. Further, separating this individual-level perception from the contents of EE and NVI makes it easier to account for different stakeholders holding varying levels of OC with regard to the same stimulus.

Figure 8.3 illustrates how external enablers, new venture ideas, and opportunity confidence are connected (cf. Davidsson, 2015). As noted before, not all NVIs have a clear link to an identifiable EE. This is why the arrow from external enabler to new venture idea is dashed.

Five individuals, 1–5, are depicted. These may be (potential) business founders (agents) or other stakeholders, such as investors. Individuals 1 and 2 are pondering the same external enabler, e.g., a sociodemographic shift or a regulatory change. As can be seen, they arrive at different levels of opportunity confidence in relation to this EE. The reason for this might be a difference in prior knowledge, dispositional optimism, or something else. The OC assessment may be based on the assumption

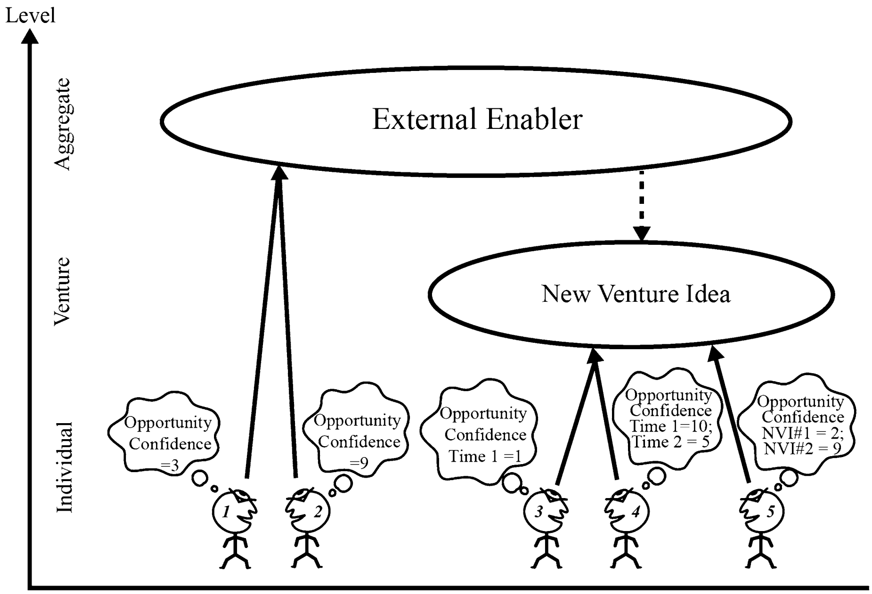


Fig. 8.3 External enablers, new venture ideas, and opportunity confidence

that the agent is the self or a specific other individual or team, or refer to "people in general". Hence, this would have to be specified in empirical research. Individuals 3 and 4 evaluate a new venture idea and do so at the same point in time. While individual 3 discards the NVI, individual 4 finds it fantastic. Again, this would be due to some individual difference. Inspired by the idea, individual 4 acts on it. Hence, he/she has reason to assess it again later on. In this case, the second assessment is less enthusiastic, perhaps because of learning that others are also trying to realize the same type of venture. Individual 5 concurrently evaluates two NVIs. He/she finds NVI#2 quite promising but is lukewarm about NVI#1. This is presumably due to differences in the characteristics of these ideas and perhaps how they fit with individual 5's skills and interests. As a result, individual 5 would be prone to act, or recommend action, on NVI#2 but not on NVI#1.

8.6 Summary and Conclusion

Prior literature on "entrepreneurial opportunities" contains some real gems. Shane (2000) and his emphasis on prior knowledge certainly had a fundamental impact on my thinking about these phenomena. The same goes for McMullen and Shepherd (2006), and I'm sure large parts of the entrepreneurship research community would nod in agreement. Dimo Dimov has shared many a deep thought and significant observation on the subject, and more recently Matt Wood has delighted us with conceptual insights as well as clever empirical designs. I'm in absolute awe of

Grégoire's work and Grégoire and Shepherd (2012) in particular. To name a few. So this chapter has not been meant to denounce prior work. On the contrary, it is only thanks to prior work that I can identify existing weaknesses as well as potential improvements.

This said, progress on important aspects of the "entrepreneurship nexus" idea and the salient characteristics of its non-agent component has been slow, and much effort has been spent on unwinnable debates about the nature or proper conceptualizations of "opportunities"; debates that should perhaps rather have been framed in terms of different types of entrepreneurial processes (cf. Bhave, 1994; Sarasvathy, 2001). The suitability for scholarly purposes of the "opportunity" concept itself has not been sufficiently questioned, and its elusiveness and complexity—especially its dual nature of substance and favorability—have led to a conceptual mess. In the chapter I argued that we need to make clear distinctions where previous conceptualizations have been blurred: between subjective perceptions and external circumstances, between independent and dependent variables, between the two nexus components, and between substance and its favorability. Further, I argued that good theory more generally requires clear concepts and tried to put my money where my mouth is by developing three new(-ish) concepts and subjecting them to Suddabinian scrutiny and elaboration. The result is the triplet of concepts *external enabler*, *new venture idea*, and *opportunity confidence*. I also tried to sneak in a few other adjustments of the entrepreneurship nexus framework in passing, in line with themes introduced in Chap. 2. This includes, e.g., generalizing "individual" to "agent" (or actor) and putting a little more emphasis on action. But the three constructs EE, NVI, and OC make up the main point of this chapter. Now you go and play with your new toys!

* * *

…but please don't wreck them! The development cost was higher than you might think! As long as you keep up the conceptual distinctions, you can call them *opportunity sources*, *opportunity ideas*, *and opportunity confidence* if you feel an irresistible urge to use the o-word and want a catchier triplet. But you can rest assured there are good reasons why I did not do so.

References

Alvarez, S. A., & Barney, J. B. (2007). Discovery and creation: Alternative theories of entrepreneurial creation. *Strategic Entrepreneurship Journal, 1*(1-2), 11–26.

Alvarez, S. A., & Barney, J. B. (2010). Entrepreneurship and epistemology: The philosophical underpinnings of the study of entrepreneurial opportunities. *Academy of Management Annals, 4*(1), 557–583.

Alvarez, S. A., & Barney, J. B. (2013). Epistemology, opportunities, and entrepreneurship: Comments on Venkataraman et al. (2012) and Shane (2012). *Academy of Management Review, 38*(1), 154–157.

Alvarez, S. A., Barney, J. B., & Anderson, P. (2013). Forming and exploiting opportunities: The implications of discovery and creation processes for entrepreneurial and organizational research. *Organization Science, 24*, 301–317.

Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly, 50*(3), 329–366.

Bakker, R. M., & Shepherd, D. A. (2015). Pull the plug or take the plunge: Multiple opportunities and the speed of venturing decisions in the Australian mining industry. Paper accepted for publication in *Academy of Management Journal*.

Baron, R. A. (2006). Opportunity recognition as pattern recognition: How entrepreneurs "connect the dots" to identify new business opportunities. *Academy of Management Perspectives, 20*, 104–119.

Baron, R. A., & Ensley, M. D. (2006). Opportunity recognition as the detection of meaningful patterns: Evidence from comparisons of novice and experienced entrepreneurs. *Management Science, 52*(9), 1331–1344.

Barreto, I. (2012). A behavioral theory of market expansion based on the opportunity prospects rule. *Organization Science, 23*(4), 1008–1023.

Bhave, M. P. (1994). A process model of entrepreneurial venture creation. *Journal of Business Venturing, 9*, 223–242.

Casson, M. (1982). *The entrepreneur*. Totowa, NJ: Barnes & Noble Books.

Companys, Y. E., & McMullen, J. S. (2007). Strategic entrepreneurs at work: The nature, discovery, and exploitation of entrepreneurial opportunities. *Small Business Economics, 28*(4), 301–322.

Cornelissen, J. P., & Clarke, J. S. (2010). Imagining and rationalizing opportunities: Inductive reasoning and the creation and justification of new ventures. *Academy of Management Review, 35*, 539–557.

Dahlqvist, J., & Wiklund, J. (2012). Measuring the market newness of new ventures. *Journal of Business Venturing, 27*(2), 185–196.

Davidsson, P. (2015). Entrepreneurial opportunities and the entrepreneurship nexus: A reconceptualization. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.002.

Dimov, D. (2007). From opportunity insight to opportunity intention: The importance of person-situation learning match. *Entrepreneurship: Theory and Practice, 31*(4), 561–583.

Dimov, D. (2010). Nascent entrepreneurs and venture emergence: Opportunity confidence, human capital, and early planning. *Journal of Management Studies, 47*(6), 1123–1153.

Dimov, D. (2011). Grappling with the unbearable elusiveness of entrepreneurial opportunities. *Entrepreneurship: Theory and Practice, 35*(1), 57–81.

Eckhardt, J. T., & Shane, S. A. (2003). Opportunities and entrepreneurship. *Journal of Management, 29*(3), 333–349.

Eckhardt, J. T., & Shane, S. A. (2010). An update to the individual-opportunity nexus. In Z. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research* (2nd ed., pp. 47–76). New York: Springer.

Eckhardt, J. T., & Shane, S. A. (2013). Response to the commentaries: The Individual-Opportunity (IO) Nexus integrates objective and subjective aspects of entrepreneurship. *Academy of Management Review, 38*(1), 160–163.

Fiet, J. O. (2007). A prescriptive analysis of search and discovery. *Journal of Management Studies, 44*(4), 592–611.

Gaglio, C. M. (1997). Opportunity identification: Review, critique and suggested research directions. In J. Katz & J. Brockhaus (Eds.), *Advances in entrepreneurship, firm emergence, and growth* (Vol. 3, pp. 139–202). Greenwich, CT: JAI Press.

Gartner, W. B. (1988). "Who is an Entrepreneur?" is the wrong question. *American Small Business Journal, 12*(4), 11–31.

Grégoire, D. A., Barr, P. S., & Shepherd, D. A. (2010). Cognitive processes of opportunity recognition: The role of structural alignment. *Organization Science, 21*(2), 413–431.

Grégoire, D. A., & Shepherd, D. A. (2012). Technology-market combinations and the identification of entrepreneurial opportunities: An investigation of the opportunity-individual nexus. *Academy of Management Journal, 55*(4), 753–785.

Grégoire, D. A., Shepherd, D. A., & Schurer Lambert, L. (2010). Measuring opportunity-recognition beliefs. *Organizational Research Methods, 13*(1), 114–145.

Hansen, D. J., Shrader, R., & Monllor, J. (2011). Defragmenting definitions of entrepreneurial opportunity. *Journal of Small Business Management, 49*(2), 283–304.

Hiatt, S. R., Sine, W. D., & Tolbert, P. S. (2009). From Pabst to Pepsi: The deinstitutionalization of social practices and the creation of entrepreneurial opportunities. *Administrative Science Quarterly, 54*(4), 635–667.

Holcombe, R. G. (2003). The origins of entrepreneurial opportunities. *The Review of Austrian Economics, 16*(1), 25–43.

Jackson, S. E., & Dutton, J. E. (1988). Discerning threats and opportunities. *Administrative Science Quarterly, 33*, 370–387.

Jenkins, A. (2012). *After firm failure: Emotions, grief, and re-entry*. Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

John, O., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. Pervin & O. John (Eds.), *Handbook of personality: Theory and research* (pp. 139–153). New York, NY: Guilford Press.

Katz, J., & Gartner, W. B. (1988). Properties of emerging organizations. *Academy of Management Review, 13*(3), 429–441.

Kirzner, I. M. (1973). *Competition and entrepreneurship*. Chicago, IL: University of Chicago Press.

Klein, P. G. (2008). Opportunity discovery, entrepreneurial action, and economic organization. *Strategic Entrepreneurship Journal, 2*(3), 175–190.

Korsgaard, S. (2013). It's really out there: A review of the critique of the discovery view of opportunities. *International Journal of Entrepreneurial Behaviour & Research, 19*(2), 130–148.

Low, M. B. (2001). The adolescence of entrepreneurship research: specification of purpose. *Entrepreneurship: Theory and Practice, 25*, 17–25.

McMullen, J. S., & Shepherd, D. (2006). Entrepreneurial action and the role of uncertainty in the theory of the entrepreneur. *Academy of Management Review, 31*(1), 132–152.

Mitchell, R. K. (2011). Increasing returns and the domain of entrepreneurship research. *Entrepreneurship: Theory and Practice, 35*(4), 615–629.

Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. *Administrative Science Quarterly, 55*(3), 439–471.

Plummer, L. A., Haynie, J. M., & Godesiabois, J. (2007). An essay on the origins of entrepreneurial opportunity. *Small Business Economics, 28*(4), 363–379.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed.). New York, NY: The Free Press.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–240). Orlando, FL: Academic Press.

Sarason, Y., Dean, T., & Dillard, J. F. (2006). Entrepreneurship as the nexus of individual and opportunity: A structuration view. *Journal of Business Venturing, 21*(3), 286–305.

Sarasvathy, S. D. (2001). Causation and effectuation: Towards a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review, 26*(2), 243–288.

Sarasvathy, S. D., Dew, N., Velamuri, R., & Venkataraman, S. (2003). Three views of entrepreneurial opportunity. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research*. Dordrecht, NL: Kluwer.

Shane, S. A. (2000). Prior knowledge and the discovery of entrepreneurial opportunities. *Organization Science, 11*(4), 448–469.

Shane, S. A. (2003). *A general theory of entrepreneurship: The individual-opportunity nexus*. Cheltenham, UK: Edward Elgar Publishing.

Shane, S. A. (2004). Encouraging university entrepreneurship? The effect of the Bayh-Dole Act on university patenting in the United States. *Journal of Business Venturing, 19*, 127–151.

Shane, S. A. (2012). Reflections on the 2010 AMR Decade Award: Delivering on the promise of entrepreneurship as a field of research. *Academy of Management Review, 37*(1), 10–20.

Shane, S. A., & Eckhardt, J. (2003). The individual-opportunity nexus. In Z. J. Ács & D. B. Audretsch (Eds.), *Handbook of entrepreneurship research* (pp. 161–194). Dordrecht, NL: Kluwer Academic.

Shane, S. A., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *Academy of Management Review, 25*(1), 217–226.

Shane, S. A., & Venkataraman, S. (2001). Entrepreneurship as a field of research: A response to Zahra and Dess, Singh, and Erikson. *Academy of Management Review, 26*(1), 13–16.

Shepherd, D. A., & DeTienne, D. R. (2005). Prior knowledge, potential financial reward, and opportunity identification. *Entrepreneurship: Theory and Practice, 29*(1), 91–112.

Short, J. C., Ketchen, D. J., Jr., Shook, C. L., & Ireland, R. D. (2010). The concept of "opportunity" in entrepreneurship research: Past accomplishments and future challenges. *Journal of Management, 36*(1), 40–65.

Sine, W. D., & Lee, B. H. (2009). Tilting at windmills? The environmental movement and the emergence of the US wind energy sector. *Administrative Science Quarterly, 54*(1), 123–155.

Sorenson, O., & Stuart, T. E. (2008). Entrepreneurship: A field of dreams? *Academy of Management Annals, 2*(1), 517–543.

Suddaby, R. (2010). Editor's comments: Construct clarity in theories of management and organization. *Academy of Management Review, 35*(3), 346–357.

Venkataraman, S. (1997). The distinctive domain of entrepreneurship research: An editor's perspective. In J. Katz & J. Brockhaus (Eds.), *Advances in entrepreneurship, firm emergence, and growth* (Vol. 3, pp. 119–138). Greenwich, CT: JAI Press.

Venkataraman, S., Sarasvathy, S. D., Dew, N., & Forster, W. R. (2012). Reflections on the 2010 AMR Decade Award: Whither the promise? Moving forward with entrepreneurship as a science of the artificial. *Academy of Management Review, 37*(1), 21–33.

Wiklund, J., Davidsson, P., Audretsch, D. B., & Karlsson, C. (2011). The future of entrepreneurship research. *Entrepreneurship: Theory and Practice, 35*(1), 1–9.

Wood, M. S., & McKinley, W. (2010). The production of entrepreneurial opportunity: A constructivist perspective. *Strategic Entrepreneurship Journal, 4*(1), 66–84.

The Power of Replication 9

"All studies have limits. It is only in their combination that evidence reveals itself"

(Rousseau, Manning, & Denyer, 2008, p. 50)

Abstract

How can we develop more solid knowledge about entrepreneurship? Like in other fields of research, the truth is that we never know the truth, but that we can arrive at increasingly accurate approximation of it. In this collective quest of knowledge development, statistical significance testing is a sadly overused tool, while replication of prior research is a better but sadly underused tool. After reiterating the limitations and frequent misuse of significance testing, this chapter illustrates how we can make progress by replicating others (traditional replication studies), each other (harmonized research collaboration), and ourselves (using multiple samples or sub-samples; robustness testing). The chapter ends on a high note with observation of several signs that our research culture may finally be about to start embracing the importance of replication and reproducibility.

9.1 Sampling and Significance Testing Revisited

In Chap. 5, I argued that there is no way we can sample probabilistically directly from the theoretically relevant population. This is because that population does not exist empirically in one place at one time. As a corollary, I emphasized that statistical significance testing is not the ideal teller of truth that we would like it to be. I have also hinted at (or been whingeing about) the shortcomings of statistical significance testing and its application elsewhere in previous chapters. It is now time to take a deeper look at this problem.

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_9

Before I resume my whingeing,[1] let me reiterate that I am very impressed by the *statistical estimation* tools that clever statisticians and econometricians have put to our disposal. Developing the statistical inference apparatus was no mean feat, either—it is with the (mis-)*application* of the statistical significance part in the social sciences that I have a problem. Let me also make clear that I fully acknowledge that statistical significance testing is an aspect of a *theory* (Liero & Zwanzig, 2013). Just like a substantive theory can be useful even if it does not offer a perfect or complete explanation of our focal phenomena of interest, I acknowledge that the method theory of statistical inference may be useful also in situations when its underlying assumptions are not fulfilled.

9.1.1 Statistical Significance as Statistical Nonsense

Notwithstanding the above, I cannot escape the conclusion that the mindless and incorrect ways in which statistical significance testing is actually used and assessed by authors, reviewers, and editors is a big and ongoing scandal. Due to its ubiquity, the problem of misapplication of statistical significance testing may actually be a greater threat to the credibility and productivity of our "industry" than are other serious—appalling!!!—problems like plagiarism and fabricated data. Moreover, it is not a problem only in entrepreneurship or management research; it flourishes all across the social sciences and beyond (Ioannidis, 2005).

So, what's the problem? Oh, my, where do I even start? Maybe I should begin by pointing out that this is not just me having some quirky pet peeve. I would say every serious scholar, who actually knows to some depth what statistical significance testing is and is not, is deeply concerned about our current practices. Below are a few examples. Study the titles as well as the author and journal credentials. If that makes you feel that you should be a bit concerned, too, then proceed to reading some of these works. If you want to be a serious, high-quality scholar, you can't just follow established practices when it comes to statistical significance testing. Current practices are deeply flawed, and that may apply also to the practices of the role models you hold in the very highest regard.

- The Earth is round ($p < 0.05$) (Cohen, 1994)
- What statistical significance testing is, and what it is not (J. Shaver, 1993)
- The case against statistical significance testing (Carver, 1978)
- Why most published research findings are false (Ioannidis, 2005)
- From significant difference to significant sameness: Proposing a paradigm shift in business research (Hubbard & Lindsay, 2013a)
- The significant difference paradigm promotes bad science (Hubbard & Lindsay, 2013b)

[1] It's 'Strailian, mate! And for you Aussies: yes, with both an "e" and an "h." Stop whingeing about my spelling…

- Researchers should make thoughtful assessments instead of null-hypothesis significance tests (Schwab, Abrahamson, Starbuck, & Fidler, 2011)
- False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant (Simmons, Nelson, & Simonsohn, 2011)
- The chrysalis effect: How ugly initial results metamorphosize into beautiful articles (O'Boyle, Banks, & Gonzalez-Mulé, 2014)
- Our scholarly practices are derailing our progress: The importance of "nothing" in the organizational sciences (Landis & Rogelberg, 2013)
- The presence of something or the absence of nothing: Increasing theoretical precision in management research (Edwards & Berry, 2010)

Here is what "statistically significant at 5 % risk level" (i.e., $p < 0.05$) tells you when the data support a previously stated directional hypothesis against the conventional alternative that the true effect or difference is zero:

1. If you have drawn a random sample from a well-defined population and you have a response rate of 100 %, you know that *if the actual difference/effect in the sampled population is exactly zero* (and the variance in the sample matches the true variance in the population), then the probability of obtaining a hypothesis-supporting difference/effect of the observed size (or larger) in a sample of the chosen size is less than 5 %.[2] Although the purpose of the test is to gain information on external validity (i.e., what might be true for cases *not* studied), you only learn something about *the probability of the sample result given an arbitrary assumption about the population*, namely that the true difference/effect is exactly zero. We know without collecting data that this assumption is never exactly true (because there is always a difference/effect is some direction, if only at the *n*th decimal); whether it is (nearly) true or not, we can only learn for sure by investigating (nearly) the entire population. The 5 % risk denotes the likelihood that the obtained, supportive result is entirely due to random sampling error if there is no real difference/effect in the sampled population. If the population is not well-defined and/or the response rate not 100 %, we can *perhaps* be generous and say that the test is valid in relation to *the proportion of members of the sampling frame used, who would have participated in the study if they had been asked*, i.e., the test result speaks to the probability of obtaining the observed result if the effect in this cooperative subpopulation is zero. If you have not sampled randomly, then the significance test does not tell you anything meaningful at all that you have not already learnt with greater precision and clarity from examining the size of the effect/difference observed in your sample.

[2] I assume here that you apply a one-tailed test, i.e., that you halve the "associated probability" (p-value) typically reported by your statistics package. Reporting one-tailed tests for directional hypotheses is not a "dubious practice" as some geniuses out there would have it—it is a logical and linguistic necessity if you want "5 % level of risk" to mean "5 % risk of reporting a false positive result."

2. If you have an experimental study applying random assignment of participants to alternative experimental conditions, what you learn from the test is that using these participants in this experiment, there is less than 5 % risk of obtaining the observed (or larger) difference/effect solely because of "random assignment error" when there is no real effect of the experimental manipulation. That is, it speaks to the risk that although you apply random assignment, you may happen to get people in the different experimental groups who were different already with respect to the outcome you are investigating, and this alone is what drives your supportive result. Since the participants are typically a nonrandom collection of volunteers rather than a random (probabilistic) sample, the purpose of statistical testing in experiments pertains only to *internal* validity. We cannot use statistical significance to make any statements relating to cases not investigated, be they another possible set of volunteer participants or a (hypothetical) representative sample from the entire, theoretically relevant population of human beings. Neither does the test tell us anything about the outcome of other possible tests based on other experimental designs aiming at the same underlying theoretical issue, even if we assume using the same participants.

In neither case, does the test tell us how serious it would be to make an error of inclusion (Type I) versus an error of exclusion (Type II), which is something we really need to know in order to justify the chosen significance criterion (critical *p*-value)? You cannot adjust the critical *p*-value in order to reduce the risk of one of these errors without increasing the other risk. The simple truth is that we have no solid basis for applying the significance criteria we typically apply (Leahey, 2005). Applying the same criteria regardless of research question and sample size is just stupid laziness (or resignation to perception of such characteristics in other powers that be). Further, an effect being "statistically significant" does not mean that it is true for the relevant theoretical population or—if it happens to be true—that it is theoretically of practical importance (Edwards & Berry, 2010; Kirk, 1996). In short, statistical significance is not a very powerful truth criterion.

In stark contrast, many researchers act as if statistical significance were definitive (or at least very strong) proof of an effect of a certain size in the theoretically relevant population. They also tend to express themselves as if lack of significance conclusively proved the absence of such an effect. The *p*-value is often explicitly or implicitly interpreted as a measure of the size and/or importance and/or truth of an effect, although the fact of the matter is that the outcome of a significance test is contingent also upon a range of factors other than the effect size in the population, e.g., (a) the chosen risk level, (b) the size of the sample, (c) random sampling error—how much the effect in the sample deviates from the true effect in the population, (d) model specification—the form and structure of analyzed relationships and what variables are included in and excluded from the analysis, (e) the variance of the variables concerned, (f) the quality of the measures, (g) the power of the specific testing procedure, and (h) probably a few more things I have forgotten right now.

To make matters worse, many researchers seem to believe that "significant on the 5 % risk level" means that the result has a 95 % chance of being replicated (Oakes, 1986). This is gross exaggeration of the power of significance testing and perhaps an explanation why business researchers underemphasize the importance of carrying out replications (Evanschitzky, Baumgarth, Hubbard, & Armstrong, 2007; Hubbard, Vetter, & Little, 1998). The truth is that if the associated probability is exactly 0.05 and we happened to be right on target regarding effect size and variance (i.e., the effect and variance in the original sample exactly match the population parameters), the chance of successful replication—meaning the replication also gives a "significant" result according to the "$p<0.05$" criterion with an equally sized sample—is a mere 50–50.[3]

But surely, problematic application of statistical testing appears only in older research and/or in low tier journals? Or only in fields other than entrepreneurship? Wrong again. I tend to find reason to object to the practice in almost every empirical article I read, including those I use as examples of (otherwise) "exemplary entrepreneurship research" in doctoral training. As an example of the latter, take Navis and Glynn's (2010) paper in *Administrative Science Quarterly*. This fascinating, mixed-method piece studies the *population* of firms in the US satellite radio industry (namely two firms), and to test their Hypothesis 4 ("When a new market category achieves legitimacy, audiences will shift the emphasis of their attention from the collective identity of the category to the organizational identities of the individual members of the category"), the authors content code the *entire population* of analyst reports from investment firms that issued reports on both firms in the category. The authors report their findings as follows:

> We conducted a series of unpaired *t*-tests to compare audience attention before and after 2002; results for the combined analysts' reports show that the patterns in figure 3 are statistically significant. All *t*-values are significant at the $p<0.05$ level and in the expected direction. (p. 459)

What they fail to mention in the text is that for the number of market category mentions per analyst report, the early-to-later period ratio was 3:1 whereas for mentions of individual organizations the corresponding ratio was 1:3. That's a pretty strong result that should be clearly stated in the text. Statistical significance has nothing to do with it; what the authors found is either a fact about this industry during the studied period or the result of some hugely biasing measurement error about which a significance test provides absolutely no information. Neither the industry nor the satellite radio companies nor the reporting investment firms nor the reports

[3] It almost brings me to tears that otherwise smart and knowledgeable PhD students never find the right answer to this problem when I pose it to them. As the problem is formulated, the only insight you really need is that if you repeatedly draw random samples from a population with effect size x, half of the samples are going to show an effect larger than x, whereas the other half will yield estimates smaller than x. Sampling variation in the variance of variables may mean that this example's assumed 1:1 relationship between effect size and significance does not hold to the full, but these deviations would cancel out for a fifty-fifty end result.

were subject to any probability sampling. Statistical inference theory is of *no help at all* in establishing external validity in this case. Whether we should expect the result to hold up for other emerging "market categories" at other places and times is something we would have to assess and argue for in other ways.

While the above exemplifies a small blemish on otherwise very well-executed research, it demonstrates that when it comes to statistical significance you cannot trust authors and editors even in the best of places. You might think *Organizational Research Methods* (ORM)—our leading methods journal—would be safe ground, but it just ain't so. A few years ago, ORM published a special issue on research methods in entrepreneurship. One of the articles in the issue surveyed research practices in our field by coding the content in *all* (entrepreneurship) articles in our leading journals during two time periods. The purpose was to determine whether there was any development over time. Again, these are *populations* of research articles, so there is no relevant statistical uncertainty concerning whether observed effects are real or not. Yet, on the reporting of correlation matrices, the authors (Crook, Shook, Morris, & Madden, 2010) state that (read "PUI" simply as "percent/100"):

> Although there was an increase in the number of studies that reported matrices in the later period (108 [PUI: .78] vs. 72 [PUI: .70] for early—$p<0.01$), there was not a significant difference in the number of studies that reported full matrices. (p. 198)

All that needs to be said here is that reporting of correlation matrices in entrepreneurship articles in these journals increased from 70 to 78 % whereas the reporting of full matrices only increased from 53 to 55 %. We would also appreciate commentary on whether the authors find these rates of increase encouraging or disappointing, or worthy of any interpretation at all. Whether the results reflect trends that are true before or after the studied time periods, or for entrepreneurship articles published in other outlets, is something the statistical test cannot help answer *at all.* The test result is irrelevant information or rather dis- or at least mis-information.[4] In a similar vein, the authors refrain to comment on a doubling—admittedly from a

[4] In the next paragraph (also p. 198), the authors offer the following, peculiar analysis "As indirect indicators of substantive and external validity, we coded the number of independent-to-dependent variables in each study that were statistically related (…) the ratios of statistically related to unrelated variables for the periods were 0.87 and 1.05, a significant increase over time ($p<0.01$)." Again, the increase over time is a fact about the studied population of articles or the result of some measurement error about which the statistical test is silent. More importantly, I am mystified as to how this trend is supposed to reflect improvements in "substantive and external validity." Assuming that by "statistically related" they mean that the relationships "achieve statistical significance," the effect probably reflects an increase in average sample size, which likely indicates an improvement in research quality. However, the effect probably also reflects that authors and/or editors are becoming less prone to submit and accept for publication, respectively, papers with (many) non-supported hypotheses. If so, this indicates increased confirmation bias. While such bias is a pervasive human trait, it is certainly not an indicator of research quality (Davidsson & Wahlund, 1992; Fanelli, 2010).

low base—of the prevalence of experimental research, presumably because this fact is not associated with the irrelevant significance asterisk. Undeterred, the authors continue a few pages later with a "Study 2" of what they themselves characterize as (the 35 % cooperative part of) the "population of experts," happily letting entirely unnecessary and uninformative significance tests accompany the reporting of facts about this population's answers to various questions. There simply is no larger, underlying population to which we can draw inferences here, so there is no need to lean on statistical inference theory.[5] In the ORM issue following the special issue, one of the special issue guest editors continues the practice of overfocusing on statistical significance (Holcomb, Combs, Sirmon, & Sexton, 2010):

> From the results, we conclude (a) the linear trend in employment growth and ROA following an IPO is negative and statistically significant, (b) ventures nested within industries differ in terms of their initial performance levels, (c) the linear rate of performance change among ventures within industries following an IPO is statistically significant for ROA but not for employment growth, and (d) differences in the linear rate of performance change among industries are not significant for either measure of post-IPO performance. (p. 364).

It would have been more useful to get an assessment of whether the effects obtained in this population (all single-product firms going through an IPO in the USA in 1996) are of a theoretically and practically meaningful magnitude. A discussion of whether the same effects should be expected in other theoretically relevant populations would also be in place. The significance test provides neither. There seems to be nothing that makes otherwise brilliant scholars turn their brains off like the lure of the significance asterisk.

Predictably, more average scholars fare no better. I would not normally name-and-shame a particular example of altogether bad research, but the example below can at least achieve some value as a pedagogical illustration of the type of statistical significance insanity we should all be able to detect and avoid. Chan, Bhargava, and Street (2006) study a small population (or nonrandom sample) of award-nominated "gazelle" firms in order provide evidence of the homogeneity of high-growth small firms. That is, their hypotheses have the form "The key organizational challenges of high growth small firms are *not* influenced by [the firm's size or industry]." Contrasts are performed across subcategories consisting of some 15–45 firms per contrasted category.

The results? Well, they show, for example, that "leadership" is a challenge according to 16.7 % in the smallest (by employees) size class, by 8.7 % of the middle size class, and by 4.1 % in the large (over 300 employees) size class. That is, the challenge was reported more than four times as often by the smallest firms

[5] Some colleagues would refer to "a hypothetical population" and/or "safeguarding against the influence of some unknown stochastic process" (as the culprit behind the observed difference/effect) to justify statistical testing. I have occasionally done so myself, but I think we are just kidding ourselves when we try such defenses. For lack of better alternatives, we continue to fantasize that significance testing is that strong, truth-telling tool that we need, but which simply does not exist.

compared to the largest. Comparing problem categories within size classes, the respondents in the smallest firm size class are 75 % more likely to report leadership as a challenge compared to human resource management (16.7 vs. 9.5 %). In the largest firm size class, the latter type of problem is reported nearly six times as often as the former (4.1 vs. 23.5 %). Quite a difference, I'd say! Further, "managing business growth and development" was reported as a problem by 45.0, 38.1, and 32.2 % of firms in revenue size classes from smallest to largest. This seems to me a pretty clear, negative relationship by firm size. Across industries, "customer management" is reported as a problem more than twice as often by manufacturing and by "agriculture and other businesses," compared to retail and wholesale, or services (over 10 % compared to just over 4 %), and "human resource management" was seen as a problem by 27.1 % of service firms but only 16.2 % of manufacturers, i.e., reporting occurrence was 2/3 higher in the former category.

The interpretation? Well, since none of the tested differences were statistically significant, they obviously support the hypotheses as stated, right? That's what the authors think:

> To sum up, an important contribution of our study was the empirical finding that successful small firms, having attained high growth, were indeed largely homogenous with respect to their key challenges. (p. 437)

I hope every reader understands that exactly the same results would have been "statistically significant" with flying colors had the total sample size been 910 firms instead of 91. If you are inclined to interpret absence of "statistical significance" as "proof that there is no effect at all," you only need to reduce your sample size to be guaranteed of "evidence" in favor of a hypothesis of noneffect. It works every time!

To state support for hypotheses of noneffect based on a conveniently small, nonrandom sample where the results suggest clear and sometimes large effects is about as bad as published statistical-significance-nonsense gets. However, interpreting absence of significance as evidence for absence of effect is commonplace. When they do not get support at conventional significance levels, authors quite often say things like "Hypothesis *x* is rejected; A does not have a positive influence on B." The first part of that sentence *may* be sound; however, the latter part is definitely not sound if the estimated effect is in the expected direction and of a magnitude that—*if true*—would be theoretically and/or practically important. What the test shows is that by the criteria you have chosen to apply, you cannot *exclude the possibility* that A has no positive effect on B in the population from which your sample was drawn. The fact that it doesn't come out significant says something about the suitability of your sample size in relation to the magnitude of meaningful effects (i.e., statistical power), but it does not *prove* that there is no effect in the population. If there is only, say, 6–10 % risk of obtaining your observed result when sampling from a population where the effect is zero, then the probability is pretty small that zero effect is the correct post hoc assumption to hold about the population you actually sampled from.

9.1.2 Naïve Belief in Fishermen/Women and Their Stories

The biggest problem with statistical significance as truth criterion is not one of those discussed above, but instead that researchers are to a considerable extent "fishing" for statistically significant findings. By this I do not mean that they collect data at random, and then correlate everything with everything, and proceed to building a story solely around those relationships that came out as "statistically significant." More likely, researchers collect their data based on research questions that interest them, and equipped with a priori ideas about how—at least in terms of sign and direction of influence—some key variables are causally related. However, with the possible exception of some experimental research situation, they have probably not made irreversible commitments regarding what specific relationships to focus on in the next conference paper or journal submission, decided on an exact analysis model in terms of all variables and relationships to be included, or selected one and only one precise operationalization of each variable in the model.

More likely, they have included a number of "suspected" explanatory variables in the research, and they start by exploring which of these "suspects" seem more "promising." Based on this early exploration, they may drift into probing further into x_1–x_3 rather than x_4–x_6. Perhaps this is because in the light of the results, they start to doubt the validity of some of the latter, and perhaps these suspicions are sound. They may come to realize that an alternative operationalization of the dependent variable may be better than the one originally planned, and this may also be (or seem) justified. They may find reason to trim the sample of some "outliers" or other cases suspected of potentially "ruining" the results. They may also refine their ideas regarding the form and structure of the relationship. A vague initial idea of a positive influence, first tested as a direct and linear effect, may become a mediated and/or curvilinear effect. Control variables may be added to or removed from the model in order to make it "work better"—which usually means "show significant results for at least some of our key variables."

So, the end product which we get to see may focus on three "independent variables" from an original set of ten or so, while the others are either left out or included only as controls. The dependent variable has been modified, as have the structure and form of the relationships (in graver cases, the sign may have changed from a positive to a negative influence). These focal relationships support the hypotheses as stated, and to each hypothesis has been retrofitted a theoretical rationale that was not fully developed a priori (in graver cases, the theoretical rationale is completely different from what was originally [vaguely] conceived). This is how much research is actually done, and much of this wrestling between theory and data may actually be a sound—albeit risky—way of developing a theoretical understanding of what is going on. We are all guilty of it to varying degrees, me included. Unfortunately, it is a practice not just driven by individual researchers' (biased) sense-making attempts but also by an academic publishing culture that actively fuels this great game of confirmation bias through various overt or covert editorial policies.

What does this mean for the interpretation of "statistical significance"? Well, to begin with, it means that the probability of reporting false-positive results is far larger than the supposedly "conservative" 5 %. Remember that one out of 20 tests will be "statistically significant" when the true effect is zero. Therefore, if being statistically significant is any part of your criteria for selecting a relationship for inclusion in your paper, then the probability of false positives will tend towards 100 % (Armstrong, 1970; Ioannidis, 2005; Simmons et al., 2011). "Statistically significant" only means "highly unlikely to be a false positive" if all hypotheses and test procedures (models, operationalizations, analysis techniques) are decided in advance, and we only get a fair chance to evaluate the findings if the results for all hypotheses—including those not supported by the data—are included in the reporting. Ergo, a *very* large proportion of all published tests of statistical significance tests are invalid.

In most real-world contexts, we investigate there would probably exist some real relationships—i.e., our data are not like matrices of random numbers—and this means that the likelihood of false positives does not reach 100 %. But how bad does it get in reality? Let's look at a realistic example. In Table 9.1 I report some regression models predicting entrepreneurial intentions within six separate samples of individuals in the same age brackets, in the same country, and at the same time (cf. Davidsson, 1995b).

According to my conceptual model, the most direct influence on intentions to start a business should be (a) their degree of conviction that this is a suitable career choice for them and (b) their current employment status, with those in permanent employment being less inclined to strike out on their own any time soon. As indirect antecedents, I included certain general attitudes as well as some domain-specific beliefs. As the most distal influences, I modelled certain personal background factors like sex and access to role models. Thus, the regressions in Table 9.1 represent a quick and dirty way of exploring the proposed relationships, because entrepreneurial intention is here regressed directly on all proximal as well as distal antecedents. Further, they display only the antecedents that turned out significant in the full sample analysis (a few additional ones were included in the conceptual model). For our current purpose, I should have rerun the analysis with these nonsignificant variables, but I'm not even sure current versions of SPSS would be able to read those prehistoric data files….

The results tell us a great deal about what kind of results are and which are not replicable in a "normally" sized study. The overall explanatory power is fairly stable, as is the strong effect of *conviction.* For the other variables, the results vary quite a bit in terms of relative effect size and—especially—in terms of statistical significance. If you weren't already convinced that "significant at $p<0.05$" does NOT mean "95 % chance of being replicated in another sample of the same size from the same population," perhaps this example can do the trick?

Now imagine that we had the usual situation of only having one of those samples or that six researchers were independently examining one sample each, all starting from a vague hunch that with the exception of "lacks role model" all these variables would have a positive influence on entrepreneurial intentions. Imagine further that the researchers apply the type of semi-exploratory search for

Table 9.1 Determinants of entrepreneurial intentions in six separate samples of Swedish adults (cf. Davidsson, 1995b)

Variable (hyp. effect)	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Full sample
Conviction (+)	0.49***	0.56***	0.54***	0.51***	0.46***	0.53***	0.52***
Situation							
Temporary empl. (+)	0.14**	0.05	0.10*	–0.00	0.11*	0.07	0.08***
Unemployed (+)	0.03	–0.01	0.03	0.01	0.03	0.09	0.04*
Attitudes							
Change-orientation (+)	0.15**	0.08	0.01	0.06	0.09	–0.00	0.07***
Achievement (+)	0.07	0.04	0.12*	0.10	–0.02	0.05	0.06**
Beliefs							
Societal contribution (+)	0.17**	0.06	0.10*	–0.00	0.08	0.12*	0.09***
Know-how (+)	0.06	0.14**	0.04	0.04	0.20**	0.05	0.08***
Personal background							
Sex ($m=1$; $f=0$) (+)	0.06	0.13**	0.14**	–0.00	0.04	–0.01	0.05*
Small firm work experience (+)	–0.01	0.11	0.00	0.15**	0.04	0.07	0.06**
Lacks role model (–)	–0.08	–0.04	–0.09	–0.08	–0.16***	–0.17**	–0.11***
Positive model (+)	0.11*	0.04	–0.02	0.09	–0.02	11*	0.05**
R^2	0.53	0.58	0.50	0.46	0.53	0.46	0.51
N	189	170	183	177	170	169	1062

Note: Forced entry of independent variables is used. Standardized regression coefficients are displayed in the table. *$p<0.05$; **$p<0.01$; ***$p<0.001$

a model and story that "works" as sketched above. Regardless of sample, they would all tell stories of "conviction"—or realize that this represents a near tautological relationship that may not be all that interesting, despite all those asterisks. Apart from this, idiosyncrasies of the sample results would lead to very different stories.

With Sample 1, the researcher might build a story around employment status, pointing out that not being committed to an ongoing employment position has a positive effect, but only for those having a temporary job and not among the unemployed. Further, the importance of change-oriented attitudes and beliefs that entrepreneurs make an important societal contribution would be highlighted—and probably offered quite a bit of front-end theoretical rationale in the manuscript. This researcher might also want to include the importance of role models, making the point that it is not the mere presence of a role model that matters but the fact that they have relayed a positive image of entrepreneurship (quite unlike the stories that would emerge from working with samples 5 or 6).

Working from Sample 2, the theorizing would likely center on "know-how," or perhaps "prior knowledge" or "entrepreneurial self-efficacy" in current parlance. This would possibly be added to a "gender" story, unless the latter be seen as old hat. The situational employment variables would probably not be part of the theoretical story; they might or might not be retained as controls. With Sample 4, the big fish allegedly caught by the researcher would be the importance of prior work experience in small firms, while possibly presenting the unexpected result for female sex as evidence that no sex difference exists. Alternatively, this researcher might have determined that the results were too dull and leave the evidence in the file drawer (Landis & Rogelberg, 2013; Rosenthal, 1979), and so on. For each sample, there is a different story. Normally, you would only hear one of them, because the others would not be undertaken or end up published in places where you would not care to look. How worthwhile are really our elaborate theorizing and discussions of practical implications derived from normally sized single studies?

"Wait a minute!," you might say. Looking at the full sample results, it appears that ALL of the variables actually have a "significant" effect in the expected direction. Hence, as far as supportive results are concerned, these researchers would all be telling truths rather than biased lies. The full sample analysis indeed shows statistical significance for all the variables (that's why I retained them—duh!) reminding us that statistical significance has a great deal to do with the size of the sample. But if we try to disregard significance for a while and start at the right end—effect size—what do we *then* see? Do our analyses suggest there is an effect of meaningful size *in the studied sample*? What is a meaningful magnitude of effect should be carefully assessed for each variable (Edwards & Berry, 2010; Schwab et al., 2011), but for convenience, let us assume that we have concluded for each of these variables that a standardized coefficient of 0.10 corresponds to the minimum size of a theoretically and/or practically meaningful effect.[6] Under that assumption, you can see that apart from "conviction," it is only "lacks role model" that has an effect worth writing home about. This also implies that what we are fed with in published research based on small-ish samples risks being exaggerated effects of some factors and negligence of other factors that are in fact equally important, whereas in large

[6] Note that absence of a meaningfully strong effect can also be theoretically and practically important, but then of course it is the small magnitude of the effect that should be highlighted in the reporting of results. You might argue that for some type of variables (operationalizations) it is very hard to say what size an effect needs to be in order to be meaningful. This is true, because we cannot measure everything in easily interpretable units like numbers of dollars or people. But shouldn't you then apply the same logic to the cutoff for what is to be regarded "significant"? In a regression context, the magnitude of unique contribution to R^2 can always be used. Further, you can find ways to give more meaning to results referring to an arbitrary scale. For example, in Davidsson (1995c) I explain: "For example, for 'Need for Autonomy' the difference is 0.38 [$p=0.02$] on a scale with possible values from 4 to 16. All that is needed to obtain such a difference is for 20 of the respondents in Region A [i.e., about 10 % of them] to choose a response alternative one step further towards the 'entrepreneurial' end of the scale on each of the four items in the index than do a set of 20 respondents from Region B, while the average for 'all others' in the two regions is identical." This portrayal of a "significant" difference is quite far removed from conveying the image that "in Region A people in general hold more entrepreneurial attitudes than do people in Region B."

studies we risk being fed with a lot of "statistically significant" effects that the authors fail to say are too small to be theoretically or practically important. Not so good.

Much of the statistical-significance-nonsense discussed above would be avoided if we did what I try to profess in Chap. 4, namely concentrating the interpretation and reporting on "Our Study" (internal conclusion validity) before turning to discussing the extent to which we should believe these results to hold also for other parts of the theoretically relevant population—"The World about Which We Wish to Know and Tell."[7] *Do the relationships seem to be strong enough to be important in that little part of the world that we actually investigated*[8]*?* Only after we have established that we have found something potentially important in "Our Study" is it time to discuss external validity. If experimental: does our experiment sufficiently mimic behavior in the real world? If we have studied an entire, empirical population: is our population part of the theoretically relevant population? If we have studied a nonrandom sample: same thing. Do we have any ground for believing that other parts of the theoretically relevant population are similar to ours? This has nothing to do with statistical inference tools, but it may have something to do with the size and composition of the sample you actually studied. Is it conceivable that your overall results and conclusion would have been radically different, had you studied more cases from various parts of the theoretically relevant population? If you have a random (probability) sample from a well-defined empirical population as well as a high response rate, and if you also stated your hypotheses before the analysis, you can justifiably use significance tests as part of your evidence to support your case for external validity. A conventional test will tell how (un)likely it is to obtain your result, if the true effect in the empirical population you sampled from were zero. This said, we have to realize that sometimes true predictions are not borne out due to random sampling error or insufficient statistical power. In other cases, statistically significant results—theoretically predicted or not—appear in the sample although they are not true for the population from which it was randomly drawn.

The above indicates how little or weak evidence of external validity you can provide in a single study. Other, important parts of the work towards better evidence would have to be addressed on the collective level. Translating our example in

---

[7] Of course, there is a previous round of considering how "our study" relates to "the world…" at the design stage, in setting up our experiment, selecting our cases or interviewees, defining a sampling frame and drawing a sample from it, and in operationalizing theoretical constructs.

[8] If the research is experimental, statistical significance has a role here: is there considerable risk that the results within our sample are due to an unfortunate distribution of participants to experimental conditions so that the supportive results may be spurious? I find the use of statistical testing in an experimental context relatively unproblematic. It is fairly clear what you are going to test before you analyze the data; there is typically no large pool of correlations to potentially over exploit, and there is (I sincerely hope) no fantasizing that "significant" means "true" for an outside population—it merely means the results are unlikely to be wholly attributable to preexisting differences among your experiment participants. But as pointed out by Simmons et al. (2011), there is quite a bit of fishing potential in the experimental pond as well.

Table 9.1 to actual research practice, and assuming that the type of investigated relationships were novel, what would happen is probably that the authors of one of those six potential studies would be faster or better at packaging their results for high tier publication, or be "lucky" enough to get more "interesting" results in their particular study. After they had published their biased study in a high tier outlet, others researching the same relationships would find it difficult to achieve the same, because they are not making a "theoretical contribution" (never mind that their results may call into question the theoretical story told in the original study) and merely undertake "a replication." But if we are not merely fishermen and storytellers and there really is a "World about Which We Wish to Know and Tell," then the above should have convinced you that we really, really need (a) replications and (b) acceptance of "null" findings. Anything less leads to cherry-picking of exaggerated support for the hypotheses put forward in published research. If you do not take my word for it, others have provided evidence of publication bias in entrepreneurship research (Bae, Qian, Miao, & Fiet, 2014; O'Boyle, Rutherford, & Banks, 2014b).

Before closing the book on statistical significance,[9] let me iterate once more how wrong it can lead to (a) hypothesize "different from zero" rather than something more precise and meaningful, (b) apply the same significance criteria regardless of the type of data, and (c) interpret the results only in terms of sign and "significance" of the results, disregarding effect size. Let us first imagine you have hypothesized "A has a positive effect on B" in a study using archival data on the entire population of 500,000 cases existing in a particular country at a particular time. You find an effect in the expected direction at $p<0.01$, one-tailed. Very small effects attain statistical significance when you use such a large "sample." Further, because the data are observational, causality can always be questioned. Therefore, it is entirely possible that rather than reporting "Hypothesis 1 is supported; A has a positive effect on B," what you really should report is: "The observed effect of A on B is positive as expected. However, at least in Country X at time Z, this effect—if at all causal—is so small that it appears to be of little theoretical and practical importance" (I would suggest you don't mention "statistically significant" at all, as it does not apply here. You can be reasonably confident that there is some statistical association, but even more confident that there is no strong, causal effect in the studied population).

Let us now imagine you have hypothesized "A has a positive effect on B" in an experimental study with 20 participants per treatment group. You find an effect in the expected direction, but only at $p<0.07$, one-tailed. With such small groups, it takes a large effect to attain statistical significance at conventional levels. However, if true, an effect of the size observed in your experiment could be a life or death

[9] Don't even try! When I vented my views on statistical significance on Facebook, a colleague–friend replied "This is one of the best justifications I have known to undertake qualitative research. Thanks Per." To which I replied, "Sorry NN (…) The problem with statistical significance (as applied) is a within-paradigm problem. You certainly do not gain any credibility for external validity claims by reducing, per se, the number of cases studied (…) Qualitative (small n) research has its roles, but securing external validity is not one of its strong points."

matter, and the threats to correct attribution of causality are far fewer than when you rely on observational data. Therefore, reporting "our results show that A does not affect B" would be incorrect by any standards and dumb by my standards. Rather, if the effect size appears important, I suggest we report such a finding along the following lines: "Based on our hypothesis test, we cannot exclude the possibility that there is no effect of A on B among our experimental participants" [remember significance has nothing to do with external validity in this case]. "However, the estimated effect is of such magnitude that if a similar effect occurs broadly in real-life contexts it is certainly worthy of policy [or managerial] attention. The result definitely calls for further investigation into this relationship."

After this half-chapter long lead-up, it is time to turn to the real topic of this chapter: the power of replication. The true acid test of theory is that the theoretically postulated effect is demonstrated again and again in empirical samples drawn from theoretically relevant populations. That is, the veracity of a theory is demonstrated through replication. If a theory is any good, it will show its effects in several, slightly different empirical samples and also be robust against variations in operationalizations. I do not think you would want your doctor to give you a risky treatment just because a substance was ascribed a statistically significant effect in a small, single study while serious side-effects came out "marginally nonsignificant." Neither do I think you would like them to deny you treatment that works just because one study was too small for the theoretically predicted effect to reach statistical significance. Similarly, we should not tell students, policy-makers, and business practitioners "truths" that have not been shown to be replicable. If they do not stand the test of academic replication, our recipes are unlikely to work in practical application, either. In the remainder of this chapter, I will use examples from my own work to demonstrate how various forms or replication can boost one's justified confidence in a theoretically proposed relationship.

9.2 Replicating Others

What first comes to mind when you think about replication is to copy what somebody else has done in already published work in order to either confirm or question the findings. This is what my first two examples are about. However, my examples are not pure replications because the studies were not designed with that main purpose, and hence the operationalizations differ, as do the sampled populations. This decreases the value of the replication as test of internal validity while it actually increases its value as a test of external validity (Hubbard et al., 1998). That is, the theory test is a tougher one because if the support for the theory in the original study was to some part due to an artifact of the specific sample or operationalizations, this "benefit" does not carry over to the replication.

My first example was part of my dissertation study (Davidsson, 1989a) and my very first conference presentation outside of Europe (Davidsson, 1988). The starting point for this research was that I realized when reading Smith's (1967) then oft-cited study that I had data in my study to test many of his propositions about types of entrepreneurs and the firms they create (yes, Fauchart and Gruber, 2011, were not the first to suggest this type of idea—but they do it well!). More specifically, Smith made claims that relative to craftsman entrepreneurs, opportunistic entrepreneurs have the following characteristics, for which I also had data available (cf. Davidsson, 1989a, pp. 143–144 and 159–160):

– They are more likely to have run (an)other firm(s) prior to the sampled one.
– They are more likely to currently run more than one firm.
– They are more likely to have management experience prior to starting their own firm.
– They are less likely to have considerable experience from the specific industry prior to becoming CEO of the sampled firm.
– They have higher average level of general education.
– They have higher average level of business education.
– They have more internal locus-of-control.
– They have higher need for achievement.
– They have more self-confidence.
– They are less concerned that their firm might become overly dependent on a small number of customers, suppliers, or investors.
– Personal control and surveillance of the firm's activities are relatively less important to them.
– Ownership control is relatively less important to them.
– They find recruiting easier.
– They are less likely to have their spouse employed in the firm.
– They have more positive attitudes towards growth.

Further, Smith (1967) postulated that opportunistic entrepreneurs create adaptive firms, implying the following testable (in my study) characteristics relative to the rigid firms that craftsman entrepreneurs create.

– The firm is less likely to be wholly owned by the respondent.
– A smaller share of the firm's sales is generated within the home county.
– The share of the firm's sales that is generated on export markets is more likely to be above industry average.
– The share of the firm's sales that is generated by products developed "in house" is higher.
– They are currently more likely to be involved in product development.
– The firm has higher historical growth rate.
– The growth aspirations for the future are also higher.

This list of characteristics can be regarded a complex hypothesis saying that if we divided the sample into the two most homogeneous groups we can find, these two groups would be split in accordance with the 22 statements above. So I applied cluster analysis in a confirmative fashion, and the results were largely supportive. It turned out that in a two-group, hierarchical cluster analysis 20 of the 22 differences were in the predicted direction. For all the firm variables, the differences were substantial, which was also the case for the educational and psychological characteristics of the individuals. The differences were in the right direction but rather unimpressive for habitual entrepreneurship and the experience variables. The two instances of differences in the "wrong" direction—ownership and supervisory control—were also very small and best regarded as indicating "no important difference."

Smith derived his taxonomy from in-depth study of a relatively small (52 cases), all-male, all-manufacturing sample in Michigan in the 1960s. Nonetheless, my test on a much larger mixed-sex, mixed-industry sample in Sweden 20 years later suggested his taxonomy was a meaningful way to distinguish conceptually and empirically between groups of business owner-managers. My conclusion was that:

> Despite temporal, cultural, sampling and operationalization differences, the groups that emerge in a cluster analysis of this new sample show considerable resemblance to the entrepreneurial groups suggested by Smith (1967). This result provides fairly strong support for the usefulness of his typology.[10] (Davidsson, 1989a, p. 155)

This example illustrates the mutually beneficial nature of replications. Smith's theory gains credibility and generalizability by being replicated in a different empirical context. My study becomes a much more meaningful contribution by being framed as a test of Smiths existing taxonomy rather than as a stand-alone, exploratory attempt to find distinct subgroups among business founders.[11] Replications refine our knowledge either by supporting or questioning established "truths."

Incidentally, my second example of "replicating others" also involves work by Arnold Cooper and Carolyn Woo. Based on human capital theory as well as previous empirical research on entrepreneurial performance, Cooper, Gimeno-Gascon, and Woo (1994) derived ten hypotheses about how initial conditions influence new venture performance. The hypotheses predicted effects of four broad categories of initial capital: *general human capital, management know-how, industry-specific know-how*, and *financial capital.* They tested their hypotheses on a large, longitudinal data set representing new ventures across all US industries and regions in the mid-1980s. They got very limited support for the effect of management know-how; otherwise the results largely supported their hypotheses.

We realized that with *The 1994 Start-up Cohort Study* we had access to a similarly composed but even larger, Swedish sample from the mid-1990s (Dahlqvist, Davidsson, & Wiklund, 2000). Although we did not always have access to the exact

[10] Sic! It should be "taxonomy."

[11] Like I said in a dissertation footnote: "Typologies [sic; see the above note!] arriving at 2–11 groups, on the basis of different approaches and with more or less of systematic empirical backing, may be found in: [nine references]. Conclusion: adding another one would be superfluous" (Davidsson, 1989a, p. 158). However, see Woo, Cooper, and Dunkelberg (1991)—a paper which made the popularity of entrepreneurial taxonomies and typologies plummet—for evidence of instability of Smith's types.

same operationalizations, we did have indicators for all four groups of predictors that Cooper et al. (1994) used, and we could also add a fifth category: *access to markets and resources*. We were also able to apply their use of three outcome categories (failure, marginal survival, and growth) although we used "high performance" rather than "growth" for the best performing group (and today we would use "exit" or "discontinuation" rather than "failure"). Our results revealed the following *similarities:*

– In both studies, indicators of general human capital contributed positively to marginal survival and high performance.
– In both studies, indicators of management know-how contributed positively to marginal survival.
– In both studies, indicators of financial capital contributed positively to high performance.
– In both studies, ventures in retailing and personal services had lower probabilities of marginal survival and high performance.

The following *differences* between the studies stand out relatively clearly:

– While our model was much stronger in predicting high performance than marginal survival, their results appear more balanced in this regard.
– While they found effects of industry-specific knowledge on both survival and high growth, our analysis confirmed neither of these effects.
– Our study also lacked the following effects obtained by Cooper et al. (1994): (a) a positive effect of management know-how on high performance and (b) a positive effect of financial capital on marginal survival.

Again, the replication adds value to both studies. For example, without our follow-up readers of the Cooper et al. article can choose rather freely to interpret lack of a specific effect as a real lack of such an effect or as a shortcoming of their operationalizations. It becomes more difficult to argue that way when our results point in the same direction. Likewise, both studies arrive in some instances at the same rather subtle differential influence on marginal survival and high performance, respectively. This is the case with the female sex effect, where according to both studies ventures run by women show a lower probability of high performance, but not a higher probability of discontinuation (cf. Zolin, Stuetzer, & Watson, 2013). Further, Cooper et al. (1994) found that presence of a parental role model (vicarious learning) increases the probability of marginal survival but not of high performance. In a similar fashion, we found that previous start-up experience (experiential learning) is positively associated with survival but not with high performance. When such patterns are repeated across several studies, they achieve a much higher level of credibility than when they are afforded a couple of asterisks in a single study.

9.3 Replicating One Another: Harmonized Research Collaboration

Another type of replication—whether or not it is thought of and presented as such—is when researchers collaborate on conducting several parallel studies aimed at addressing the same research questions. The Global Entrepreneurship Monitor (GEM) may be regarded a giant example of this (Álvarez, Urbano, & Amorós, 2014; Amorós, Bosma, & Levie, 2013). Well before the dawn of GEM, one of my most gratifying professional experiences ever was when as a young researcher I got involved in a seven country international collaborative project on the regional determinants of firm start-up rates, under the competent leadership of Paul Reynolds and David Storey. Sure, we had our differences within the group, and cats from some countries turned out to be less easily herded than others. Overall, however, our meetings were joyful and rewarding because everybody was working on the same problem and therefore up to speed with the relevant theoretical and methodological issues at hand. Thus, unlike some sessions at broader meetings, the discussion could start on a very high level—and climb from there.

In short, we collaborated in the design phase in order to harmonize the data collection (or, rather, compilation of archival data) as far as possible. In many cases, it proved impossible to get exactly the same indicators, but then we tried to ensure that each country study included at least some indicator(s) of the following regional characteristics: *demand growth, urbanization/agglomeration, unemployment, personal/household wealth, small firms/economic specialization, political ethos,* and *government spending/policies.* We then related these characteristics to subsequent firm start-up rates in manufacturing only as well as across all sectors, and relative to the size of the firm population as well as relative to the size of the workforce. The country teams could conduct and present whatever analyses they liked, wherever they liked, but the main country reports as well as harmonized comparative analyses were published jointly in a special issue of *Regional Studies* (Vol. 28; No. 4, 1994).

Those who care to read the individual country reports may get a rather confused picture of what determines the regional rate of business start-ups. Different studies used different indicators and did not necessarily present them as indicators of higher-order theoretical constructs (cf. Chap. 3). The French study (Guesnier, 1994) reported unstandardized regression coefficients and positive effects of unemployment rate and small firm density. The German study (Audretsch & Fritsch, 1994) reported standardized regression coefficients and negative effects of seemingly the same variables. The Irish study (Hart & Gudgin, 1994) included only the manufacturing sector. The Italian (Garofoli, 1994) and Swedish (Davidsson, Lindmark, & Olofsson, 1994) studies reported results for start-ups relative to the size of the workforce only, stubbornly refusing to relate the number of start-ups also to the number

Table 9.2 Summary results for regional determinants of firm start-up rates (cf. Reynolds et al., 1994)

Regional determinant	All sectors	Manufacturing only
1. Demand growth	Positive (6)	Positive (6)
2. Urbanization/agglomeration	Positive (6)	Positive (5)
3. Unemployment	Positive (4)	Mixed (5)
4. Personal/household wealth	Positive (3)	None (4)
5. Small firms/specialization	Positive (6)	Positive (7)
6. Political ethos	Positive (2)	Positive (2)
7. Government spending/policies	None (4)	Positive (1)

Note: The numbers indicate the number of countries (out of seven) where one or more indicators of the process could be included

of organizations in the economy.[12] The British study (Keeble & Walker, 1994) included predictive models also for small firm growth and death, which no other study matched.

However, the harmonized analysis in the final article (Reynolds, Storey, & Westhead, 1994) nevertheless arrived at powerful, generalizable conclusions. Once the analyses were harmonized and the proper level of abstraction applied, the authors were able to convert the above mess to the summary in Table 9.2. In verbal terms, they summarized the findings as follows:

> Analysis of the processes associated with new firm births across seven advanced market economies…indicates three processes having a positive impact on firm birth rates:
>
> growth in demand, indicated by population growth and growth in income
>
> a population of business organizations dominated by small firms
>
> a dense, urbanized context, reflecting the advantages of agglomeration (…)
>
> Other processes—related to unemployment, personal wealth, liberal political climate or government actions—had weak or mixed impact. (Reynolds et al., 1994, p. 453)

At the time, there were few if any conclusions about entrepreneurship that had as solid empirical backing as these. It would not be possible to achieve the same in a single study, no matter how comprehensive and well designed. Replication rules! Research on regional variations in entrepreneurship has since moved on in terms of methods, theories, and results, but I do not think one can say that the above conclusions have been broadly refuted (Ács & Armington, 2006; Audretsch, Dohse, & Niebuhr, 2010; Boschma & Fritsch, 2009; Bosma & Schutjens, 2011; Fritsch & Storey, 2014; Plummer, 2010).

[12] We also stubbornly (valiantly?) refused to report significance tests for analyses of our population data: "It should be noted that the study covers the whole population of establishments and regions. Statistical significance thus is a non-issue and such tests are therefore not reported" (Davidsson et al., 1994, p. 397). This collaborative effort provides a good example of how stupid it would be to use statistical significance to assess the observed effects. The US study (Reynolds, 1994) included 15 times as many regions (380-ish) as did the Irish study (23-ish), meaning that if *exactly the same* coefficients were obtained for the two countries, they would likely be judged "significant" for the US but not for Ireland. Time to turn the brain back on; the comparison is between empirical *facts* for two different countries, not statistically uncertain estimates.

Convinced yet? Ready for another one? My second case of "replicating one another" exemplifies temporal rather than spatial replication. As part of my doctoral dissertation project, I developed a package of questions concerning small firm owner-managers' expected consequences of growth. Again, I saw growth as an entrepreneurship issue at the time; subsequently I have refined that view (cf. Davidsson et al., 2002 and Chap. 1 of this book). I used my pilot study and the extant literature to find aspects of small firm owner-managers' work environment that were important for their "job satisfaction," and which could be suspected to be improved or worsened as a consequence of expansion. I came up with the following dimensions (Davidsson, 1989b):

- *Workload*—would the owner-manager have to work more or less if the firm were twice as big?
- *Work tasks*—would the owner-manager get to spend a smaller or larger share of his/her time doing the most preferred work tasks?
- *Employee well-being*—would the firm be a better or worse place of work in the eyes of the employees?
- *Personal income*—would the owner-manager make more or less money, were the firm twice as big?
- *Control*—would it be easier or more difficult for the owner-manager to survey and control the operations of the firm?
- *Independence*—would the owner-manager enjoy a greater or lesser feeling of independence relative to important external stakeholders?
- *Vulnerability*—would it be easier or more difficult for the larger firm to survive a severe crisis?
- *Product/service quality*—would it be easier or more difficult to keep up high quality of the firm's products and/or services?

Albeit a side issue in all three projects, the same package of questions were reused in Frédéric Delmar's and Johan Wiklund's respective dissertation projects in the mid- to late 1990s (Delmar, 1996; Wiklund, 1998). In Wiklund, Davidsson, and Delmar (2003), we finally wrapped up the three studies, using an expectancy theory lens in relating the owner-managers' expected consequences of growth to their overall growth attitude. That is, we tried to answer the question "What specific expectations determine small business owner-managers' general positive–negative inclination towards growing their firms?" Tables 9.3 and 9.4 summarize the results.

From Table 9.3, we learn that in each study expectations concerning the effect of growth on *employee well-being* comes out as the most important determinant of growth attitude. Had this result appeared in a single study skeptics could have regarded it a peculiarity of little consequence. When replicated in three studies, the suggestion that this nonfinancial concern may be more important than financial ones (*personal income*) in determining overall growth attitude has to be taken seriously. From Table 9.3 we also learn an important lesson regarding "relative importance" of not-very-strong predictors. Other than the dominance of employee well-being, the estimated relative importance of different expectations appears to be quite

Table 9.3 Effects of expected consequences on growth attitude in three studies (cf. Wiklund et al., 2003)

Sample variable	1986 sample $n=287$	Rank order	1994 sample $n=338$	Rank order	1996 sample $n=533$	Rank order	Joint probability
Workload	0.11*	2	0.04	7	0.02	7	0.0015
Work tasks	0.04	7	0.15**	2	0.00	8	0.0003
Empl. well-being	0.27***	1	0.19***	1	0.25***	1	>0.000001
Personal income	0.07	4	0.08	5	0.12**	4	0.000007
Control	0.10*	3	0.00	8	0.13**	2	0.00003
Independence	0.07	4	0.11*	3	0.13**	2	0.000004
Vulnerability	0.07	4	0.11*	3	0.06	5	0.0002
Quality	0.04	7	0.08	5	0.03	6	0.04
Adj. R^2	0.23		0.20		0.23		

Note: Forced entry of independent variables is used. Standardized regression coefficients are displayed in the Table. *$p<0.05$; **$p<0.01$; ***$p<0.001$; one-tailed (see footnote 2 above)

unstable (cf. Table 9.1). In a single study, researchers and readers alike are often lured to over interpret small differences like these.

Moreover, the table reiterates important lessons about statistical significance. We can clearly see the non-replicability (as "significant") of results that are significant on the 5 % risk level in at least one analysis in these fairly sized samples. However, the lesson is not that we should necessarily disregard these effects. Instead, the results across studies illuminate how stupid it is to regard lack of significant result as *proof* of nonexistence of an effect, i.e., as solid evidence against the theory. We would be in serious error if we concluded from either of the first two studies that "Small firm owner-managers willingness to grow their firms has nothing to do with the expected effects of growth on their personal income." As we can see from the last column of the table, the likelihood of finding positive effects this large or larger in three separate studies of this size is actually less than seven in a million. Quite impressive for a "nonexisting" relationship, wouldn't you say? See now that we need replications?

In Table 9.4, the data from the three studies have been pooled in order to make possible breakdown analyses on large enough subsamples. This is another form of replication of results that bridge over to the next section about "Replicating Yourself." After seeing these results, our confidence in the general importance of employee well-being concerns should be further enhanced. Moreover, we can say with some confidence that no important industry differences seem to exist regarding the influence of expected consequences of growth on overall growth attitude, other than perhaps a somewhat lesser suitability of the model for explaining variance in the retailing industry (lower R^2). We do not see in the table any dramatic subsample differences for size or age, either. However, given the rather sizable subsamples, we should perhaps dare to conclude that *independence* comes to the fore only as the firm has already grown out of the smallest size class. Conversely, concerns about

Table 9.4 Effects in various pooled subsamples (cf. Wiklund et al., 2003)

Sub-sample variable	Manuf. $n=571$	Service $n=340$	Retail $n=246$	5–9 emp $n=326$	10–19 emp $n=479$	20–49 emp $n=353$	Old firms $n=771$	Young firms $n=372$
Workload	0.07*	0.08	0.00	0.07	0.08*	–0.01	0.07*	0.07
Work tasks	0.04	0.06	0.02	0.13**	0.05	–0.05	0.01	0.10*
Employee well-being	0.23***	0.27***	0.23***	0.30***	0.17***	0.29***	0.28***	0.22**
Personal income	0.10**	0.10*	0.10*	0.11*	0.06	0.13**	0.12***	0.05
Control	0.08*	0.10*	0.08	0.07	0.12**	0.08	0.10**	0.04
Independence	0.11**	0.09*	0.14*	0.02	0.15***	0.13**	0.10**	0.13**
Vulnerability	0.09*	0.10*	0.04	0.07	0.07	0.13**	0.09**	0.06
Quality	0.06	0.04	0.09	0.06	0.09*	0.00	0.03	0.11*
Adj. R^2	0.22	0.28	0.16	0.26	0.24	0.21	0.25	0.20

Note: Standardized regression coefficients are displayed in the table. $^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$; one-tailed (see footnote 2 above) neither *workload* nor *work tasks* seem to be a factor at all for the "largest" firms. This explains the modest overall results for these dimensions.

I hope the above examples have demonstrated that replicating one another and publishing the results jointly is not a bad idea. It is actually a pretty good one.

9.4 Replicating Yourself

Regrettably, a straight replication of somebody else's work is difficult to get published in a highly ranked journal. If you ask me, this shouldn't be the case—the outlets that published the original work should also make room for succinct replication manuscripts that support or question it. But the paucity of replication research is reality; in a quick search, I could only find three explicit replications in entrepreneurship published since 2010, namely Frank, Kessler, and Fink (2010); Obschonka, Andersson, Silbereisen, and Sverke (2013); and Weismeier-Sammer (2011). To these authors, I say "Good on ya, mates!"[13]

Apart from exaggerated faith in the truth value of significance tests, I suspect—as mentioned in a previous chapter—that one reason for the low regard for replication studies is that colleagues with editorial powers to some extent confuse two things: (a) yet another study on an "old" topic, which does not build sufficiently on theoretical, methodological, and empirical insights gained in previous research on the same topic, and (b) an explicit attempt to replicate previous research and retest the ideas it puts forward. The former type is of questionable value whereas the second type is an absolutely essential ingredient in a field of research that is serious about developing reliable and useful knowledge.

I also suspect editors think that replication research is less interesting and therefore will be less cited, thereby potentially hurting the holy journal impact factor. If so, I believe they are plain wrong. It is typically influential work that gets replicated, and the replications therefore will get many "free-rider" citations jointly with the original articles. The four articles cited above seem to do *much* better than the average article in the same journals and years. Our abovementioned replication of Cooper et al. (1994) is currently the fourth most cited article ever in the (admittedly short-lived) journal in which it was published. Replication rules!

Regrettably again, you don't always have friends who want to play your game and harmonize their research with yours. So what to do when editors don't appreciate straight replications as they should; when you can't get colleagues to conduct parallel studies, and this annoying Davidsson guy claims single studies aren't worth that much, anyway? Replicate yourself! That is, make your study large enough so that you can prove your propositions on several, separate data sets or subsamples. The following examples are intended to show how this can dramatically increase the value and credibility of individual articles.

[13] Honig and Samuelsson (2014) is an extended reanalysis rather than a replication on new data.

Table 9.5 Performance transition matrix; Australian data

	1994/1995 group						
		Poor ($n=620$)	Middle ($n=1018$)	Growth ($n=654$)	Profit ($n=538$)	Star ($n=816$)	Total ($n=3646$)
	Exit	33.9	17.0	25.7	28.8	20.3	23.9
	Poor	21.3	14.5	23.2	10.6	10.3	15.7
1997/1998 group	Middle	16.8	36.8	17.3	14.7	19.9	22.8
	Growth	13.7	9.1	19.3	4.1	7.8	10.7
	Profit	5.3	11.8	5.5	25.1	20.0	13.4
	Star	9.0	10.7	**9.0**	**16.7**	21.7	13.5
	Total	100.0	100.0	100.0	100.0	100.0	100.0

In 2005, my coauthors and I won a best paper award at the Academy of Management for an article with what some find to be a controversial message. We asked which firms end up in the privileged position of being above-average performers both in terms of growth and profitability. Do firms sacrifice profitability in order to grow and then become highly profitable as a result of their larger size? Or do they fix the profit problem first and then scale up while maintaining their above-average profitability? We approached this question by assigning firms according to their relative position compared to the industry median on the two performance dimensions: “Poor” = low on both; “Star” = high on both; “Profit” = high on profitability but low on growth; “Growth” = high on growth but low on profitability. To avoid large effects of small changes, we also defined a “Middle” category for those that did not deviate markedly from the industry median on either dimension. Table 9.5 reiterates the main results as reported in the conference paper.

As portrayed in the table, the difference may not look that large, but it does in fact show that the Profit→Star transition is 85 % more common than the Growth→Star transition. Although not highlighted in the table, a closer examination reveals that Growth firms frequently neither sustain their growth nor become highly profitable as a result of it, but instead become underperformers on both dimensions (Poor). In the main, then, the results suggest that firms do not become more profitable because they grow larger. Instead they can grow profitably because they have already proven an ability to achieve above-average profits at a smaller scale.

This made great sense to me all along: if your offer is not that special, you won’t enjoy high margins. If you nevertheless want to grow, you need to spend more on marketing and/or sell at a lower price in order for customers to prefer your product over others’. Unless growing will make you enjoy major cost advantages of scale, this would lower rather than increase your profitability as you grow.

With clear results and an award under our belts, we just had to submit it to *Journal of Fantabulous Research*, reluctantly accept their letter of acceptance (after considering asking them for a suitable publication fee), and earn eternal glory, right? Not quite. It took a few years and journals for Davidsson et al. (2009) to see the light of day. Partly this was because it took some work for us to convince others *and ourselves* that we really had got the results and interpretations right. Because the results were seen as controversial, it is fair enough to demand that we strengthened the evidence. Here is what we did:

1. We performed the same analyses on another data set from a different country.
2. We reran the analyses for shorter and longer transition periods.
3. We reran the analyses for various subgroups in the data.
4. We also improved our theoretical story by focusing on a resource-based, view-based line of argumentation.

As shown in Table 9.6, the results hold up in terms of the direction of the difference across most of our internal replications, and often showing that Profit → Star is twice as likely (or more) as Growth → Star. That many of the tests come out as "statistically significant" in isolation isn't really the main point; to get 50 out of 52 contrasts run in the same direction is improbable beyond belief in statistical terms. The repeated evidence makes the main finding much more solid and generalizable and therefore harder to dismiss. The results for transition to Poor resemble the overall results in Table 9.5 and are thus the mirror image of those reported for the Growth → Star transition in Table 9.6.

As a second example of "replicating yourself," consider Erik Hunter's dissertation work (Hunter, 2009; Hunter & Davidsson, 2007; Hunter, Burgers, & Davidsson, 2009). Erik identified the growing phenomenon of "celebrity entrepreneurship"—that celebrities increasingly appear not just as paid product endorsers but are (apparently) involved in roles such as initiators, founders, owners, product designers, and strategic advisors of the companies they endorse. Erik built his thesis in the borderland of marketing and entrepreneurship around this neglected phenomenon. His core idea was that the "celebrity effect" would be stronger if the celebrity was (or was portrayed as) entrepreneurially involved in the venture rather than just endorsing its products for a fixed dollar fee.

After some initial work, we had a growing realization that this effect was not mainly transmitted via the traditional endorser qualities *attractiveness, trustworthiness,* and *expertise* (Ohanian, 1990). Although there likely was some effect on perceptions of these, it seemed a new endorser characteristic, (perceived) *emotional involvement,* needed to be considered. So Erik developed a scale to measure this characteristic and ran a series of experiments where he manipulated the celebrity's role as either entrepreneurially engaged or merely being a paid endorser, to test the following three main hypotheses[14]:

H1: Emotional Involvement is a conceptually and empirically distinct characteristic of communicators relative to the traditional characteristics trustworthiness, attractiveness, and expertise.

[14] For ease of communication, the numbering of experiments and hypotheses as well as the exact wording of the hypotheses has been adapted for this example but stay true to the essence of the underlying research.

Table 9.6 Transitions to "Star" by different performance groups of origin, countries, and subgroups (cf. Davidsson et al., 2009)

	Australia						Sweden					
	1-year transitions			3-year transition			1-year transitions			2-year transition		
	Initial year performance group			Initial year performance group			Initial year performance group			Initial year performance group		
Subgroup	Growth[a]		Profit[a,b]	Growth[a]		Profit[a,b]	Growth[a]		Profit[a,b]	Growth[a]		Profit[a,b]
Manufacturing	9.8	***	25.2	10.7	n.s.	13.4	15.0	***	41.1	8.3	***	44.6
Prop./bus. serv.	10.6	***	28.0	8.7	**	22.6	13.7	**	26.0	15.5	n.s.	25.8
Retail[c]	11.7	**	23.1	13.3	n.s.	9.8	13.5	n.s.	21.1	8.3	n.s.	23.3
Wholesale	12.0	***	23.8	8.0	*	17.2		N/A			N/A	
Other	11.7	***	28.9	10.7	*	19.6	19.1	n.s.	28.0	11.9	**	34.0
Size 1[d]	14.6	***	31.4	9.4	**	19.6	13.0	*	26.0	5.7	**	27.1
Size 2	14.0	***	29.4	13.6	n.s.	11.9	18.4	*	30.0	16.7	**	35.4
Size 3	11.3	***	30.2	8.3	*	18.5	16.8	**	30.3	13.7	**	33.3
Size 4	8.9	***	27.5	9.9	*	16.2	10.6	***	34.6	4.3	***	35.7
<2 years[e]	13.3	*	24.0	6.5	*	21.7		N/A			N/A	
2–5 years	9.2	***	36.4	9.1	*	18.1	20.4	n.s.	26.8	14.8	n.s.	19.0
5–10 years	13.8	***	28.8	14.5	n.s.	17.1	16.7	**	30.2	9.8	***	35.7
10–20 years	13.9	***	28.7	6.8	**	16.6	16.0	*	27.6	18.0	*	33.3
>20 Years	7.7	***	30.1	11.5	n.s.	13.2	13.3	***	31.6	8.1	***	35.1

Note: The table was compiled by Paul Steffens and has only been lightly edited

[a]Table entries are percentages of specified initial performance group that transition to Star

[b]Significance levels: *$p<0.05$; **$p<0.01$; ***$p<0.001$ (this application of statistical significance testing is subject to some of the issues raised in this chapter, but as you know, many reviewers and editors are starstruck)

[c]Retail and Wholesale combined for Swedish sample

[d]Size classes for Australian sample based on revenue in AUD: size 1 <\$300 K; size 2 \$300 K-\$1 M; size 3 \$1 M-\$3 M; size 4 >\$3 M. Size classes for Swedish sample based on revenue in SEK: size 1 <15 M; size 2 15 M-\$50 M; size 3 \$50 M-\$150 M; size 4 >\$150 M

[e]The Swedish results are not reported as the sample included only 12 firms less than 2 years old

Table 9.7 Summary results for celebrity entrepreneurship experiments

Experiment	Product	Celebrity	Participants	Country	Result		
					H1	H2	H3
1	Surfing gear	Female movie star	Students	Sweden	S	S	S
2	Snowboarding gear	Female movie star	Students	Sweden	S	S	S
3	Surfing gear	Female movie star	Students	Latvia	S	S	S
4	Fast food chain	Male eating champion	Students	Sweden	S	S	S
5	Fast food chain	Male eating champion	Students	Sweden	S	R	S
6	Food supplement	Male sports star/ TV host	Retirees	Sweden	S	S	R

H2: Greater perceived emotional involvement will lead to higher communication effectiveness.

H3: When the celebrity is entrepreneurially engaged (i.e., part founder-owner) in the venture rather than being a paid endorser, this has a positive effect on the audience's perception of the celebrity's emotional involvement.

Of course, nothing about initiating or owning was included in the emotional involvement items; these centered on perceptions that the celebrity genuinely liked and used the product. Communication effectiveness was measured with traditional marketing variables "attitude towards the brand" and "attitude towards the ad."

Rather than just running one, big experiment to test these hypotheses, Erik conducted a series of six experiments. Across these experiments, he varied the celebrity, the product, the type of participants, and the country where the experiment took place (meaning language was also varied). There is very good replication logic behind this. If you get strong support in one, big experiment, you cannot exclude the possibility that somehow the results were driven by idiosyncrasies of the particular design and setting, no matter how long a row of significance asterisks you try to impress us with. As you can see in Table 9.7, Erik obtained fairly consistent support for his hypotheses across varying conditions (S denotes support whereas R means the hypothesis was rejected). This makes a much stronger case for the generalizability of the effects.

Interestingly, the effects hold up also with the hot dog eating champion Takeru Kobayashi as the celebrity—a person whose "celebrity" status was not even known to most participants in advance, and whose claim to celebrity would not necessarily trigger unambiguously positive feelings in every participant.[15] Further, snowboarding gear is arguably a more relevant product category to these Northern European participants than is surfing gear, but this did not make a difference to the results.

[15] One examiner suggested the analyses should be broken down by the sex of the participants. Seems a reasonable suggestion….

The results also include one interesting case of nonsupport. The stronger effect on perceived emotional involvement when the celebrity is entrepreneurially engaged did not come through with Swedish retirees. A plausible explanation is that this is a cohort effect. Swedes of that generation were politically more leftist oriented and therefore less prone to have a heroic view of entrepreneurs. They may perceive the difference as one of people trying to profit as much as they can (bad thing!), this way or that way, and they remain unimpressed either way. By contrast, the nonsupport of H2 in Experiment 5 seems more like happenstance in the context of the overall results.

I take the final example of "replicating yourself" from a methods piece, and the internal replications therefore largely coincide with the validation exercises I exemplified in Chap. 6. However, since developing a new operationalization was the core contribution in a piece published in the *Strategic Management* Journal, I think it (Brown, Davidsson, & Wiklund, 2001) has earned its place here.

The rationale behind the research, which was originally Terrence Brown's idea, was that despite its popularity at the time, Howard Stevenson's conceptualization of entrepreneurship or "Entrepreneurial Management" (Stevenson, 1984; Stevenson & Gumpert, 1991; Stevenson & Jarillo, 1990) had never been systematically tested. In order to make the conceptualization testable, someone had to first operationalize Stevenson's dimensions of entrepreneurial management. Terrence and I took on the challenge when designing the project *Entrepreneurship in Different Organizational Contexts*. Johan Wiklund joined us as coauthor at a later stage and did a great job not least on validation issues, which relates to our current emphasis. Table 9.8 displays our (i.e., Brown's et al., 2001) main results.

We were pleased that it turned out possible after some addition and deletion of items to arrive at a solution where factors that accorded with Stevenson's explicit and implicit dimensions of entrepreneurial management came out quite clearly. In most cases, the corresponding computed indices also show a satisfactory degree of internal consistency. To be honest, we were also somewhat surprised; as all of the dimensions aim at capturing some aspect of corporate entrepreneurship, they should not necessarily be expected to come out this clearly in an orthogonally rotated factor analysis. But they did. So far so good. But were our results stable or had we just been lucky and/or used stochastic variation to the max when we dropped and added items until we arrived at this clean factor structure?

Again, we could use the fact that we had a large and stratified sample. As internal replication, we reran the factor analysis for different strata for a total of ten subsample analyses. Displaying factor loadings from ten separate analyses is a bit over the limit, so in Table 9.9 I have summarized the essence of the results.

By and large, the results of the subsample analyses were very encouraging. The cumulative variance explained is very similar in every analysis; the right number of factors is extracted in all analyses but one, and—importantly—the extracted factors remain the same across subsamples. There are some problems with the *resource orientation* factor, but otherwise our operationalization appears successful at least from a technical point of view. With the internal replication on different groups of firms, I would argue we made a much, much stronger case for the validity of this

Table 9.8 Factor analysis results for Stevenson's conceptualization of entrepreneurial management (cf. Brown et al., 2001)

Factor variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Strategic orientation 1	0.79					
Strategic orientation 2	0.85					
Strategic orientation 3	0.82					
Resource orientation 1		0.56				
Resource orientation 2		0.80				
Resource orientation 5		0.72				
Resource orientation 6		0.49				
Management structure 1			0.75			
Management structure 2			0.80			
Management structure 3			0.68			
Management structure 4			0.67			
Management structure 5			0.65			
Reward philosophy 1				0.74		
Reward philosophy 2				0.66		
Reward philosophy 3				0.73		
Growth orientation 1					0.84	
Growth orientation 2					0.86	
Entrepreneurial culture 1						0.82
Entrepreneurial culture 2						0.66
Entrepreneurial culture 3						0.84

Note: $n = 1233$. Absolute values less than 0.30 have been suppressed

operationalization than if we had only had access to, say, the 399 independent firms or the 372 manufacturers.

Apart from Hunter's dissertation work, the above examples of "replicating yourself" consisted mainly of performing subgroup analyses "post hoc." Although this is valuable, it means that all of the analyses are subject to the same journey of gradual adaptations and ad hoc choices that typically occur to a greater or lesser extent between data collection and published article. In this sense, a better strategy of "internal replication" may be to explicitly allow the element of exploration using only half the data (tryout sample or learning sample) and then test the model on the other half of the data (holdout sample). Adding a second data set, like we did in Davidsson et al. (2009), largely serves the same function; the outcomes of the new analysis is not driven by patterns in the data at hand that are already known to the analyst.[16] Those who use survey data may feel they cannot "afford" not to use all the

[16] Actually, adding a separate second sample is even better, because then neither the data source nor the process of developing the model are subject to stochastic idiosyncrasies that are capitalized on in model fitting. In addition, a trusted colleague questions the logic of tryout/holdout altogether on the grounds that if split randomly, the only source of difference in results is stochastic variation. I would counterargue that the strategy mitigates the tendency to step-by-step make adaptation that

Table 9.9 Stability of factor analysis results across sampling strata

Subsample	No. of cases	No. of factors w. Eigenv. >1	Cum. var. explained. by 6 factors	No. of incorrect loadings Type I[a]	Type II[b]
Governance					
Independent	399	6	61.1 %	0	1
Part of "small" group (<250 empl.)	446	6	61.4 %	0	2
Part of large group (250+ empl.)	433	6	59.6 %	0	1
Size					
10–49 employees	655	6	59.8 %	0	1
50–249 employees	623	6	61.2 %	0	0
Industry sector					
Manufacturing	372	6	61.3 %	0	1
Prof. services	366	6	61.5 %	0	3
Retail/wholesale	226	6	60.2 %	1	4
Other services	314	7[c]	60.2 %	0	1

[a]Number of occurrences that the highest loading is on the "wrong" factor
[b]Number of "side-loadings" > 0.30
[c]The resource orientation dimension split into two factors

data available in their main analysis, but this can of course be solved by doing fewer and larger studies for the same monetary cost. Those who work with large, archival data sets can usually well afford to develop their model in a semi-exploratory fashion and test it on untouched data. Since significance testing does not apply to population studies, this is a much better way to get a sense of the robustness and true predictive ability of the estimated model.

9.5 Some Encouraging Signs

I have been whingeing above about misapplication of and overreliance on significance testing as well as underappreciation of replication studies, and I don't think these ills will vanish any time soon. This said, I also see a lot of indicators that we are making progress; we are moving in the direction of collectively building more solid support for our theoretical ideas about entrepreneurial phenomena. Here are some of the positive signs:

The increased emphasis on context in management and entrepreneurship research (G. George, 2014; Johns, 2006; Welter, 2011; Zahra, 2007; Zahra & Wright, 2011). This trend mitigates overreliance on the "statistically significant" evidence provided for some group of entrepreneurial ventures or agents at some place at some

capitalize on stochastic relationships in the data, but agree that a separate, second sample that did not affect model development is even better.

point in time. As long as authors can develop theoretical insights into the reasons for context-driven differences, journals will be interested, and we will all benefit from refined understanding of the boundary conditions of theoretical propositions. The remaining problem is that we also need confirmatory evidence on replicability of initial results *within the same and similar* contexts (Hubbard & Lindsay, 2013a). However, in the efforts to find differences, we will also find sameness, namely those effects that do not seem to vary strongly by context.

The general quality increase in entrepreneurship research (Crook et al., 2010) reduces all kinds of threats to conclusion validity due to larger survey samples and increased use of archival data (usually with large numbers of cases), better measures, better modelling, etc. The larger samples (or populations) allow for increased use of subgroup analyses as a means of "internal replication." Some researchers also provide evidence from more than one sample or data set, like Grégoire and Shepherd (2012) do in their exemplary study of the "individual-opportunity nexus"[17] and Obschonka et al. (2015) do in their analysis of entrepreneurial regions.

Part of the quality increase also shows in terms of increased requests for and provision of *robustness testing*. This serves a similar purpose as (and sometimes coincides with) the "replicating yourself" examples given above. That is, authors provide better evidence that the support for their theory holds up across various variations in the testing procedure and thus that it is not happenstance of a very particular model specification, analysis technique, or operationalizations of critical variables.

Meta-analyses that sum up the collective evidence have become commonplace in entrepreneurship, and we now know much better what is the main thrust of results reported on a range of topics (Brandstätter, 2011; Brinckmann, Grichnik, & Kapsa, 2010; Martin, McNally, & Kay, 2013; Rauch & Frese, 2007; Rosenbusch, Brinckmann, & Bausch, 2011; Rosenbusch, Rauch, & Bausch, 2013; Unger, Rauch, Frese, & Rosenbusch, 2009; Zhao & Seibert, 2006). This is very good! In addition, recent improvements make meta-analysis an excellent tool for detecting publication bias (O'Boyle et al., 2014b). This said, meta-analyses cannot really test theoretical models but only relatively simple, empirical generalizations. In addition, because of the lack of a replication culture, the studies that are aggregated may not have focused on the relationship the meta-analysis is after, and what gets aggregated are results based on relatively poor operationalizations originally meant to capture some other theoretical construct or intended to be used only as control variables. This may be one reason why effects of indicators of human capital (which are often thrown in as controls) come out as relatively weak compared to effects of personality (which researchers are unlikely to include as single-item control variables) (cf. Rauch & Frese, 2007; Unger et al., 2009). O'Boyle et al.'s (2014b) compilation illustrates the perils of meta-analyzing studies that were not originally addressing the theoretical construct that is the subject of the meta-analysis. As a result of these limitations,

[17] In addition, they do all they can to reduce experimental research's eternal problem of questionable external validity by using samples of real entrepreneurs as well as real technologies and real market needs in the experimental design. It's an awesome piece of entrepreneurship research.

meta-analysis is a useful tool but does not replace the need for explicit replication studies or narrative reviews.

The emergence and increased recognition of the evidence-based entrepreneurship movement (Frese, Rousseau, & Wiklund, 2014), of which the surge in meta-analyses is part. The evidence-based philosophy is largely in line with everything I have professed in this chapter. It is also a perspective that may open up to a very sympathetic type of stock-taking of our collective evidence across different research paradigms (van Burg & Romme, 2014).

The increasing adoption (albeit slowly) of Bayesian approaches and other alternatives to mindless testing solely against a "null" alternative (Edwards & Berry, 2010; Johnson, van de Schoot, Delmar, & Crano, 2014; Schwab et al., 2011). If nothing else, this might force researchers to think a little bit more about what they are doing when they engage in hypothesis testing. It also provides interesting alternatives for those who are willing to learn and break new ground.

The emergence of new journals, which show a broader appreciation for different types of scholarly contributions. For example, both *Academy of Management Discoveries (AMD)* and *Journal of Business Venturing Insights (JBVI)* explicitly encourage replication studies; JBVI also welcomes "non-findings" (i.e., evidence of lack of an effect of meaningful magnitude). The latter journal has had a flying start with interesting and provocative findings as well as robust debates involving established researchers (see Coad, Frankish, Roberts, & Storey., 2015; Crawford, McKelvey, & Lichtenstein, 2014; Davidsson, 2015; Delmar, 2015; Derbyshire & Garnsey, 2014; Honig & Samuelsson, 2014). It will also be interesting to see how new initiatives like *Heliyon* ("A home for all sound research") and *PLoS One* will fare in the future—in general and as forums for entrepreneurship research.

The event of journals being so fed up with misuse of statistical significance testing that they actually ban it. You don't believe what you just read? Then check Trafimow and Marks (2015)! My heroes!

And then, literally days before I submit this manuscript, appears an even greater example of academic heroism when s*cience*—no less—published the article *Estimates of the Reproducibility of Psychological Research* (Aarts et al., 2015).[18] This open science collaboration under Brian Nosek's leadership reports on *one hundred* systematic replications of studies published in leading psychology journals. This large-scale effort confirms many of the points I have argued in this chapter and provides some heuristics to hold on to until we have more precise evidence from entrepreneurship specifically: (a) despite reaching the holy significance ($p < 0.05$) in the original study, under 40 % of studies successfully replicate; (b) the average effect size in careful replications is about half of what was obtained in the original study; (c) large effects are more reproducible; and (d) surprising effects are less reproducible. Should this make us sad or embarrassed? Hardly. As the authors

[18] Special thanks to A. A. Aarts; some name-changing ancestor; Brian Nosek, and the Science editors for giving me the chance to put this important work as the very first entry in this book's rather long list of references!

argue: “Any temptation to interpret these results as a defeat for psychology, or for science more generally, must contend with the fact that this project demonstrates *science behaving as it should*” (pp. aac4716-7; italics added).

Aah! That was a fun section to write! Feeling much better now!

9.6 Summary and Conclusion

In this chapter, I have argued against switching one’s brain off in the presence of asterisks and instead argued for effect size assessment and—above all—replication. Before jumping to (implicit) discussion of “The World about Which We Wish to Know and Tell” (cf. Chap. 4), we should really take a closer look at “our study” as such. What seems to be true within the confines of our study? If that is true in that context, would that seem to be important for the actors in the study? If it *were* true also *outside of* our studied context, how important would that be? *If and only if* the answer to that question is that “yes, it would indeed be important,” do we have reason to address the question we fantasize that statistical significance answers: *How likely* is it that the results of our study are valid for (large parts of) the entire world that we are interested in? The honest answer is that we do not know how likely that is and that statistical significance testing does not change that fact by much at all. Hence, we need to replicate our studies within similar and different contexts in order to really develop some solid knowledge.

I have tried to support this view with examples from empirical research. So much have I praised the virtue of replication and trashed statistical significance testing that I will not burden the reader with any more canonizing or dishonoring, respectively, of them here. Suffice it to repeat that replication provides us with much better truth criteria than other tools at our disposal. Replication therefore facilitates the building of collective and cumulative knowledge, which is what research is all about (if you ask me). Also importantly, replication has a sound, humbling effect that may make us less prone to over interpret single study results regarding relative importance of explanatory variables, prematurely disregard antecedents that do not turn out significant in an individual study, or show an undue level of confidence in a result that happens to be (marginally) statistically significant in a single study.

References

Aarts, A. A. et al. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251) (28 August 2015). Doi: 10.1126/science.aac4716. (270 co-authors under B. Nosek’s leadership).

Ács, Z. J., & Armington, C. (2006). *Entrepreneurship, geography, and American economic growth*. Cambridge, UK: Cambridge University Press.

Álvarez, C., Urbano, D., & Amorós, J. E. (2014). GEM research: Achievements and challenges. *Small Business Economics, 42*(3), 445–465.

Amorós, J. E., Bosma, N., & Levie, J. (2013). Ten years of global entrepreneurship monitor: Accomplishments and prospects. *International Journal of Entrepreneurial Venturing, 5*(2), 120–152.

Armstrong, J. S. (1970). How to avoid exploratory research. *Journal of Advertising Research, 10*(4), 27–30.

Audretsch, D. B., Dohse, D., & Niebuhr, A. (2010). Cultural diversity and entrepreneurship: A regional analysis for Germany. *The Annals of Regional Science, 45*(1), 55–85.

Audretsch, D. B., & Fritsch, M. (1994). The geography of firm births in Germany. *Regional Studies, 28*(4), 359–365.

Bae, T. J., Qian, S., Miao, C., & Fiet, J. O. (2014). The relationship between entrepreneurship education and entrepreneurial intentions: A meta-analytic review. *Entrepreneurship: Theory and Practice, 38*(2), 217–254.

Boschma, R. A., & Fritsch, M. (2009). Creative class and regional growth: Empirical evidence from seven European countries. *Economic Geography, 85*(4), 391–423.

Bosma, N., & Schutjens, V. (2011). Understanding regional variation in entrepreneurial activity and entrepreneurial attitude in Europe. *The Annals of Regional Science, 47*(3), 711–742.

Brandstätter, H. (2011). Personality aspects of entrepreneurship: A look at five meta-analyses. *Personality and Individual Differences, 51*(3), 222–230.

Brinckmann, J., Grichnik, D., & Kapsa, D. (2010). Should entrepreneurs plan or just storm the castle? A meta-analysis on contextual factors impacting the business planning-performance relationship in small firms. *Journal of Business Venturing, 25*(1), 24–40.

Brown, T., Davidsson, P., & Wiklund, J. (2001). An operationalization of Stevenson's conceptualization of entrepreneurship as opportunity-based firm behavior. *Strategic Management Journal, 22*(10), 953–968.

Carver, R. P. (1978). The case against statistical significance testing. *Harvard Educational Review, 48*(3), 378–399.

Chan, Y. E., Bhargava, N., & Street, C. T. (2006). Having arrived: The homogeneity of high growth small firms. *Journal of Small Business Management, 44*(3), 426–440.

Coad, A., Frankish, J. S., Roberts, R. G., & Storey, D. J. (2015). Are firm growth paths random? A reply to "Firm growth and the illusion of randomness". *Journal of Business Venturing Insights, 1*(3), 5–8.

Cohen, J. (1994). The earth is round (p < .05). *American Psychologist, 47*(12), 997–1003.

Cooper, A. C., Gimeno-Gascon, F. J., & Woo, C. Y. (1994). Initial human and financial capital as predictors of new venture performance. *Journal of Business Venturing, 9*(5), 371–395.

Crawford, G. C., McKelvey, B., & Lichtenstein, B. B. (2014). The empirical reality of entrepreneurship: How power law distributed outcomes call for new theory and method. *Journal of Business Venturing Insights, 1*, 3–7.

Crook, T. R., Shook, C. L., Morris, M. L., & Madden, T. M. (2010). Are we there yet? An assessment of research design and construct measurement practices in entrepreneurship research. *Organizational Research Methods, 13*(1), 192.

Dahlqvist, J., Davidsson, P., & Wiklund, J. (2000). Initial conditions as predictors of new venture performance: A replication and extension of the Cooper et al. study. *Enterprise and Innovation Management Studies, 1*(1), 1–17.

Davidsson, P. (1988). Type of man and type of company revisited: A confirmatory cluster analysis approach. In B. Kirchhoff, W. Long, W. McMullan, K. Vesper, & W. Wetzel (Eds.), *Frontiers of entrepreneurship research 1988* (pp. 88–105). Wellesley, MA: Babson College.

Davidsson, P. (1989a). *Continued entrepreneurship and small firm growth*. Doctoral dissertation, Stockholm School of Economics, Stockholm.

Davidsson, P. (1989b). Entrepreneurship—and after? A study of growth willingness in small firms. *Journal of Business Venturing, 4*(3), 211–226.

Davidsson, P. (1995b). *Determinants of entrepreneurial intentions working paper 1995:1*. Jönköping: Jönköping International Business School. http://eprints.qut.edu.au/2076/1/RENT_IX.pdf.

Davidsson, P. (1995c). *Kultur och Entreprenörskap - en uppföljning (Culture and Entrepreneurship - A Follow-up)*. Örebro: Stiftelsen Forum för Småföretagsforskning.

Davidsson, P. (2015). Data replication and extension: A commentary. *Journal of Business Venturing Insights, 3*, 12–15.

Davidsson, P., Delmar, F., & Wiklund, J. (2002). Entrepreneurship as growth; growth as entrepreneurship. In M. A. Hitt, R. D. Ireland, S. M. Camp, & D. L. Sexton (Eds.), *Strategic entrepreneurship: Creating a new mindset* (pp. 328–342). Oxford, UK: Basil Blackwell & Mott, Ltd.
Davidsson, P., Lindmark, L., & Olofsson, C. (1994). New firm formation and regional development in Sweden. *Regional Studies, 28*, 395–410.
Davidsson, P., Steffens, P., & Fitzsimmons, J. (2009). Growing profitable or growing from profits: Putting the horse in front of the cart? *Journal of Business Venturing, 24*(4), 388–406.
Davidsson, P., & Wahlund, R. (1992). A note on the failure to use negative information. *Journal of Economic Psychology, 13*, 343–353.
Delmar, F. (1996). *Entrepreneurial behavior and business performance*. Stockholm: Stockholm School of Economics.
Delmar, F. (2015). A response to Honig and Samuelsson (2014). *Journal of Business Venturing Insights, 3*, 1–4.
Derbyshire, J., & Garnsey, E. (2014). Firm growth and the illusion of randomness. *Journal of Business Venturing Insights, 1*, 8–11.
Edwards, J. R., & Berry, J. W. (2010). The presence of something or the absence of nothing: Increasing theoretical precision in management research. *Organizational Research Methods, 13*(4), 668.
Evanschitzky, H., Baumgarth, C., Hubbard, R., & Armstrong, J. S. (2007). Replication research's disturbing trend. *Journal of Business Research, 60*(4), 411–415.
Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US States Data. *PloS One*, *5*(4), doi: 10210.11371/journal.pone.0010271.
Fauchart, E., & Gruber, M. (2011). Darwinians, communitarians, and missionaries: The role of founder identity in entrepreneurship. *Academy of Management Journal, 54*(5), 935–957.
Frank, H., Kessler, A., & Fink, M. (2010). Entrepreneurial orientation and business performance: A replication study. *Schmalenbach Business Review, 62*, 175–198.
Frese, M., Rousseau, D. M., & Wiklund, J. (2014). The emergence of evidence-based entrepreneurship. *Entrepreneurship: Theory and Practice, 38*(2), 209–216.
Fritsch, M., & Storey, D. J. (2014). Entrepreneurship in a regional context: Historical roots, recent developments and future challenges. *Regional Studies, 48*(6), 939–954.
Garofoli, F. (1994). New firm formation and regional development: The Italian case. *Regional Studies, 28*(4), 381–393.
George, G. (2014). Rethinking management scholarship. *Academy of Management Journal, 57*(1), 1–6.
Grégoire, D. A., & Shepherd, D. A. (2012). Technology-market combinations and the identification of entrepreneurial opportunities: An investigation of the opportunity-individual nexus. *Academy of Management Journal, 55*(4), 753–785.
Guesnier, B. (1994). Regional variations in new firm formation in France. *Regional Studies, 28*(4), 347–358.
Hart, M., & Gudgin, G. (1994). Spatial variations in new firm formation in the republic of Ireland, 1980–90. *Regional Studies, 28*(4), 367–380.
Holcomb, T. R., Combs, J. G., Sirmon, D. G., & Sexton, J. (2010). Modeling levels and time in entrepreneurship research an illustration with growth strategies and post-IPO performance. *Organizational Research Methods, 13*(2), 348–389.
Honig, B., & Samuelsson, M. (2014). Data replication and extension: A study of business planning and venture-level performance. *Journal of Business Venturing Insights, 1*, 18–25.
Hubbard, R., & Lindsay, R. M. (2013a). From significant difference to significant sameness: Proposing a paradigm shift in business research. *Journal of Business Research, 66*(9), 1377–1388.
Hubbard, R., & Lindsay, R. M. (2013b). The significant difference paradigm promotes bad science. *Journal of Business Research, 66*(9), 1393–1397.
Hubbard, R., Vetter, D. E., & Little, E. L. (1998). Replication in strategic management: Scientific testing for validity, generalizability, and usefulness. *Strategic Management Journal, 19*, 243–254.

Hunter, E. J. (2009). *Celebrity entrepreneurship and celebrity endorsement.* Doctoral dissertation, Jönköping International Business School, Jönköping, Sweden.

Hunter, E. J., Burgers, J. H., & Davidsson, P. (2009). In G. T. Lumpkin & J. Katz (Eds.), *Celebrity capital as a strategic asset: Implications for new venture strategies* (Advances in entrepreneurship, firm emergence, and growth, Vol. 1, pp. 137–160). Bingley, UK: Emerald.

Hunter, E. J., & Davidsson, P. (2007). Celebrity entrepreneurship: Communication effectiveness through perceived involvement. *International Journal of Entrepreneurship and Small Business, 5*(4), 505–527.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine, 2*(8), e124.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review, 31*(2), 386–408.

Johnson, A. R., van de Schoot, R., Delmar, F., & Crano, W. D. (2014). Social influence interpretation of interpersonal processes and team performance over time using Bayesian model selection. *Journal of Management*. doi:10.1177/0149206314539351.

Keeble, D., & Walker, S. (1994). New firms, small firms and dead firms: Spatial patterns and determinants in the United Kingdom. *Regional Studies, 28*(4), 411–442.

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement, 56*(5), 746–759.

Landis, R. S., & Rogelberg, S. G. (2013). Our scholarly practices are derailing our progress: The importance of “nothing” in the organizational sciences. *Industrial and Organizational Psychology, 6*(3), 299–302.

Leahey, E. (2005). Alphas and asterisks: The development of statistical significance testing standards in sociology. *Social Forces, 84*(1), 1–24.

Liero, H., & Zwanzig, S. (2013). *Introduction to the theory of statistical inference*. Boca Raton, FL: CRC Press.

Martin, B. C., McNally, J. J., & Kay, M. J. (2013). Examining the formation of human capital in entrepreneurship: A meta-analysis of entrepreneurship education outcomes. *Journal of Business Venturing, 28*(2), 211–224.

Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. *Administrative Science Quarterly, 55*(3), 439–471.

Oakes, M. (1986). *Statistical inference: A commentary for the social and behavioural sciences*. Chichester, UK: John Wiley & Sons, Inc.

O’Boyle, E. H., Banks, G. C., & Gonzalez-Mulé, E. (2014). The chrysalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*. doi:10.1177/0149206314527133.

O’Boyle, E. H., Rutherford, M. W., & Banks, G. C. (2014). Publication bias in entrepreneurship research: An examination of dominant relations to performance. *Journal of Business Venturing, 29*(6), 773–784.

Obschonka, M., Andersson, H., Silbereisen, R. K., & Sverke, M. (2013). Rule-breaking, crime, and entrepreneurship: A replication and extension study with 37-year longitudinal data. *Journal of Vocational Behavior, 83*(3), 386–396.

Obschonka, M., Stuetzer, M., Gosling, S. D., Rentfrow, P. J., Lamb, M. E., Potter, J., et al. (2015). Entrepreneurial regions: do macro-psychological cultural characteristics of regions help solve the “knowledge paradox” of economics? *PloS One, 10*(6), e0129332.

Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers’ perceived expertise, trustworthiness, and attractiveness. *Journal of Advertising, 19*(3), 39–52.

Plummer, L. A. (2010). Spatial dependence in entrepreneurship research challenges and methods. *Organizational Research Methods, 13*(1), 146–175.

Rauch, A., & Frese, M. (2007). Let’s put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners’ personality traits, business creation, and success. *European Journal of Work and Organizational Psychology, 16*(4), 353–385.

Reynolds, P. D. (1994). Autonomous firm dynamics and economic growth in the United States 1986–1990. *Small Business Economics, 28*(4), 429–442.

Reynolds, P. D., Storey, D. J., & Westhead, P. (1994). Cross-national comparisons of the variation in new firm formation rates. *Regional Studies, 28*(4), 443–456.

Rosenbusch, N., Brinckmann, J., & Bausch, A. (2011). Is innovation always beneficial? A meta-analysis of the relationship between innovation and performance in SMEs. *Journal of Business Venturing, 26*(4), 441–457.

Rosenbusch, N., Rauch, A., & Bausch, A. (2013). The mediating role of Entrepreneurial Orientation in the task environment–performance relationship: A meta-analysis. *Journal of Management, 39*(3), 633–659.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638.

Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in management and organizational science: Assembling the field's full weight of scientific knowledge through syntheses. *Academy of Management Annals, 2*(1), 475–515.

Schwab, A., Abrahamson, E., Starbuck, W. H., & Fidler, F. (2011). Researchers should make thoughtful assessments instead of null-hypothesis significance tests. *Organization Science, 22*(4), 1105–1120.

Shaver, J. P. (1993). What statistical significance testing is, and what it is not. *Journal of Experimental Educational Research, 61*(4), 293–316.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.

Smith, N. R. (1967). *The entrepreneur and his firm: The relationship between type of man and type of company*. East Lansing, MI: Michigan State University.

Stevenson, H. H. (1984). A perspective of entrepreneurship. In H. H. Stevenson, M. J. Roberts, & H. Grousebeck (Eds.), *New business venture and the entrepreneur*. Boston, MA: Harvard Business School.

Stevenson, H. H., & Gumpert, D. E. (1991). The heart of entrepreneurship. In W. A. Sahlman & H. H. Stevenson (Eds.), *The entrepreneurial venture* (pp. 71–80). Boston, MA: Harvard Business School.

Stevenson, H. H., & Jarillo, J. C. (1990). A paradigm of entrepreneurship: Entrepreneurial management. *Strategic Management Journal, 11*, 17–27.

Trafimow, D., & Marks, M. (2015). Editorial. *Basic and Applied Social Psychology, 37*(1), 1–2.

Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2009). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing, 26*(3), 341–358.

van Burg, E., & Romme, A. G. L. (2014). Creating the future together: Toward a framework for research synthesis in entrepreneurship. *Entrepreneurship: Theory and Practice, 38*, 369–397.

Weismeier-Sammer, D. (2011). Entrepreneurial behavior in family firms: A replication study. *Journal of Family Business Strategy, 2*(3), 128–138.

Welter, F. (2011). Contextualizing entrepreneurship—conceptual challenges and ways forward. *Entrepreneurship: Theory and Practice, 35*(1), 165–184.

Wiklund, J. (1998). *Small firm growth and performance: Entrepreneurship and beyond.* Doctoral dissertation, Jönköping International Business School, Jönköping.

Wiklund, J., Davidsson, P., & Delmar, F. (2003). What do they think and feel about growth? An expectancy-value approach to small business managers' attitudes towards growth. *Entrepreneurship: Theory and Practice, 27*(3), 247–269.

Zahra, S. A. (2007). Contextualizing theory building in entrepreneurship research. *Journal of Business Venturing, 22*(3), 443–452.

Zahra, S. A., & Wright, M. (2011). Entrepreneurship's next act. *Academy of Management Perspectives, 25*(4), 67–83.

Zhao, H., & Seibert, S. E. (2006). The big five personality dimensions and entrepreneurial status: A meta-analytical review. *Journal of Applied Psychology, 91*(2), 259.

Zolin, R., Stuetzer, M., & Watson, J. (2013). Challenging the female underperformance hypothesis. *International Journal of Gender and Entrepreneurship, 5*(2), 116–129.

A Quick Look at Analysis Method 10

Abstract

How do we find out what our data really are saying? Once all the evidence is collected, the researcher still has many choices to make in the analysis of the data, and these choices affect conclusions. This short chapter discusses some challenges that are particularly pronounced in entrepreneurship research. These are the process- and multilevel nature of the phenomenon, and the uncertainty and heterogeneity that signifies it. It is also observed that many developments toward increased "sophistication" in the analysis only address the validity of significance tests that are often invalid already for more fundamental reasons.

10.1 Let's Make This a Short One

I used to be the analysis wizard. Used to, admittedly. It's been a while. I had more formal stats training than most of my contemporaries; the equivalent of more than 1 year of full-time study. Already as an undergraduate, I was forced through calculating multiple regression examples by hand. Apart from the standard stuff on sampling, operationalization, and analysis techniques, the mandatory doctoral coursework included a course in matrix algebra.[1] Ever felt the need for a bordered

[1] The course instructor was one of those great masters of pedagogy who learn math themselves in no time while munching their breakfast cereals, but who couldn't work out how to teach it if even if given a lifetime to do so. I also recall one of the short, supplementary books he assigned us to read as written by one of his fellow masters of pedagogy. It was not until the last few pages that I realized that the whole work was repetition of the basic rules of calculus, on which I had spent months and months in high school. However, learning matrix algebra is actually useful. For me, the best use has been as confidence booster. This is how it works: when I come across matrix notation in an article, I can no longer read it, but neither am I intimidated by it because I realize that if I really tried, I *could* understand it, and if I did I would *probably* find the notation to be shorthand for something rather simple.

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3_10

Hessian, anyone? I applied not only cluster analysis, discriminant analysis, and multiple regression (including a moderation analysis of sorts) in my doctoral thesis but also partial least squares analysis (PLS), which in the late 1980s meant working on the basis of documentation like impenetrable draft manuscripts by Herman Wold and Lohmöller's arbitrarily incomplete manual. I learnt LISREL—the first, major structural equations modeling (SEM) software—straight from the horse's mouth; a course held by K.G. Jöreskog and Dag Sörbom. Similarly, I was taught PLS by the young stallion in the stable next door (Claes Fornell, as in Fornell & Larcker, 1981). And at a very early stage, I bought from Sawtooth Software some Conjoint Analysis software on 5.25-inch (360 kilobytes!) flexible floppy disks. Yeah, it's been a while….

In my early collaborations with more senior people, I was definitely the methods guy (among other things). I kept that up for a decade or so after graduation, but eventually that role started to drift to my doctoral students and other junior partners. That's the way of us old fossils. The drift probably started with Frédéric Delmar's learning and applying CHAID for Delmar and Davidsson (2000) and later his cross-validation approach to cluster analysis in Delmar, Davidsson, and Gartner (2003). It then continued with others such as Mikael Samuelsson finding and using longitudinal growth modeling in Samuelsson and Davidsson (2009); Lucia Naldi's application of negative binomial and fractional logit regression techniques in Naldi and Davidsson (2014) and Scott Gordon's use of the difference-in-difference (DID) analysis in Davidsson and Gordon (2015). Occasionally, we brought in a trained econometrician as coauthor, such as Girma in Lockett, Wiklund, Davidsson, and Girma (2011) and Hatemi-J in Davidsson, Kirchhoff, Hatemi-J and Gustavsson (2002).

In short, I'm not the analysis technique wizard any more. Further, *a lot* has happened on the analysis method frontier in the last decade or so, and I'm not the best source from which to capture that knowledge. So let's make this chapter a short one. I will refrain from elaborate examples and instead confine my exposition to giving (a) hints about general problems that have to be dealt with and (b) references to works by scholars in possession of deeper methods knowledge than mine regarding solutions to these problems. This said, I may still have some relevant insights to offer in brief. I have mentioned several of these here and there throughout the book. Among the more basic and more important ones are these:

1. Although application of correct and sophisticated analysis techniques can help get the most out of your data, trying to fix fundamental problems is a tad late at this stage.[2] The GIGO principle (Garbage In = Garbage Out) is alive and well.

[2] In my estimation, doctoral training focuses relatively too much on the later stages of the research process, i.e., on analysis and "packaging" (publication). These are important, too, but not of much use if you are trying to answer an uninteresting or unimportant question, or work with data that are not good enough to answer whatever question it is that you are trying to answer. Since others are putting so much emphasis on the tail end, I won't offer any publishing advice in this book. On that topic, see instead Fayolle and Wright (2014).

2. Take note of the fact that much of the developments toward greater methods sophistication concern the statistical inference part. For example, if it is about correcting standard errors, then it is about the validity of the significance test. For reasons elaborated in Chap. 9, that test is usually invalid already due to much greater problems than those addressed by the new and sophisticated procedure. In your search for new and better methods, try to focus on those that provide estimates with better internal validity than the old approaches.
3. Remember that no matter how modern or sophisticated your analysis approach, our focus on statistically significant effects tends to lead us astray (Hubbard & Lindsay, 2013). A positive regression coefficient that is internally valid and statistically significant does not necessarily mean that "A has such an important effect on B that it is worth investing money in increasing A in order to get more of B." An internally valid and statistically significant group difference does usually not mean that "members of group A are in general like this, while members of group B are generally like that" (see also footnote 6 of Chap. 9).
4. Various types of corrections and checks that are introduced at the analysis stage, while somewhat helpful, can never fully solve the underlying problem that they are meant to remedy. The fact that the correction was developed by one heck of a man, and is routinely demanded and accepted by journal reviewers, does not mean it can fully solve such problems as retrospection bias, survivor bias, or inability to experimentally manipulate the variance of the explanatory variable in focus (Semadeni, Withers, & Certo, 2014).
5. By and large, if it takes a lot of econometric trickery to "prove" the existence of an effect, then the effect is probably not very large or important. Really important effects tend to scream at you in much simpler analyses. We may need the more sophisticated analyses to rule out other explanations, make a stronger case for causality, and get closer to the truth about the form and magnitude of the effect, but if it does not show much at all in simpler analyses, there is a real risk that what is reported is the result of analytic wizardry rather than the effect of A on B in the sample or in the world about which we wish to know and tell (Chap. 4).[3]

Revisiting my verbose domain delineation from Chap. 2, we find statements about *uncertainty and heterogeneity,* about *processes,* and about antecedents and effects on *different levels of analysis.* All of these characteristics (or choices) have implications for analysis. Let's take a closer look.

[3] This is genuinely tricky, because admittedly there are also cases where the multivariate pattern of relationships conceal true and important effects, and we really need complex and sophisticated techniques in order to unveil them.

10.1.1 Uncertainty, Heterogeneity, and Analysis Method

The first implication of uncertainty and heterogeneity is that we should never expect anything near full explanation, even if all important variables are included in the explanatory model, with perfect operationalizations. Hence, omission of other important, expected outcomes and poor operationalizations of those actually included are not the only reasons—and probably not even the main reasons—for reaching only 20–23 % explained variance in the analysis in Table 9.3. Neither do the eight variables come out with the coefficients they do because each of the eight factors influence overall growth attitude exactly this much for every small business owner-manager. The reason why *employee well-being* seems more important than *independence* is not necessarily that for every respondent the first factor is more important than the second. It may just as well be the case that relatively more managers consider the first factor at all in their evaluation of the prospect of growth. For a distinct minority, concerns about independence may be *the* defining growth motivator or growth deterrent. As a result, regression coefficients normally represent *average* effects for members of the investigated population. As per our heterogeneity assumption, the true effect is different for every case, and by virtue of Knightian and even Heisenbergian uncertainty, the effect for each case is not an unshakeable law of nature, either. Which means part of the variance will remain unexplained.

What are the ways in which we can deal analytically with this causal heterogeneity (cf. Chap. 4) in the analysis? One is, of course, to accept it and its effect on explanatory power. Another obvious way to deal with it is to conduct subsample analyses in order to see what the relationships look like for different, more homogeneous, subgroups (cf. the "replicate yourself" examples in Chap. 9). A very popular line of attack in recent years has been to apply moderated regression; i.e., to include interaction effects in the explanatory model. However, this can only do so much: the number of interactions you can include in one model is narrowly restricted if you also want to be able to interpret your coefficients. What if it is *not* a useful simplification that the effect for *men* or *online start-ups* is β_1 whereas the effect for *women* or *brick-and-mortar start-ups* is β_2, but instead each case has its own, unique regression equation across all explanatory variables? Trying to find all these unique functions would at best lead us toward complete description instead of development or validation of something worth calling a theory, but we may still want to take strides toward acknowledging more causal heterogeneity than just a couple of interaction effects in an otherwise global model.

One way of finding out how explanatory variables matter differently for different parts of the population is to run *automatic interaction detection* of some kind. This classification tree technique works when the dependent variable is dichotomous, such as being or not being a business founder. The sample also needs to be large; otherwise, the branching of the tree quickly leads to small subgroups and therefore unreliable results. For an entrepreneurship application, see Davidsson and Delmar (2000). When the dependent variable is continuous and you suspect models are different for known categories in the data, a crude but relevant approach is to perform separate subgroup analyses. For example, in Samuelsson and Davidsson (2009), we showed that regression models were radically different for imitative vs. innovative start-ups. In that work, we actually went further than just using known (empirical) categories, because *Latent Class Analysis* on four indicators was used to classify cases as imitative vs. innovative in the first place.

Many years ago, I played with the idea of combining principles of cluster analysis and regression analysis in order to find subgroups in the population whose outcomes are explained by regression models that are the same within the subgroup but different across groups. Apparently, this is what is now done with *regression trees* in the world of data mining (see Bou Kheir, Nekhili, Jokung, Chtioui, & Bellalah, 2015). That is, the analysis does not start from predefined groups or with an assumption that a global regression model will fit the entire population reasonably well. Instead, the procedure finds subgroups where each subgroup has a different regression model. A result of this is that total explanatory power is greatly enhanced (by maximizing the capitalization on chance, some would say). In a sense, cases are grouped not because they have similar characteristics across a range of variables as in cluster analysis. Instead, cases are grouped because they share similar *effects* of the independent variables on the dependent variables.

Another heterogeneity problem discussed in Chap. 4 is that of *unmeasured heterogeneity.* One approach to dealing with this when the omitted variables cannot be identified and added to the model is *fixed effects multiple regression* (Allison, 2009). This technique requires longitudinal data and makes the important leap to rely only on within-case variation over time in the independent variables as a cause of their value on the dependent variable. In regular regression models, it is assumed that the variance *across* cases is caused by the variables in the model. In a complex world, that assumption is quite a leap of faith. This said, fixed effects regression is not always the right savior. It assumes that the case-specific effects bundled in the omitted variables are stable over time, which may be unrealistic for long time series. Further, if you study, e.g., the effect of country-level institutional factors on the level of entrepreneurial activity, there may not be enough within-country variation in the IVs to make for a credible model or strong explanation.

Propensity score matching (e.g., Elert, Andersson, & Wennberg, 2015) is another device which has come into use in recent years to account for (otherwise) unmeasured heterogeneity. While useful, this is one of those tools for improvements at the analysis stage one should realize is unlikely to *fully* solve the problem it is supposed to help remedy.

Entrepreneurship research takes an interest in the exceptional, and should do so. Often the analysis is aimed at learning about a small minority that stands out from a larger population: individuals currently involved in a start-up, venture capital-backed ventures, IPOs, rapidly growing firms, or some other select minority. Again, this has implications not only for sampling but also for analysis method.

For example, one major research question in GEM- and PSED-type research concerns who is more likely to be a nascent entrepreneur. We are then talking about a single-digit minority of the general population. The typical analysis method for this type of research question, logistic regression analysis, does not provide unbiased estimates when the group sizes are this uneven. In addition, with standard application of logistic regression, you may end up with a function that reaches a high correct classification rate overall by putting almost all cases in the less entrepreneurial group, i.e., by performing poorly with respect to what was the researcher's key interest. *Rare Events Logistic Regression* or *ReLogit* can be the right solution in such situations (Bergmann, Mueller, & Schrettle, 2014).

Although my perspective acknowledges relatively mundane, imitative venture start-ups as instances of entrepreneurship, there is no doubt that the most interesting (and infrequent) cases are found at the other outskirt of distributions. This points to an inconvenient truth and a very fundamental problem, namely that the standard package of statistical methods—this method knowledge in which we have all invested sweat equity—is ill-adapted to entrepreneurship research problems. These variance-explaining techniques typically focus on central tendencies, preferably for normally distributed variables. Outliers are technical problems to be eliminated, not thrilling empirical phenomena of the highest societal importance. In sharp contrast, our key interest may rest with the rare cases at the high end of a highly skewed distribution. This clash behooves us to show the guts and marshal the energy to disregard our vested interest in the conventional. Much like the problem of secondary data that are fundamentally inadequate for the purpose (remember the drunkard under the streetlight in Chap. 4?), we have to face it when the methods we have learnt don't do the job properly. We have to look elsewhere. There are alternatives to opening the tin can with a hammer.

Crawford, McKelvey, and Lichtenstein (2014) and Crawford, Aguinis, Lichtenstein, Davidsson, and McKelvey (2015) have recently highlighted and discussed this problem. In the latter work, we first demonstrate that quite a range of core variables in entrepreneurship are power law distributed rather than being anything near the Gauss curve. That is, the distributions resemble the gray area in Fig. 10.1. This is true not just for outcome variables but also for some common candidates for an explanatory variable role. In a power law distribution, lots of cases have very low values, while a select few have extremely high values, orders of magnitude beyond anything remotely "normal." This renders the arithmetic mean pretty useless. It is not a "middle" value and not a particularly common one, either. These problems carry over to most of our usual analysis techniques.

The fact that power law distributions are much more common in our field than are so-called "normal" distributions comes with strong implications for entrepreneurship research. We discuss these at some length in Crawford et al. (2015). The implications go beyond the analysis stage and involve also theorizing and the nature of the data that need to be collected. Focusing *only* on high-end cases (if we knew where to find them) is not the solution, because this would not allow insights into the generative mechanisms that produce the extreme cases. Dichotomizing the skew but continuous distribution in order to apply ReLogit to the important but low-frequency "high end" of entrepreneurship would be doable, but unsatisfactory. One suggestion is that we are better equipped to analyze outcomes on the aggregate level; i.e., we may be able to develop theory and analysis methods that can explain why the extremely skew distributions emerge, but not necessarily what particular cases end up in the right tail. Computer simulations are likely to be useful for

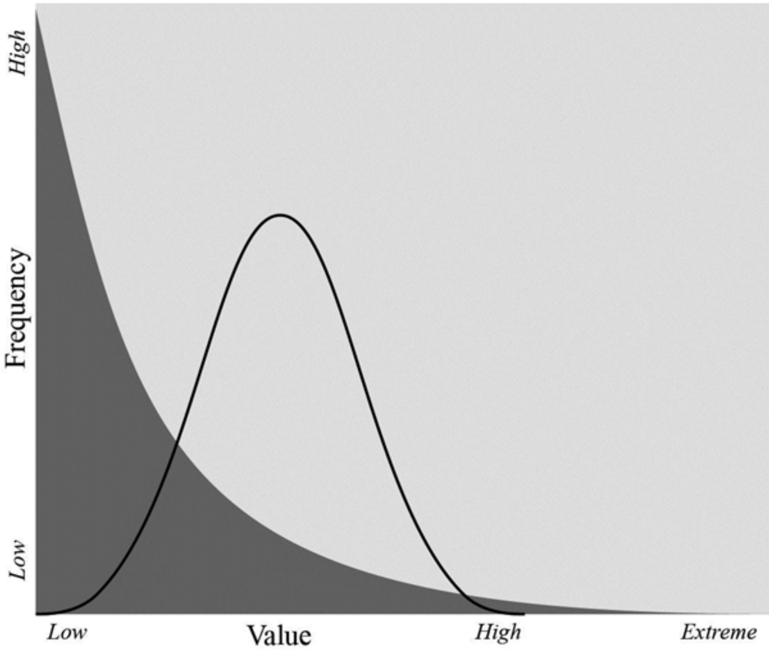


Fig. 10.1 Normal distribution vs. power law distribution (from Crawford et al., 2015; adapted from O'Boyle & Aguinis, 2012)

understanding what microlevel processes are consistent with observed, aggregate level phenomena, but "picking winners" on the microlevel is likely to remain elusive. However, approaches of the kind recently championed by Guzman and Stern (2015a, 2015b) actually afford a glimmer of hope to the latter as well.

10.1.2 Analysis Implications of Entrepreneurship as Process

This book has advocated a process perspective on entrepreneurship. This, too, has implications not only for data collection but also for choice of analysis techniques. It is possible to attain new insights about entrepreneurial processes by applying conventional techniques to longitudinal data (e.g., Davidsson & Honig, 2003). However, this is not the hope for the future and not even acceptable current practice. In order to better deal with the specific data challenges, and to make full use of the longitudinal aspects of the data, other techniques may have to be applied.

Delmar and Shane (2004) were among the pioneers to introduce *event history analysis* (EHA; Blossfeld, Hamerle, & Mayer, 2014) into entrepreneurship research. In event history analysis, the data set is organized as monthly (or weekly, bimonthly, yearly, etc.) spells. The technique makes use of the longitudinal aspect of the dependent as well as independent variables. Independent variables can be entered as time invariant or time variant. In the latter case, the value of the independent variable is allowed to change over time, which is what makes the technique truly longitudinal rather than just time-ordered. The dependent variable is always categorical. It changes its value in the period (e.g., month) when the event to be predicted has occurred. Thus, EHA can be regarded a longitudinal alternative to logistic regression. Cases where the event has still not occurred when the last data collection is made are treated as right censored, a problem the technique is designed to deal with.

The original logic of the technique assumes all cases are heading toward an inevitable outcome (death, it's just a matter of time!) which makes it especially suited for predicting abandonment vs. continuation of the start-up processes. However, EHA has developed into a family of techniques for different situations, and now can handle multiple, possible outcomes (see Bakker & Shepherd, 2015). General and entrepreneurship-specific data requirements for EHA, and the consequences of violating them, have also become much better understood and documented over the past decade (Delmar, 2015; Yang & Aldrich, 2012).

Bengt Muthén (Muthén & Curran, 1997; Muthén & Khoo, 1999) has taken the Uppsala University tradition (e.g., Jöreskog's LISREL and Wold's PLS) of structural equation modeling with latent variables into longitudinal territory. Thus, when the dependent variable is continuous, *longitudinal growth modeling* (growth curve analysis; latent growth modeling) can be a particularly interesting alternative. My then doctoral student Mikael Samuelsson stumbled over this technique early in his studies. Later he came to use the Swedish PSED data, with their problems of unequal entry stage and likewise differential pace of development. This made the technique almost uncannily well suited for his needs because of its ability to model and predict both initial state and development over time (Samuelsson, 2004; Samuelsson & Davidsson, 2009). A shortcoming of the technique—at least in the form we applied it back then—is that unlike EHA it cannot include cases that discontinue during the studied period in the analysis. This calls for supplementary analyses in order to avoid erroneous conclusions based on success bias.

One aspect of process is the problem of analyzing sequences of events. Liao, Welsch, and Tan (2005) pioneered using data mining software for this purpose. I have not seen any follow-ups using their particular approach (a software called Clementine®), but I'm sure there are modern data mining routines that provide interesting alternatives. Scott Gordon applied sequence analysis (Abbot, 1995)—well known from genetics—in his dissertation work (Gordon, 2012) with promising results. This technique makes it possible to compare how similar empirical sequences are to target sequences (and to each other). In Scott's case, each start-up sequence of discovery and exploitation actions over time was compared with five theoretical sequences, five random sequences, and five empirical norm sequences (similar to cluster centroids). Likewise working with "gestation activities" data from PSED-type research, Hak, Jaspers, and Dul (2013) review and discuss some of the pros and cons of a range of techniques for analyzing sequences: *event structure analysis* (ESA), *optimal matching* (OM), *temporal qualitative comparative analysis* (TQCA), and *necessary condition analysis* (NCA). The name of the latter technique also brings to mind Eckhardt, Shane, and Delmar's (2006) multistage selection approach to the problem of new venture finance. Specifically, these authors do not regress predictors directly onto venture-funding success. Instead they first theorize predictors of seeking funding; without this voluntary act, the process cannot continue toward success at obtaining external funding.

A serious interest in process issues naturally leads to theorizing and testing that the effects of explanatory variables differ by stage of development. Some things may be important early on while others come more to the fore at later stages. The influence of some variables may even change signs from early to late in the process. In Davidsson and Honig (2003), we could only hint at such patterns, while recent examples in the entrepreneurship domain better show the potential of such approaches (Bakker & Shepherd, 2015; Johnson, van de Schoot, Delmar, & Crano, 2014). When the process is disrupted by some external shock (which applies to the entire sample) we have a natural experiment. Assuming before- and after-measures are available, this sets the stage for *difference-in-difference analysis* (DID). The method is described, e.g., by Lechner (2011). For an entrepreneurship application, see Davidsson and Gordon (2015).

A process focus calls for longitudinal data, and this introduces new issues compared to cross-sectional data. While missing data (nonresponse) is always a problem that has to be dealt with in data analysis, it is aggravated when the data are longitudinal. With multiple waves of data, the likelihood that a case has complete information on every variable we want to include in the analysis asymptotically approaches zero. The problem of attrition—that some cases are lost entirely over time—is bad enough; when we add loss of cases due to partially missing data, we may end up having nothing left to analyze. So, we must find ways to include cases with missing information. The quick and dirty tricks like replacing missing data with the mean or with a predicted value from a regression may be defensible when but a tiny percentage of the cases are manipulated in this way. However, such techniques reduce the error variance, and therefore using them amounts to cheating when we have a lot of partially missing data.

Luckily, method experts have developed more sophisticated techniques for data imputation that can be applied. Back in 2003, *Organization Research Methods* ran a special issue on this topic (see, e.g., Fichman & Cummings, 2003). Newman (2014) recently followed up in the same outlet with some practical advice. Serious methods research underlies the techniques for data imputation that are now increasingly being used. However, a word of caution is in place. The solutions always build on simplifying assumptions. These assumptions are rarely met, and the missing data may be a bit more complex than assumed. For example, in longitudinal entrepreneurship research, you will have the following types of missing data, with implications of varying gravity depending on the analysis method and the purpose of the analysis: (a) cases that should have been included in the original sample (wave 1) but did not participate, (b) cases that are "validly" missing in later waves because they are known to be terminated, (c) cases that can no longer be found or refuse an entire wave of data collection, and (d) cases with missing values on individual questions, whether because of refusal, inability to answer, interviewer mistakes, or programming error. None of this is likely to happen for completely stochastic (random) reasons. Moreover, the methods experts who developed the imputation methods probably did not envisage you using the same imputation method across all or most of those cases.

Missing data are missing. The data you impute aren't the real values for these cases. If you can't back up your conclusions also with data only from complete cases (perhaps just for parts of the process), or at least show they are robust enough to hold up across varying assumptions about the attrition, I reserve the right to remain a sceptic.

10.1.3 Analysis Implications of Entrepreneurship as a Multilevel Phenomenon

In this book, I have repeatedly portrayed entrepreneurship as a multilevel phenomenon. And I'm not alone; *Journal of Business Venturing* also defines the field in this manner. Acknowledging more than one level presents an interestingly expanded set of analysis alternatives, as well as challenges to overcome.

The simplest manner in which to consider multiple levels in the analysis is to include explanatory variables pertaining to different levels (e.g., individual, venture, industry, region) or to examine the effects of a given set of predictors on outcome variables on different levels. However, nowadays researchers apply specific, multilevel analysis techniques which allow insights into the interplay across levels. Still not ubiquitous, explicit multilevel approaches are no longer a rare occurrence in entrepreneurship research. A quick search on Google Scholar with a combination of "multilevel" and entrepreneurship terms in the *title* yielded a dozen or so journal article hits since 2010, some of which appear in leading, mainstream journals. There would be many more examples where the term does not appear in the title. For example, Kwon and Arenius (2010) apply multilevel modeling to analyze both individual- and country-level GEM data. This approach has the distinctive advantage of making it possible to avoid misattributing effects to the wrong level. To illustrate, the low entrepreneurial activity of a country is not necessarily due to it being particular in any other way than having a high proportion of individuals who would show low entrepreneurial inclination in any context. If the analysis were limited to the aggregate level, this distinction could not be made.

Although many applications involve the national level in combination with individual or firm levels, multilevel analyses can be applied also in more micro-oriented research. For example, Shepherd (2011) discusses compelling options open to researchers interested in entrepreneurial decision making. In undertaking such research, which might involve, for example, some of the levels *task*, *individual*, *group*, and *organization*, the analyst can draw on the rich experiences from the field of organizational behavior (Klein & Kozlowski, 2000; Kozlowski, Chao, Grand, Braun, & Kuljanin, 2013). In fact, all multilevel work in entrepreneurship can probably benefit from taking a peek in that direction, since they have many decades of experience of the issues at hand (Rousseau, 1985).

Apart from analysis techniques, the levels quest also requires us to take a closer look at our data. In Chap. 4 (around Fig. 4.4), we discussed design fallacies leading to unmeasured heterogeneity pertaining to levels. Specifically, the data may be lacking information about other team members or about resource investments the focal person makes in ventures other than the sampled one. Similar lapses in the logic of the data may occur on other levels of analysis. In this situation, if it is at least known which cases are team-based and which founders are portfolio entrepreneurs, then the analyst can choose to delimit the test only to solo founders who are engaged in a single venture (cf. Dimov, 2010). At least, they should be able to show that in terms of direction and magnitude (if not in terms of statistical significance, due to reduced statistical power), the results hold up for this subsample. If it does not, then the supportive results for the full sample are likely due to reasons other than those hypothesized.

10.2 Summary and Conclusion

In this short and incomplete treatment, I have pointed at some ways in which the perspective on entrepreneurship I advocate influences the choice and application of analysis methods. We dealt in particular with three aspects: the uncertainty and heterogeneity of entrepreneurship, its process character, and its multilevel nature. The toughest-to-accept conclusion arises from the discussion of heterogeneity of outcomes and the extremely skew distributions of core entrepreneurship variables. This is the insight that the standard sets of variance-explaining techniques, which focus on central tendencies, assume normal distributions, and regard outliers a problem, are fundamentally inadequate tools for many analysis tasks in entrepreneurship research.

In relation to new developments, I suggested you focus on improvements to estimation rather than sophistication only in terms of fine tuning the already doomed significance calculations. I have barely touched upon the Big Issue of Big Data and data mining. A focus on such things might have looked a bit anachronistic in this age of "theoretical contributions," but pendulums swing, and it may actually be the next Big Thing in entrepreneurship research.

At any rate, have fun squeezing the most out of your data!

References

Abbott, A. (1995). Sequence analysis: New methods for old ideas. *Annual Review of Sociology, 21*(1), 93–113.

Allison, P. D. (2009). *Fixed effects regression models* (Vol. 160). Newbury Park, CA: Sage.

Bakker, R. M., & Shepherd, D. A. (2015). Pull the plug or take the plunge: Multiple opportunities and the speed of venturing decisions in the Australian mining industry. Paper accepted for publication in *Academy of Management Journal*.

Bergmann, H., Mueller, S., & Schrettle, T. (2014). The use of Global Entrepreneurship Monitor data in academic research: A critical inventory and future potentials. *International Journal of Entrepreneurial Venturing, 6*(3), 242–276.

Blossfeld, H. P., Hamerle, A., & Mayer, K. U. (2014). *Event history analysis: statistical theory and application in the social sciences*. East Sussex, UK: Psychology Press.

Bou Kheir, R., Nekhili, M., Jokung, O., Chtioui, T., & Bellalah, M. (2015). Evaluating financial performance in non-profit microfinance institutions: A regression-tree approach. *Working Paper.* Available at SSRN and Research Gate.

Crawford, G. C., Aguinis, H., Lichtenstein, B., Davidsson, P., & McKelvey, B. (2015). Power law distributions in entrepreneurship: Implications for theory and research. *Journal of Business Venturing*. doi:10.1016/j.jbusvent.2015.01.001.

Crawford, G. C., McKelvey, B., & Lichtenstein, B. B. (2014). The empirical reality of entrepreneurship: How power law distributed outcomes call for new theory and method. *Journal of Business Venturing Insights, 1*, 3–7.

Davidsson, P., & Gordon, S. R. (2015). Much ado about nothing? The surprising persistence of nascent entrepreneurs through macroeconomic crisis. *Entrepreneurship Theory and Practice*. doi:10.1111/etap.12152.

Davidsson, P., & Honig, B. (2003). The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing, 18*(3), 301–331.

Davidsson, P., Kirchhoff, B., Hatemi-J, A., & Gustavsson, H. (2002). Empirical analysis of growth factors using Swedish data. *Journal of Small Business Management, 40*(4), 332–349.

Delmar, F. (2015). A response to Honig and Samuelsson (2014). *Journal of Business Venturing Insights, 3*, 1–4.

Delmar, F., & Davidsson, P. (2000). Where do they come from? Prevalence and characteristics of nascent entrepreneurs. *Entrepreneurship & Regional Development, 12*, 1–23.

Delmar, F., Davidsson, P., & Gartner, W. B. (2003). Arriving at the high-growth firm. *Journal of Business Venturing, 18*(2), 189–216.

Delmar, F., & Shane, S. A. (2004). Legitimating first: Organizing activities and the survival of new ventures. *Journal of Business Venturing, 19*, 385–410.

Dimov, D. (2010). Nascent entrepreneurs and venture emergence: Opportunity confidence, human capital, and early planning. *Journal of Management Studies, 47*(6), 1123–1153.

Eckhardt, J., Shane, S., & Delmar, F. (2006). Multistage selection and the financing of new ventures. *Management Science, 52*(2), 220–232.

Elert, N., Andersson, F. W., & Wennberg, K. (2015). The impact of entrepreneurship education in high school on long-term entrepreneurial performance. *Journal of Economic Behavior & Organization, 111*, 209–223.

Fayolle, A., & Wright, M. (Eds.). (2014). *How to get published in the best entrepreneurship journals: A guide to steer your academic career*. Cheltenham, UK: Edward Elgar Publishing.

Fichman, M., & Cummings, J. N. (2003). Multiple imputation for missing data: Making the most of what you know. *Organizational Research Methods, 6*(3), 282–308.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Gordon, S. R. (2012). *Dimensions of the venture creation process: Amount, dynamics, and sequences of action in nascent entrepreneurship*. Doctoral dissertation, Queensland University of Technology, Brisbane.

Guzman, J., & Stern, S. (2015a). *Nowcasting and placecasting: Entrepreneurial quality and performance. Working Paper 20954*. National Bureau of Economic Research.

Guzman, J., & Stern, S. (2015b). Where is Silicon Valley? *Science, 347*(6222), 606–609.

Hak, T., Jaspers, F., & Dul, J. (2013). The analysis of temporally ordered configurations: Challenges and solutions. In P. C. Fiss, B. Cambré, & A. Marx (Eds.), *Configurational theory and methods in organizational research. Research in the sociology of organizations* (Vol. 38, pp. 107–127). Bingley, UK: Emerald.

Hubbard, R., & Lindsay, R. M. (2013). The significant difference paradigm promotes bad science. *Journal of Business Research, 66*(9), 1393–1397.

Johnson, A. R., van de Schoot, R., Delmar, F., & Crano, W. D. (2014). Social influence interpretation of interpersonal processes and team performance over time using Bayesian model selection. *Journal of Management*. doi:10.1177/0149206314539351.

Klein, K. J., & Kozlowski, W. J. (2000). *Multilevel theory, research, and methods in organizations*. San Francisco, CA: Jossey-Bass Inc.
Kozlowski, S. W. J., Chao, G. T., Grand, J. A., Braun, M. T., & Kuljanin, G. (2013). Advancing multilevel research design capturing the dynamics of emergence. *Organizational Research Methods, 16*(4), 581–615.
Kwon, S.-W., & Arenius, P. (2010). Nations of entrepreneurs: A social capital perspective. *Journal of Business Venturing, 25*(3), 315–330.
Lechner, M. (2011). The estimation of causal effects by Difference-in-Difference methods. *Foundations and Trends in Econometrics, 4*(3), 165–224.
Liao, J., Welsch, H., & Tan, W. L. (2005). Venture gestation paths of nascent entrepreneurs: Exploring the temporal patterns. *Journal of High Technology Management Research, 16*(1), 1–22.
Lockett, A., Wiklund, J., Davidsson, P., & Girma, S. (2011). Organic and acquisitive growth: Re-examining, testing and extending Penrose's growth theory. *Journal of Management Studies, 48*(1), 48–74.
Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2*(4), 371–402.
Muthén, B. O., & Khoo, S.-T. (1999). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, 10*, 73–101.
Naldi, L., & Davidsson, P. (2014). Entrepreneurial growth: The role of international knowledge acquisition as moderated by firm age. *Journal of Business Venturing, 29*(5), 697–703.
Newman, D. A. (2014). Missing data five practical guidelines. *Organizational Research Methods, 17*(4), 372–411.
O'Boyle, E. H., & Aguinis, H. (2012). The best and the rest: Revisiting the norm of normality of individual performance. *Personnel Psychology, 65*, 79–119.
Rousseau, D. M. (1985). Issues of level in organizational research: Multi-level and cross-level perspectives. *Research in Organizational Behavior, 7*, 1–37.
Samuelsson, M. (2004). *Creating new ventures: A longitudinal investigation of the Nascent venturing process*. Doctoral dissertation, Jönköping International Business School, Jönköping
Samuelsson, M., & Davidsson, P. (2009). Does venture opportunity variation matter? Investigating systematic process differences between innovative and imitative new ventures. *Small Business Economics, 33*(2), 229–255.
Semadeni, M., Withers, M. C., & Certo, S.T. (2014). The perils of endogeneity and instrumental variables in strategy research: Understanding through simulations. *Strategic Management Journal, 35*(7), 1070–1079.
Shepherd, D. A. (2011). Multilevel entrepreneurship research: Opportunities for studying entrepreneurial decision making. *Journal of Management, 37*(2), 412–420.
Yang, T., & Aldrich, H. E. (2012). Out of sight but not out of mind: Why failure to account for left truncation biases research on failure rates. *Journal of Business Venturing, 27*(4), 477–492.

Now that We're Done…

Was it good for you, too? Useful? Worthwhile? Bearable? Since you're part of the non-attrition, the survivor-biased subsample of readers who are still with me at this point, you probably found some reason to read on. I can only hope it was a good one. I regret it if I have annoyed people with my patchy, biased, and incomplete coverage of conceptual and empirical issues in entrepreneurship research, or with the numerous references to my own work. Perhaps you misunderstood my intentions? This book was an attempt to codify the knowledge—if I may so name it—that I have built up over the years; it was decidedly not an attempt to review and present the most central themes of the extant collective knowledge of entrepreneurship scholars. I'm neither smart nor energetic enough to take on the latter task.

Some may have been turned off by the nonacademic, chatty writing style that I apply in places. I regret it if you felt that the form deducted from the contents, but I remain confident that there are other readers who appreciate that we do not always have to keep a dead serious tone even if we are dead serious about the message. My intention was to make the contents more digestible. There may be a need for that—people have remarked more than once that my texts tend to be jam-packed with information. So even if I used a catchy phrase here and there, I trust you did not find a lot of empty BS.

I have put a lot of time and intellectual effort into the first two chapters of the book—those on the entrepreneurship phenomenon and the corresponding field of research. The same goes for Chap. 8. While respecting that some researchers could not care less, or find this type of effort futile, I hope others will appreciate this attempt to reconcile some of the issues that plague the field and feel inspired to continue to contribute to this debate and development.

I hope Chap. 3 gave at least some readers an appreciation of the value of theory as well as of the perils of overemphasizing theoretical contributions as the sole goal of research studies. One of the revisions I am particularly pleased with is that of Chap. 4. I hope researchers who are still in the formative stages of their career will find it useful. Well, I hope, of course, that you found every chapter useful and that you are now a significance skeptic and a great fan of replications. And that you will *act* on those convictions. We can't leave it to "the others."

P. Davidsson, *Researching Entrepreneurship*, International Studies in Entrepreneurship 33, DOI 10.1007/978-3-319-26692-3

In the long chapters, Chaps. 4, 5, and 6, I dealt with many issues of importance far outside the realm of entrepreneurship research. This was because I find them missing in most standard treatments. However, Chaps. 4, 5, and 6 also include a lot of practical, hands-on design advice that in a sense is the core of this book. Although I have tried to give advice that has applicability for a broad range of possible entrepreneurship questions, it is unavoidable that the examples become so specific that they best serve as inspiration rather than being of immediate use as exemplars for the reader to follow.

I also hope you appreciated the entirely new chapters, Chaps. 7 and 8. They highlight issues that I feel lead to a chronic state of slower-than-it-needs-to-be development of our field. I know, I know—if I look up and stand back, I can see that our field is undergoing great development, but being content and resting on laurels is just not my nature. Chapter 9 is one of my own favorites and I hope it will inspire readers to appreciate more the value of replications. Finally, Chap. 10 was merely a briefing on analysis method implications—although there is a message in there that at least I find very important.

In order not to make a fool of myself, or step on too many toes, I should, of course, have made trusted colleagues preview this manuscript, or parts thereof, to a much greater extent than I have done. However, although I nowadays appreciate the quality-enhancing effect of the peer review process ***much*** more than did the cocky, young version of myself, I maintain that it can also make a work lose some of its individual voice. I'm neither cautiously natured nor nontenured, and in this particular case, I really wanted to keep my voice. Albeit shaped in part by the environment, what you found in this book were my ideas, my experiences, my advice, my convictions, my misconceptions, my omissions, my shortcomings, my rudeness, and my outright errors. Thank you so much for bearing with me! ***P.D.***