ADVANCES IN PSYCHOLOGY

80

Mental Images in Human Cognition

Robert H. Logie
Michel Denis
Editors

North-Holland

MENTAL IMAGES
IN
HUMAN COGNITION

ADVANCES IN PSYCHOLOGY

80

Editors:

G. E. STELMACH

P. A. VROON

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

MENTAL IMAGES
IN
HUMAN COGNITION

Edited by

Robert H. LOGIE
University of Aberdeen
United Kingdom

Michel DENIS
Université de Paris-Sud
Orsay, France

1991

NORTH-HOLLAND
AMSTERDAM • NEW YORK • OXFORD • TOKYO

NORTH-HOLLAND
ELSEVIER SCIENCE PUBLISHERS B.V.
Sara Burgerhartstraat 25
P.O. Box 211, 1000 AE Amsterdam, The Netherlands

Distributors for the United States and Canada:
ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
655 Avenue of the Americas
New York, N.Y. 10010, U.S.A.

ISBN: 0 444 88894 2

Printed in The Netherlands

Preface

> ".....when I produce an image of Peter, it is Peter who is the object of my actual consciousness. As long as that consciousness remains unaltered, I could give a description of the object as it appears to me in the form of an image but not of the image as such. To determine the properties of the image as image I must turn to a new act of consciousness: I must reflect.....It is this reflective act which permits the judgment 'I have an image'" (Sartre, 1950).

This book represents the research efforts of individuals whose scientific expertise lies in reflection on what Sartre described as reflective acts. Theory in the cognitive psychology of mental imagery endeavours not only to describe the contents and nature of mental imagery, but also to understand the underlying functional cognition. Psychologists need not rely solely on the techniques of introspection, and the last two decades have seen highly creative developments in techniques for eliciting behavioural data to be complemented by introspective reports.

This level of sophistication has provided singular insights into the relationship between imagery and other consequential and universal aspects of human cognition: perception, memory, verbal processes and problem solving. A significant additional benefit of such developments has been the recognition that imagery, despite its ubiquitous nature, differs between individuals both in prevalence and in kind. Moreover the dramatic rise in cognitive science, (to include cognitive neuroscience and computational modelling as well as cognitive psychology), over the last decade has spread the sources of converging evidence to neuroscience. These exciting developments in techniques provide the additional potential for integrating our understanding of cognitive function with our understanding of neuroanatomy and of computer science.

All of these relationships, developments and issues are embraced by the included chapters, not least by some of the most distinguished authors in imagery research. This represents a truly international effort. It highlights an era of progressive collaboration among European cognitive psychologists, while recognising the substantial influence of North American theorists and researchers.

This book arose from The Third European Workshop on Imagery and Cognition held at the University of Aberdeen, United Kingdom in August 1990. The chapters represent the topics covered at that meeting, but this book was undertaken as a separate enterprise not necessarily restricted by the oral presentations at the conference. To ensure a high level of academic quality, all of the chapters were subjected to peer review. Our intention was not for the book to act as a set of conference proceedings, but rather to form a source text for imagery researchers. The list of contributors displays the diversity of countries in which this research is based, and builds on the seminal efforts of two previous workshops; the first in 1986 at the Université de Paris-Sud, France, and the second in 1988 at the Università di Padova, Italy. We are grateful to the organisers and participants of those meetings for stimulating what continues to be a Zeitgeist of international academic dialogue and exchange.

Reference

Sartre, J-P. (1950). *The Psychology of the Imagination.* New York: Rider (in English translation).

Robert H. Logie
Michel Denis

Acknowledgements

We should like to thank our co-organisers of the Third European Workshop on Imagery and Cognition, Cesare Cornoldi and Mark McDaniel. We are grateful for the support given to that meeting by the Cognitive Psychology Section of The British Psychological Society, and by the local organisers. The production of this book in camera-ready copy was greatly eased by the help of Peter Bates, Caroline Green, and Val Wynn of the Psychology Department University of Aberdeen. Their contribution is much appreciated. We are extremely grateful for financial support from the United States Office of Naval Research (Europe). However, the content of this book does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

Finally, we greatly appreciate the help of those individuals who acted as referees for the chapters in this book:

Jeanine Blanc-Garin
Daniel Chartier
Marguerite Cocude
Martin Conway
Cesare Cornoldi
Jan Deregowski
Manuel De Vega
Nick Ellis
Johannes Engelkamp
Tore Helstrup
Ken Gilhooly
Jacques Lautrey
Marc Marschark
Denis Parker
John Richardson
Norman Wetherick

This Page Intentionally Left Blank

Contents

List of Contributors

Rita E. Anderson - Department of Psychology, Memorial University of Newfoundland, St. John's NF, Canada A1B 3X9

Allessandro Antonietti - Departimento di Psicologia, Catholic University of Milano, I-20123 Milano, Italy

Christa Artner - Neurologische Universitätsklinik, Lazarettgasse 14, A 1090 Wien, Austria

Raymond Baillargeon - York University, Toronto, Canada

Kathryn Bruce - Department of Psychology, University of Lancaster, Lancaster LA1 4YF, United Kingdom

Martin Conway - Department of Psychology, University of Lancaster, Lancaster LA1 4YF, United Kingdom

Lynn Cooper - Psychology Department, Columbia University - New York, New York 10027, USA

Cesare Cornoldi - Dipartimento di Psicologia Generale, Università di Padova, 35139 Padova, Italy

Graham Dean - Department of Psychology, University of Lancaster, Lancaster LA1 4YF, England

Rossana De Beni - Dipartimento di Psicologia Generale, Università di Padova, 35139 Padova, Italy

Michel Denis - Centre D'Études de Psychologie Cognitive, Université de Paris-Sud, Orsay, France

Manuel De Vega - Departamento de Psicologia Cognitiva, Universiidad de La Laguna, Tenerife, Spain

José M. Díaz - Departamento de Psicologia Cognitiva, Universiidad de La Laguna, Tenerife, Spain

Helen Duce - Department of Psychology, University of Lancaster, Lancaster LA1 4YF, United Kingdom

Gilles O. Einstein - Department of Psychology, Furman University, Greenville, South Carolina 29163, USA

Nick Ellis - Department of Psychology, University College of North Wales, Bangor LL57 2DG, Wales

A. Elzinga-Plomp - Vakgroep Psychonomie, Rijksuniversiteit Utrecht, Postbus 80.140 3508 Utrecht, The Netherlands

Johannes Engelkamp - FR Psychologie, Universität des Saarlandes, D-6600 Saarbrücken 11, Germany

D. F. Giusberti - Department of Psychology, Viale Berti Pichat 5, 40127 Bologna, Italy

Georg Goldenberg - Neurologische Universitätsklinik, Lazarettgasse 14, A 1090 Wien, Austria

Tore Helstrup - Department of General Psychology, Cognitive Unit, University of Bergen, N-5007 Bergen, Norway

J.H. Hennink - Vakgroep Psychonomie, Rijksuniversiteit Utrecht, Postbus 80.140 3508 Utrecht, The Netherlands

Charles Huffman - Department of Psychology, University of North Carolina at Greensboro, Greensboro NC 27412 USA

Janice Johnson - York University, Toronto, Canada

Hank Kahney - The Psychology Unit, Faculty of Social Sciences, The Open University, Milton Keynes, United Kingdom

Stephen M. Kosslyn - Department of Psychology, Harvard University, Cambridge MA 02138, USA

Even Loarer - Centre de Recherche de l'INETOP, 41 rue Gay-Lussac, CNAM, 75005 Paris, France

Robert H. Logie - Department of Psychology, University of Aberdeen, Aberdeen AB9 2UB, United Kingdom

W.A. van Loon-Vervoorn - Vakgroep Psychonomie, Rijksuniversiteit Utrecht, Postbus 80.140 3508 Utrecht, The Netherlands

Clelia Marchetti - Clinica del Lavoro, 28010 Veruno, Novara, Italy

Marc Marschark - Department of Psychology, University of North Carolina at Greensboro, Greensboro NC 27412, USA

Francesco S. Marucci - Dipartimento di Psicologia, Via degli Apuli 8, 00185 Roma, Italy

Manfredo Massironi - Via Forcellini 14, 35128 Padova, Italy

Giuliana Mazzoni - Dipartimento di Psicologia Generale, Università di Padova, 35139 Padova, Italy

Dr. Mark McDaniel - Department of Psychology, University of Purdue, West Lafayette, Indiana 47907, USA

Gilbert Mohr - FR Psychologie/Bau 1.1, Universität des Saarlandes, D-6600 Saarbrücken 11, Germany

Sergio Morra - Dipartimento di Psicologia Generale, Universitá di Padova, Piazza Capitaniato 3, 35139 Padova, Italy

Peter Morris - Department of Psychology, University of Lancaster, Lancaster LA1 4YF, England

Juan Pascual-Leone - York University, Toronto, Canada

Patrick Péruch - Université d'Aix-Marseille II, URA CNRS Cognition & Mouvement , IBHOP, 13388 Marseille Cedex 13, France

Ivo Podreka - Neurologische Universitätsklinik, Lazarettgasse 14, A 1090 Wien, Austria

Dr. J.G. Quinn - Department of Psychology, University of St. Andrews, St. Andrews KY16 9JU, Scotland

Dr. Daniel Reisberg - Psychology Department, Reed College, Portland Oregon 97202, USA

Dr. John Richardson - Department of Human Sciences, Brunel University, Uxbridge UB8 3PH, United Kingdom

Frédérique Robin - Centre D'Etudes de Psychologie Cognitive, Université de Paris-Sud, Orsay, France

Pertii Saariluoma - University of Helsinki, Fabianinkatu 28, 00100 Helsinki 10, Finland

Alain Savoyant - Université d'Aix-Marseille II, CNRS, Cognition et Mouvement, IBHOP, Marseille, France

J. David Smith - New School for Social Research, New York, New York, USA

Roxann Thomson - Massachussetts Institute of Technology, Cambridge MA 02142, USA

John Warner - Department of Psychology, University of North Carolina at Greensboro, Greensboro NC 27412 USA

Norman E. Wetherick - Department of Psychology, University of Aberdeen, Aberdeen AB9 2UB, United Kingdom

Meg Wilson - University of California, Berkeley, California, USA

Part 1

Imagery and Perception

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 1

Dissociable aspects of the mental representation of visual objects*

Lynn A. Cooper
Columbia University, New York NY, USA

Much of our ordinary behavior in the environment is guided by representations of the static, dynamic, and relational aspects of the visual objects that surround us. These mental representations of visual objects and events serve many important functions. Computations based upon such representations allow us to locomote in a relatively smooth and uninterrupted fashion, by permitting prediction of the consequences of both object and observer movement. The partial information provided by momentary glances at surfaces of objects constitutes the raw material from which rough constructions of the continuing structure of hidden portions of three-dimensional objects can be made. The effectiveness of such constructed mental representations is evidenced by the fact that rarely do we collide with the concealed surfaces of three-dimensional objects or register surprise as previously hidden aspects of object structure are revealed through patterns of eye fixation or movement of the observer. Moreover, we generally accomplish acts such as these in the absence of any conscious identification or recognition of the objects in our surroundings. There are, in addition, a variety of quite different perceptual tasks for which other sorts of analyses of information about objects and events are required. These tasks include recognizing an object or its spatial location as familiar, determining the meaningful identity or function of an object from its visual appearance, and distinguishing one particular object or instance of a set of objects from another.

*The research reported in this chapter was supported in part by Air Force Office of Scientific Research Grant 90-0187 to Daniel L. Schacter and Lynn A. Cooper and in part by Office of Naval Research Contract N00014-81-C-0532 to Lynn A. Cooper.

In general, the many perceptual and cognitive functions served by mental representations of visual objects can be grouped into two distinguishable categories. The first includes the seemingly automatic and unconscious computations involved in anticipating the continuing structure of objects, particularly as the objects or the observer move in the environment or transform in space. These sorts of activities seem to require information about the layout of objects in the environment and, importantly, about the global structure of individual objects. The second category of perceptual/cognitive functions are those that rely on conscious knowledge of the identity of an object, in the senses both of the object's meaning and of the characteristics that distinguish recognizable objects one from the other.

In this paper, I argue that the information for, or aspects of, mental representations of visual objects that support these two sorts of perceptual/cognitive activities can be dissociated empirically via laboratory experiments with intact, adult individuals. Furthermore, such dissociations as well as other experimental techniques can be used to understand at a relatively detailed level the nature of the information preserved in the underlying representations. After presenting the results of relevant experimental work, I sketch an evolving picture of the properties of these dissociable aspects of mental representations of visual objects. I conclude by suggesting that the concept of "mental image" be reserved for those aspects of representations of objects of the second sort: viz., those that are analogous to the perceptual activities of recognition and determination of object identity.

At the core of the present discussion are two related lines of research addressed to questions about the nature of the mental representation of visual objects. The first line of work seems, superficially, somewhat far removed from perceptual considerations in that it seeks to unite questions about object representation with fundamental issues in the study of human memory. This area of research demonstrates that dissociable aspects of mental representations of visual objects support different kinds of cognitive tasks. It also provides an experimental technique for investigating the nature of the information preserved in the dissociable subsystems. The second line of work directly addresses the construction and recognition of representations of three-dimensional objects from two-dimensional information. A particular experimental phenomenon is used to argue that the generation of three-dimensional object representations is in some sense "obligatory", in that such representations are constructed even when information about two-dimensional shape alone would suffice for performing the required task. This phenomenon is also used to explore the question of the view-specificity of constructed three-dimensional object models. Below, each of these programs of research is described in turn.

IMPLICIT AND EXPLICIT MEMORY FOR REPRESENTATIONS OF VISUAL OBJECTS

As indicated above, the first line of experimental work combines questions about object recognition with more general issues in human memory. This work is being done in collaboration with Daniel L. Schacter, as well as several students and colleagues who are noted in connection with appropriate experiments. The central questions that we are asking concern how perceptual information about properties of visual objects is represented mentally and how this represented information is accessed when we remember things about them. To anticipate a bit, the general view that we are proposing is that systems for representing the global, structural relations among components of an object are functionally separable from systems for representing other sorts of information about object properties. We infer this functional separation from an experimental dissociation in which certain kinds of memory tasks are supported by or have access to specific information about the three-dimensional relations that define the structure of an object, whereas other tests of memory are supported by various more flexible sources of information about object properties.

The experiments described here are motivated by a number of different considerations. First, some general ideas concerning the perceptual representation of information about the structure of three-dimensional visual objects, described elsewhere (e.g., Cooper, 1988; 1989), coupled with the observations about aspects of object representation described above, provide one theoretical context for the research. Second, some recent findings in the literature on memory for verbal materials (e.g., Graf & Schacter, 1985; Tulving & Schacter, 1990) provide another conceptual framework, as well as the outlines of an empirical approach to examining the representation of nonverbal information. Finally, certain case reports and experimental work in clinical and cognitive neuropsychology (e.g., Farah, Hammond, Levine, & Calvanio, 1988; Riddoch & Humphreys, 1987a) provide concrete and suggestive instances in which dissociations between aspects of object representations (visual versus spatial; semantic/functional versus structural, respectively) have been obtained in persons with particular types of brain injury.

The relevant findings in the domain of verbal memory processes centre on Schacter's (1987) distinction between "implicit" and "explicit" forms of remembering. Explicit memory refers to our intentional or conscious recollection of recent experiences; the familiar sort of episodic memory that is generally expressed on standard tests of recognition and recall. Implicit memory, in contrast, refers to the unintentional retrieval of previously acquired information on tasks that do not require conscious recollection of a specific previous experience. This latter sort of memory is generally inferred from performance facilitation or priming effects on tasks like word identification, word fragment completion, and lexical decision (e.g., Graf & Schacter, 1985; Schacter, 1987). Experimentally-induced dissociations between tasks that tap implicit and explicit remembering have been reported by a number of investigators and for a variety of manipulations that differentially affect the two ways of assessing memory (for a review, see Richardson-Klavehn & Bjork, 1988). Perhaps even more impressive evidence for the dissociation comes from the finding that subject populations that exhibit very poor memory performance on explicit tests, including profoundly amnesic patients and elderly adults, can sometimes show virtually normal retention on a number of implicit memory tasks (e.g., Graf, Squire, & Mandler, 1984; Schacter, 1985; Warrington & Weiskrantz, 1974).

It is important to note that the tasks that show priming, or evidence for implicit memory, are all ones with strong perceptual components, even though they have been studied primarily using verbal materials. In the domain of nonverbal object recognition, another set of memory dissociation findings relevant to the present work comes from reports of patients with certain forms of visual agnosia in which structural knowledge of objects is intact, but the patients can neither name nor recognize the objects (Riddoch & Humphreys, 1987a). These last observations led Schacter and me, along with Suzanne Delaney, (Schacter, Cooper, & Delaney, 1990a; 1990b) to the following hypothesis about visual object representation in the normal memory system: *Information about the relationships among components of an object, or its global three-dimensional structure, might be represented in a system accessible primarily to implicit tests of memory, while other information about objects, in particular their meaning and function, might be available to retrieval processes underlying explicit memory performance.* How might such a hypothesis be evaluated in a normal population, in the absence of a clinical syndrome suggestive of a representational dissociation?

The task that we devised to assess implicit memory for three-dimensional objects, an "object decision task", requires that a line

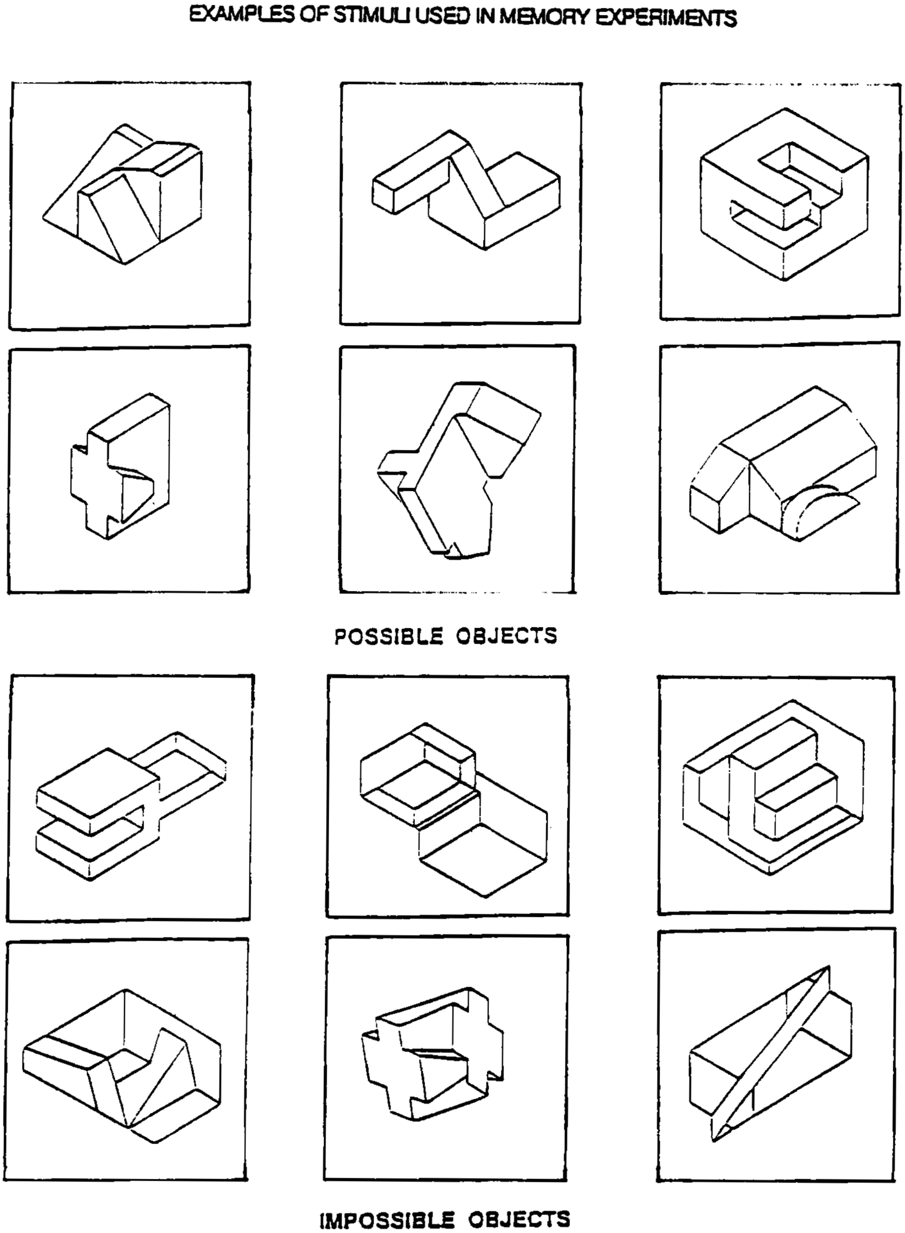

Figure 1. Examples of possible (at top) and impossible (at bottom) objects used as stimuli in the memory experiments. Adapted from Schacter, Cooper, & Delaney, 1990a.

drawing of an object be classified as to whether or not it depicts a structure that could exist in the three-dimensional world. That is, subjects studied drawings of unfamiliar three-dimensional constructions like those shown in Figure 1. Some of the drawings depicted *possible objects;* objects whose surfaces and edges are connected in such a way that they could potentially exist in the world. Other drawings displayed *impossible objects,* objects that contain subtle surface and edge violations that would make it impossible for them to exist as actual three-dimensional structures (cf. Penrose & Penrose, 1958). To assess implicit memory for these unfamiliar objects, subjects were given 100ms exposures to drawings of either previously-studied or non-studied possible or impossible structures. They then had to determine whether or not each object would be possible. This object decision task can be regarded as an implicit test of memory in that it does not make explicit reference to, or require conscious recollection of, any specific previous encounter with a presented object. So, if object decision performance shows facilitation as a result of prior study of the test objects, then implicit memory for unfamiliar, three-dimensional structures is demonstrated.

Manipulation of the conditions under which objects are initially studied or encoded is a crucial part of the logic of these experiments. Briefly, we reasoned that to perform the object decision task, information must be extracted about the three-dimensional structure of the object. That is, an object can only be judged "possible" following an analysis of the structural relations among its various components. We reasoned further that facilitation of object decision performance should only be obtained following study tasks that lead to the encoding of information about global, three-dimensional object structure. In contrast, study tasks that do not promote the acquisition of structural information should not facilitate subsequent object decision performance. To evaluate the specificity of this hypothesized implicit memory for unfamiliar objects, in all experiments we compare potential priming or object-decision facilitation effects with explicit remembering of target objects, as assessed by "yes/no" recognition. Our expectation in beginning this line of research was that considerable levels of recognition memory should be obtained following a variety of study tasks, as long as those tasks enabled subjects to acquire distinctive information about each studied object. However, while *recognition processes* should be able to make use of different types of information, we expected that *implicit memory* should only be observed following the acquisition of information specifically about the three-dimensional structure of an object.

Consider the basic experimental situation that we used to demonstrate a dissociation between representation and retrieval of structural vs. non-structural information about visual objects. Subjects studied sets of objects containing both possible and impossible items under one of two conditions. The *structural* encoding condition was designed to draw attention to information about the global structure and relations among parts of the objects. In this condition, subjects viewed each object for 5 secs, and they were asked to determine whether the object appeared to be facing to the left or to the right. This encoding task was compared with a study condition requiring the generation of *elaborative* encodings of the objects. Subjects examined each object and were asked to think of something familiar that the object reminded them of most. This task requires relating the unfamiliar objects to pre-existing knowledge structures, i.e., producing the distinctive and semantically rich encodings that are known to enhance explicit memory performance.

If our idea about functional separation of systems for representing structural and non-structural (in this case, presumably semantic) knowledge about objects has merit, then the following predictions can be made about the outcome of this experiment: The elaborative encoding task should yield significantly higher levels of recognition than the left/right task, because of the known relation between distinctive encoding of an item and level of explicit memory performance. However, these elaborative encodings, though useful for recognition, should not be expected to enhance object decision judgments relative to the left/right encoding task. This prediction is based on the idea that the left/right task promotes encoding of a structural representation of an unfamiliar three-dimensional object, and that such a structural representation might support or facilitate performance on the object decision task.

The results of this experiment, shown in Figures 2 and 3, clearly support this prediction and demonstrate the desired dissociation in memory performance attributable to conditions of encoding. Consider, first, performance on the *object decision task,* expressed as percent correct on the "possible/impossible" judgment, as a function of encoding condition (left/right vs. elaborative) , type of test object (possible vs. impossible), and whether or not a tested object had been previously displayed during the encoding or study phase of the experiment (studied vs. non-studied). A priming effect (that is, facilitation of the object decision judgment as a result of prior study), indicating implicit memory, shows up in these data as a significant advantage of decisions made to studied over non-studied objects. Note that there is no evidence that study of impossible objects, under either encoding condition, facilitates

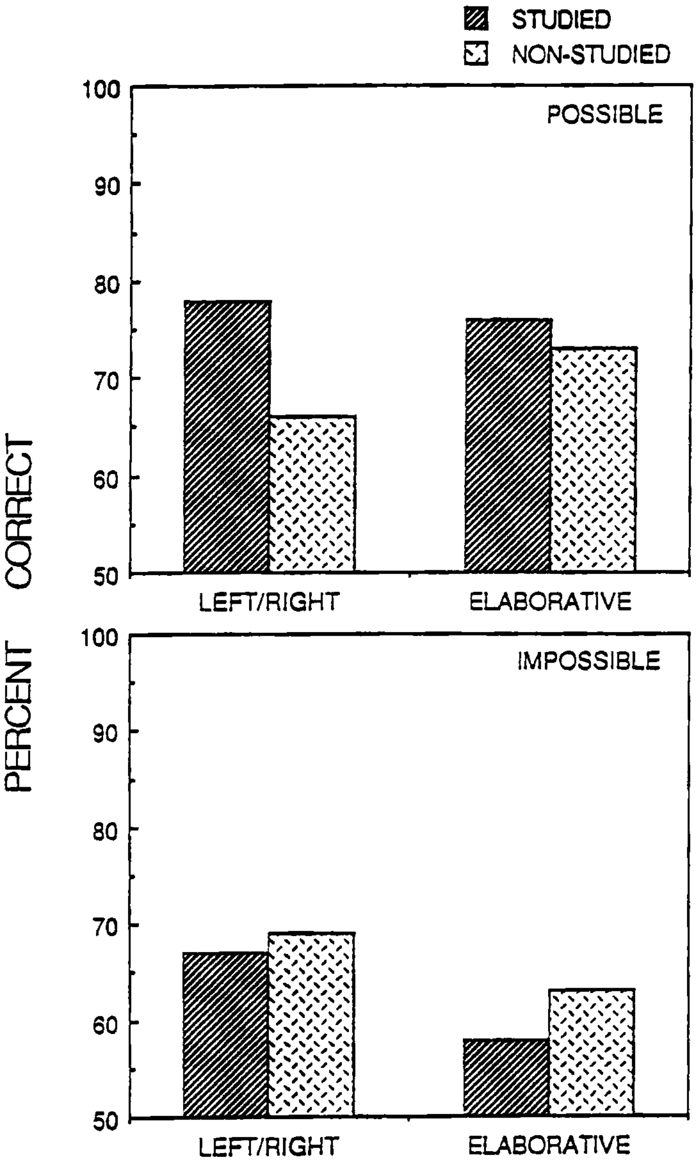

Figure 2. Mean percent correct on the object decision task as a function of encoding condition (left/right vs. elaborative), object type (possible vs. impossible), and prior presentation (studied vs. non-studied). Data from Schacter, Cooper, & Delaney, 1990a, Experiment 2.

object decision performance. This is a persistent result that will be deferred to later discussion; suffice it to say that we believe there are coherent reasons for this lack of effect on judgments of object impossibility that are consistent with our underlying theoretical notions.

For possible objects, the results accord well with our expectations. In the left/right encoding condition, studied objects were classified significantly more accurately than non-studied objects, thereby demonstrating a priming effect. In contrast, no significant priming of "possible" judgments, or evidence for implicit memory, was found under elaborative encoding conditions. This pattern of results demonstrates, first, that implicit memory can be observed for unfamiliar three-dimensional objects, and, second, that such facilitation of object decision performance depends selectively on access to a structural representation of an object (that is, by hypothesis, encoded during the left/right study task).

The key finding of this experiment emerges when *explicit recognition* performance is considered in conjunction with performance on the object decision task. Figure 3 shows recognition, expressed as percentage of hits (studied) and false alarms (non-studied) as a function of the same study and test variables. Concentrating on results for possible objects, a clear interaction between type of encoding task and type of memory test is apparent: While on the object decision task, robust priming is obtained following left/right encoding and no significant priming is observed following elaboration, on the recognition test, performance is significantly enhanced by elaborative encoding compared with left/right encoding. The fact that elaborative encoding improves recognition relative to left/right encoding, while left/right encoding primes object decision performance relative to elaborative encoding, clearly demonstrates that implicit and explicit memory for unfamiliar three-dimensional objects can be dissociated.

But, does this finding provide satisfactory support for the hypothesis that implicit memory tasks access information specifically about the global structural relations among components of an object? The data so far are consistent with a weaker hypothesis, viz., that *any* visual characteristic of an object (a feature or part, as well as a global structural relation) is represented by a system accessible to implicit memory tests, while information about the identity, meaning, or associates of an object is represented by a system accessible to explicit tests of memory. A second experiment more clearly shows that the encoding of specifically structural/relational information facilitates object decision performance.

In this next experiment, subjects again studied possible and impossible objects under one of two conditions. The left/right condition,

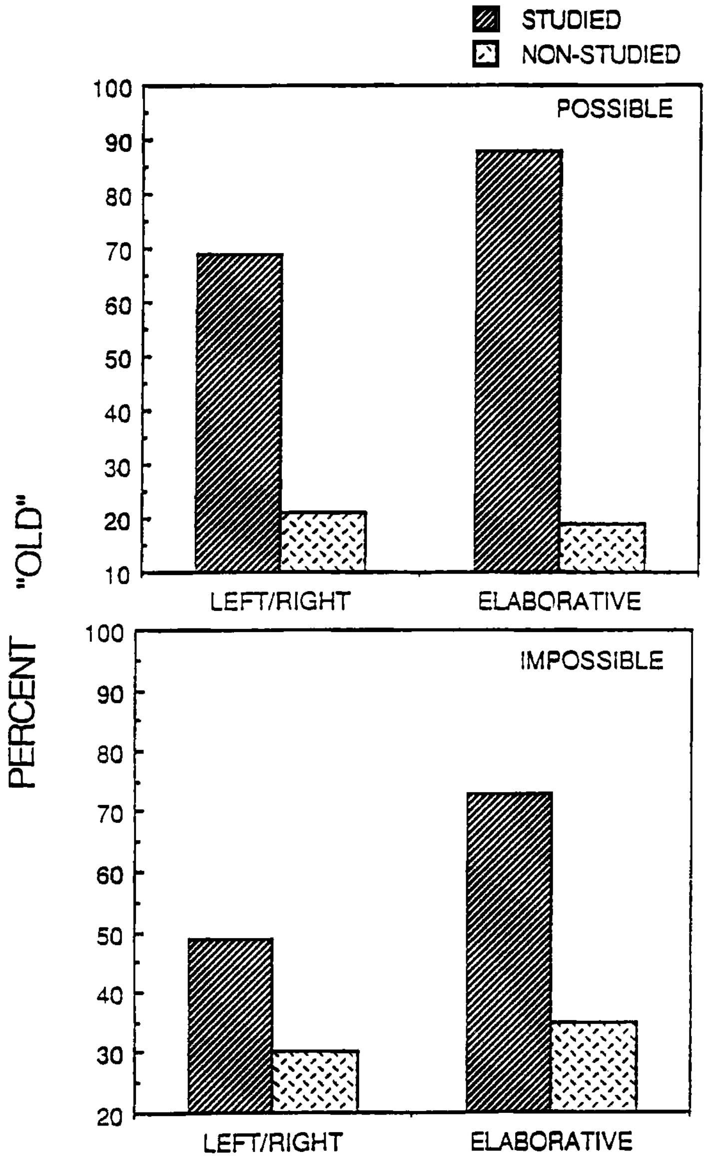

Figure 3. Mean percent "old" recognition responses as a function of encoding condition (left/right vs. elaborative), object type (possible vs. impossible), and prior presentation (studied vs. non-studied). Data from Schacter, Cooper, & Delaney, 1990a, Experiment 2.

designed to promote structural encoding, was now contrasted with a *local encoding condition.* In this study task, a different group of subjects viewed each object for 5 secs, and had to determine whether the majority of the line segments in the object were horizontal or vertical in orientation. This study task was designed to draw attention to local features of the display, without requiring an analysis of structural relations among component surfaces.

The results for the subsequent object decision task are shown in Figure 4. Again, neither condition produces priming of decisions to impossible objects, but the structural or left/right encoding task selectively facilitates judgments for possible objects. No such facilitation is present under conditions of local or horizontal/vertical encoding. Recognition results, shown in Figure 5, do not exhibit a crossover pattern like that obtained when left/right and elaborative encoding are compared, but they do produce a rather different pattern than for the object decision task. Specifically, recognition following left/right encoding does not differ from recognition following horizontal/vertical encoding. The pattern of results from this experiment confirms the dissociation between implicit and explicit memory for unfamiliar three-dimensional objects found in the first experiment. In addition, it demonstrates that this dissociation is specific to *structural representations* of encoded objects, in that only global (left/right) and not local (horizontal/vertical) study tasks produce subsequent priming of object decision performance.

One puzzling result that has been obtained consistently in these and subsequent experiments is the absence of priming for object decision judgments to objects representing physically impossible structures. Our conjecture, which is consistent with the overall theoretical perspective, is that subjects are not able to represent structural impossibility at a global level, relying instead on detection of local surface and edge violations to make an "impossible" decision. There are alternatives to this sort of explanation for the lack of priming for impossible objects that we have recently evaluated in a series of experiments that provide, as well, additional parametric information concerning the range of encoding conditions that produce reliable facilitation of object decision judgments (see Schacter, Cooper, Delaney, Peterson, & Tharan, 1991).

The particular factors that might contribute to the failure to find priming with impossible objects include an implicit bias in task instructions to regard "possible" as the positive response, particular characteristics of the sets of objects used as stimuli, and the idea that memory representations of impossible objects are generally more fragile or degraded than representations of their possible-object counterparts. This last factor is potentially a good candidate for contributing to the

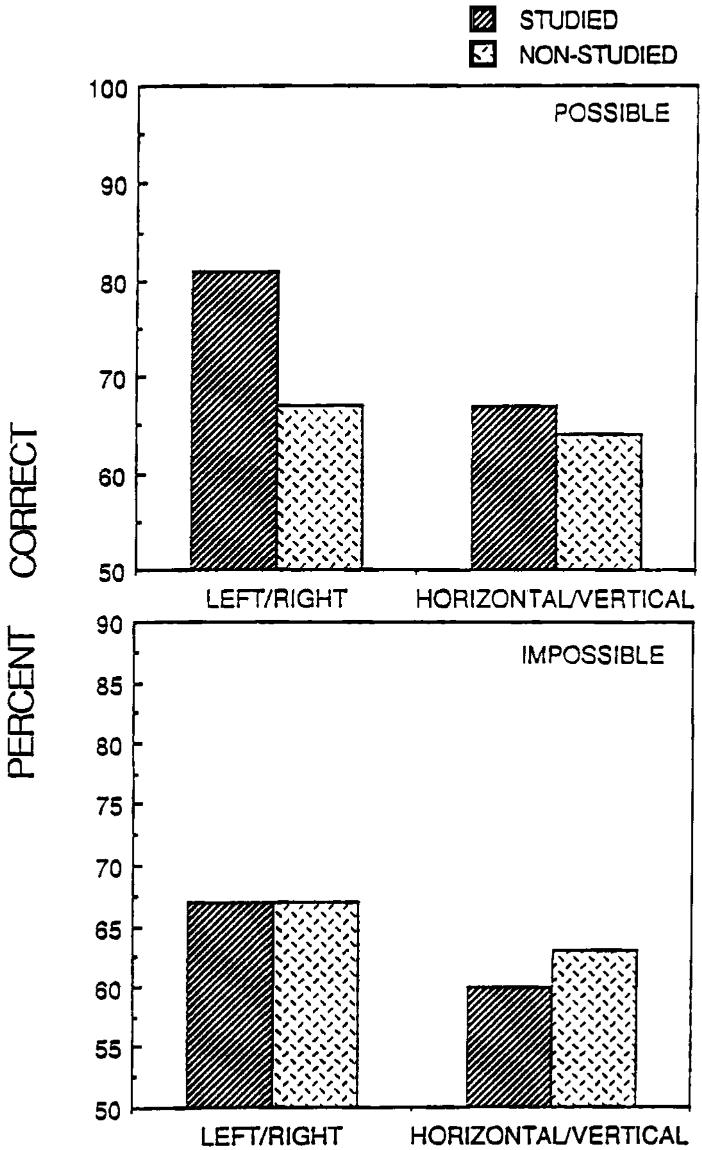

Figure 4. Mean percent correct on the object decision task as a function of encoding condition (left/right vs. horizontal/vertical), object type (possible vs. impossible), and prior presentation (studied vs. non-studied). Data from Schacter, Cooper, & Delaney, 1990a, Experiment 1.

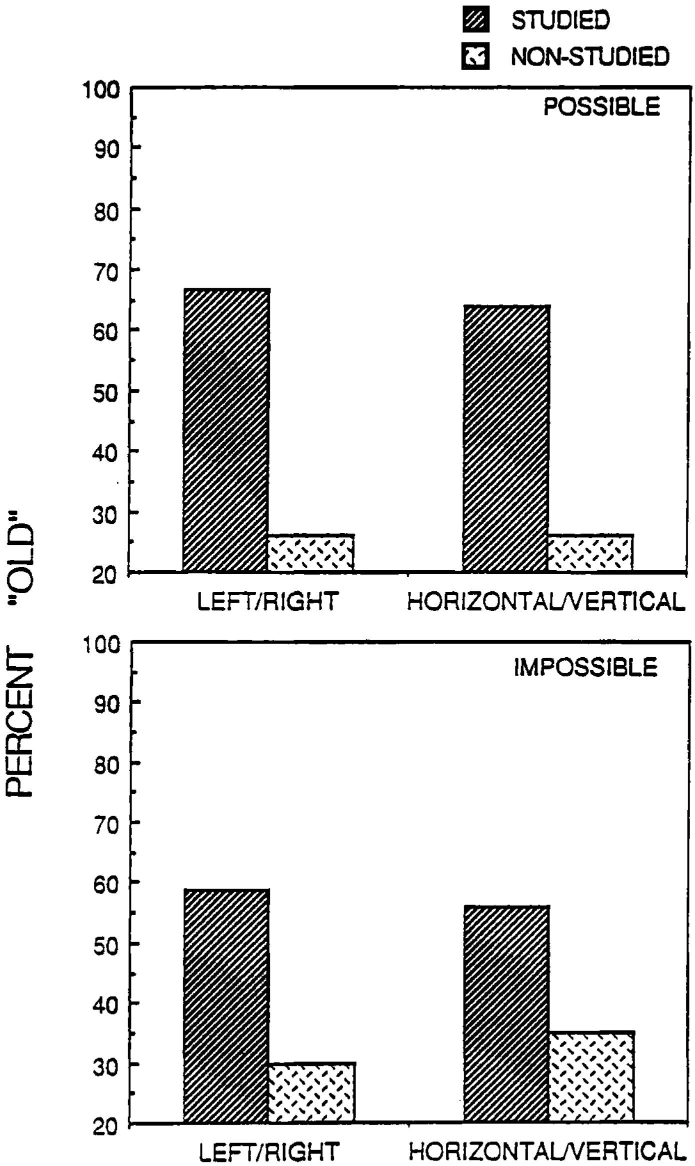

Figure 5. Mean percent "old" recognition responses as a function of encoding condition (left/right vs. horizontal/vertical), object type (possible vs. impossible), and prior presentation (studied vs. non-studied). Data from Schacter, Cooper, & Delaney, 1990a, Experiment 1.

absence of priming, in that levels of explicit memory for possible objects are consistently higher than for impossible objects in the experiments reviewed thus far. To summarize this series of experiments briefly, we have been unable to observe priming of object decision judgments to impossible structures, despite changes in task instructions, stimulus materials, and number, duration, and quality of exposures to materials in the study list. The study-list variables that we have examined include number of repetitions of the set of items (one vs. four), temporal distribution of repeated exposures (one vs. five 1-sec repetitions), and number of distinct structural encodings required (left/right encoding only vs. left/right encoding combined with an encoding task involving classification of items into different categories of three-dimensional objects).

The data from one experiment in this series (in which one 5-sec exposure of the study items, under conditions of left/right encoding, was compared with four separate such exposures) are displayed in Figures 6 and 7. The results for object decision judgments shown in Figure 6 indicate that equivalent levels of priming for possible objects are obtained after one or four full exposures of the study items, but no priming for impossible objects is produced even in the four-exposure condition. Clearly, these data infirm an account of the absence of facilitation for "impossible" judgments based on fragile memory representations of impossible objects. Results for the explicit recognition condition are shown in Figure 7. Again, a dissociation between implicit and explicit memory is apparent; recognition is significantly enhanced when the left/right encoding task is performed four times, as compared with a single study exposure. This is true for both possible and impossible objects, and can be contrasted sharply with the lack of effect of the study-task variable on implicit memory for both sorts of objects. These data, as well as the results of the other experiments in this series (cf., Schacter et al., 1991), strongly suggest that the failure to obtain priming of "impossible" judgments reflects computational constraints on the ability to form structural representations of these objects.

There are several directions in which this program of research on functionally separable forms of memory for representations of visual objects is evolving. Some of these directions include: demonstrating further dissociations between structural and functional, as well as semantic, aspects of object representations; pursuing the distinction between global/structural and local/feature-based information about objects; and, assessing implicit and explicit memory for objects in amnesic patients and individuals with particular forms of brain damage. Still another line of inquiry (described in more detail below) uses the

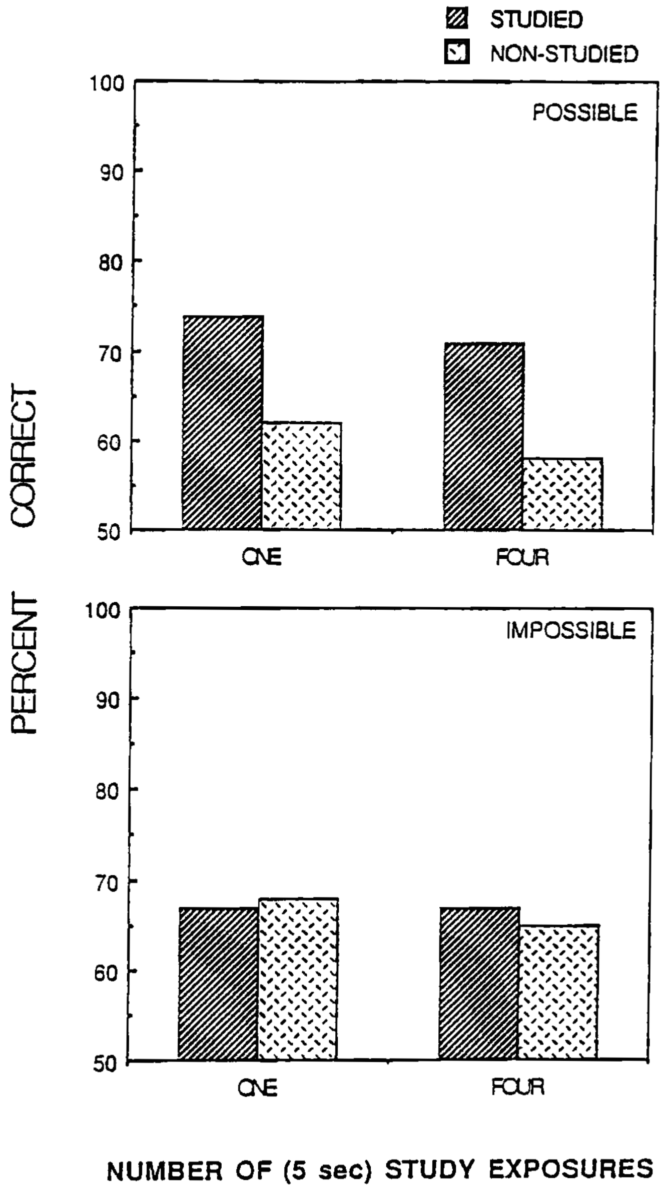

Figure 6. Mean percent correct on the object decision task as a function of number of study exposures (one vs. four), object type (possible vs. impossible), and prior presentation (studied vs. non-studied). Data from Schacter, Cooper, Delaney, Peterson, & Tharan, 1991, Experiment 1.

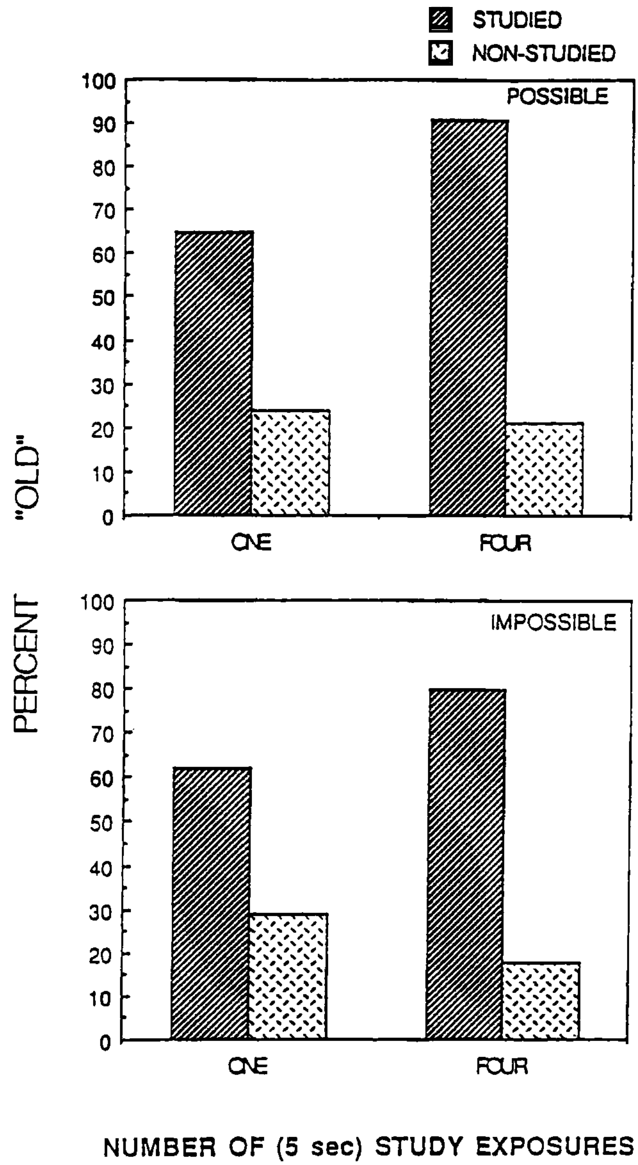

Figure 7. Mean percent "old" recognition responses as a function of number of study exposures (one vs. four), object type (possible vs. impossible), and prior presentation (studied vs. non-studied). Data from Schacter, Cooper, Delaney, Peterson, & Tharan, 1991, Experiment 1.

experimental situation that we have developed as a tool for exploring the nature of the information embodied in structural representations of unfamiliar three-dimensional objects that supports the priming observed in our object decision task.

This research effort has implications not only for the analysis of implicit and explicit memory effects, but also for models of the processes involved in object perception and recognition. One simplified view of the nature of structural representations of objects might hold that only information concerning relationships among component units (regardless of how these units are conceptualized) is preserved in the underlying representations. Under such an account, aspects of visual information *irrelevant* to the coding of component relations should not be represented in or accessible from structural descriptions of objects. If, as we suppose, structural representations of this sort mediate object decision priming, then information concerning such attributes as object size, color, overall reflectance, and contour definition should be unrelated to performance on the object decision task. Variations in other forms of information, e.g., depicted three-dimensional orientation, direction of illumination, occlusion of intersections, and global dimensional structure, might serve to enhance or to reveal certain relations while obscuring others and, thus, could affect object decision performance.

In short, this program of research is examining, experimentally, what forms of information are embodied in structural representations of objects by asking whether study-to-test changes in particular types of information affect object decision, compared with explicit recognition, performance. To the extent that study-to-test changes reduce or modify the magnitude of priming or recognition effects, we can conclude that the system accessed by the memory task *does* represent the form of information in question. If, however, priming or recognition effects persist in the face of study-to-test changes in certain forms of information about objects, then we have reason to believe that the representational system being tapped by the relevant memory test is *not* sensitive to the type of information undergoing change.

The logic of this ongoing set of experiments becomes clear in the context of some concrete examples. We have recently completed a series of experiments (in collaboration with Cassandra Moore and Soledad Ballesteros) asking whether information about object *size* and object *parity,* i.e., standard/reflected orientation is encoded in the mental representation accessed by implicit and explicit memory tasks. These experiments each contained a study phase in which possible and impossible objects were encoded under left/right or structural study conditions. In the experiment manipulating object size, half of the subjects studied objects defined as "small", and the other half of the subjects studied objects enlarged by a factor of 2.5 to 1. Half of the subjects in each encoding group viewed test objects (half studied, half non-studied) in the same size as presented during study, and the other half viewed test objects changed in size. The experiment investigating study-to-test changes in standard/reflected object orientation used a within-subject version of this design; each subject studied standard versions of the objects, and test trials included half standard and half mirror-image versions of the same (as well as non-studied) objects.

Both experiments yielded remarkably similar patterns of performance. The central results for the object decision tasks can be summarized as follows: Substantial priming of responses to possible objects was obtained when studied and tested objects were the same in size or in standard/reflected parity. This is the usual facilitation effect for structural encoding found in previous experiments. The novel finding of the present experiments is that significant priming was also exhibited under conditions of change in object size or in standard/reflected object orientation from study to test. Furthermore, the magnitude of facilitation of judgments to possible objects was roughly equivalent, regardless of whether the size or the parity of the tested object remained the same or was changed from what had been displayed during encoding. For impossible objects, as we have come to expect, no priming on the object decision task was apparent for any combination of stimulus manipulations.

Explicit recognition under the identical encoding and test conditions showed a very different pattern. Recognition of size-changed objects was clearly impaired relative to the same-size condition. Similarly, changes in standard/reflected orientation from study to test resulted in a decrease in the level of recognition for both possible and impossible objects. The finding of recognition impairment for size-transformed stimuli is consistent with results reported by Jolicoeur and his colleagues (e.g., Jolicoeur, 1987; Jolicoeur & Besner, 1987), and it is evidence once again of a dissociation between representations that are accessed by implicit and explicit tests of memory.

Following the logic outlined earlier, we can conclude from the present experiments that information about object size and object parity is *not* preserved in the representational system accessed by object decision judgments and hypothesized to encode specifically relational, structural information about the global organization of objects' surfaces. The mental representations of unfamiliar, three-dimensional objects accessed by explicit recognition, however, appear to preserve both size

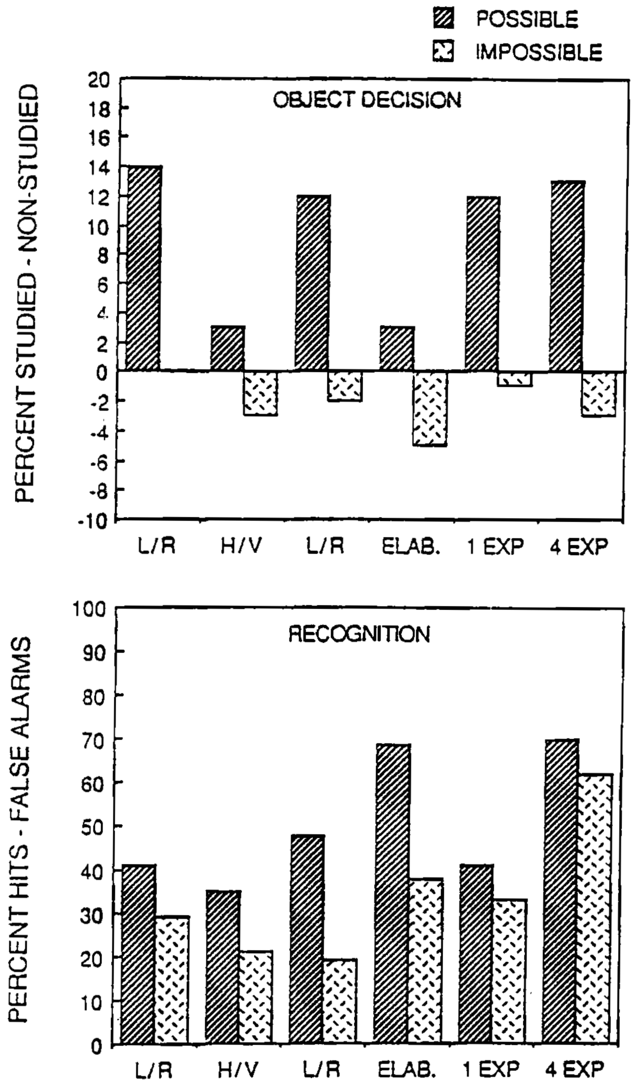

Figure 8. Summary of results of memory experiments involving manipulation of encoding conditions. At the top, object decision data are expressed as differences in percent correct to studied minus non-studied items, as a function of encoding condition and object type. At the bottom, recognition data are expressed as differences between percent hits and false alarms, as a function of the same variables. See text for further explanation.

and standard/reflected orientation as characteristics of an object's distinctive encoding. Hence, performance on recognition is impaired by changes in these aspects of object representations. Experiments currently in progress are assessing the effects of study-to-test changes in a variety of additional stimulus attributes, including the picture-plane orientation of objects and their portrayed positions under rotations in depth other than mirror-image reversal. Results of such studies should help to clarify the role, if any, of orientation, as well as other visual characteristics of objects, in their underlying structural representations.

Figure 8 provides a summary of the findings from all of the experiments described above in which conditions of encoding were manipulated. Performance *differences* are shown - for object decision, between percent correct on studied minus non-studied items, and for recognition, percent hits minus false alarms. Note from Figure 8 that robust object decision facilitation is obtained for possible objects under all conditions requiring structural (left/right) encoding. Recall, as well, that this priming persists, regardless of the relationship between the studied and tested sizes and parities of these unfamiliar three-dimensional objects. (For similar priming results using an object-naming task, see Biederman & Cooper, 1989.)

Explicit recognition exhibits some marked dissociations from object decision performance. Most notably, elaborative encoding and multiple study exposures enhance the level of recognition, while producing no facilitation or an unaffected level of priming on the object decision task. Furthermore, while not illustrated in Figure 8, changes in object size or standard/reflected orientation have a debilitating effect on explicit recognition, but not on object decision, performance.

These experiments, taken together, provide evidence for the representation of information about the global structure of three-dimensional objects that is dissociable from the representation of information about object meaning and associations, as well as more local visual properties like size and orientation of specific line-based features. Ongoing research should help to characterize further the nature of the visual information that is and is not preserved in these structural representations.

INCIDENTAL RECOGNITION OF CONSTRUCTED MENTAL REPRESENTATIONS OF THREE-DIMENSIONAL OBJECTS

The experiments described above provide considerable evidence concerning aspects of the representation of objects that can be

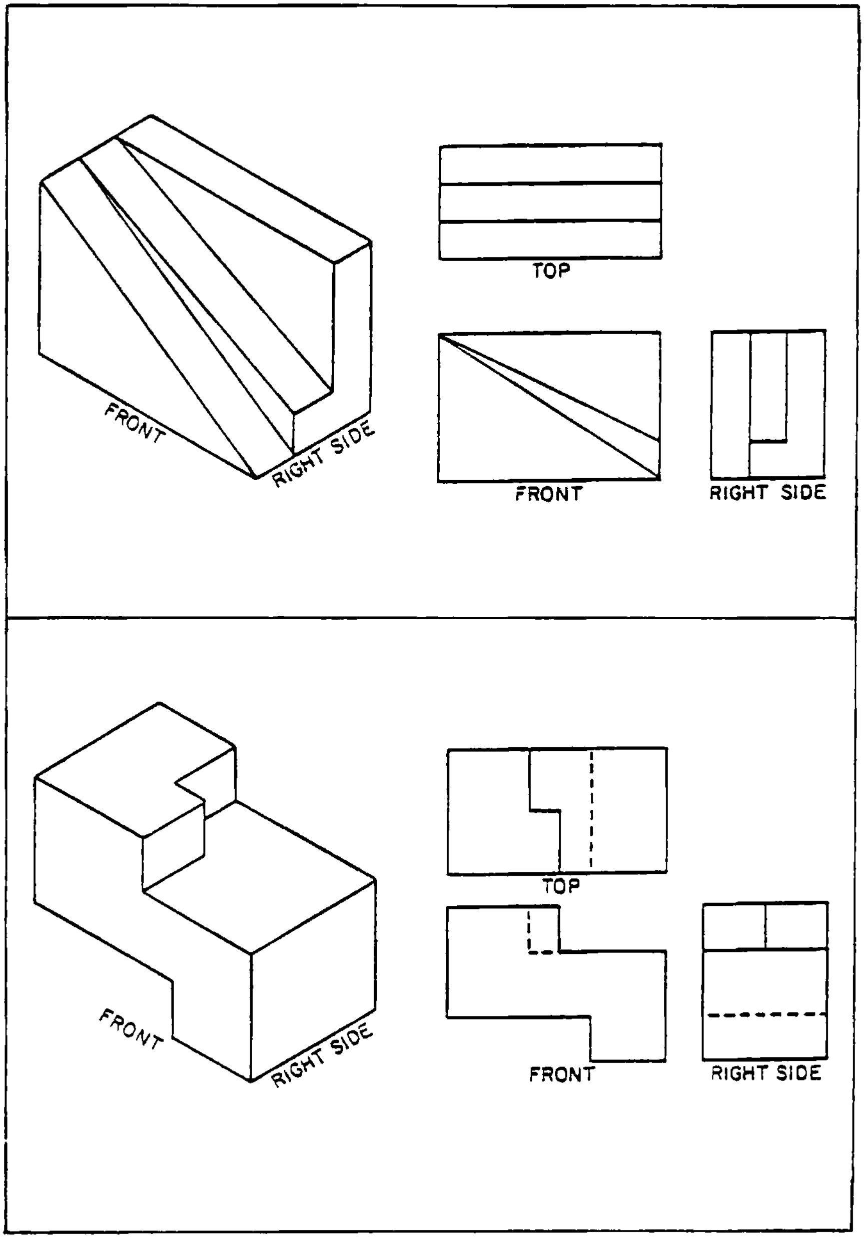

Figure 9. Isometric (left side) and orthographic (right side) projections of two three-dimensional objects. Corresponding object sides are labelled on the two sets of projections.

dissociated experimentally. However, the results of these experiments should *not* be interpreted as showing that mental representations of structural relations in objects are inaccessible to explicit recognition processes. Everyday experience as well as psychological research informs us otherwise. We seem able to recognize objects on the basis of their global structures alone, and conditions also exist under which we can recognize objects (albeit sometimes with considerable difficulty) despite changes in size, from novel points of view, and over other kinds of transformations. The memory research presented above effectively provides an experimental procedure for examining paths of access to information in representations of objects that can be functionally separated.

The second major line of research that I consider here uses explicit, but incidental, recognition techniques to assess the question of whether constructed mental representations of objects embody information about three-dimensional structure that is specific to a particular point of view, or whether the accessible information in such representations is more general, in the sense of being view-independent This question is an important one for theories of object representation; although it is currently an active area of investigation in perceptual and cognitive research, there is little consensus on the answer (see, for various positions, Biederman & Cooper, 1989; Corballis, 1988; Marr, 1982; Rock & Di Vita, 1987; Tarr & Pinker, 1989).

An experimental situation that I (along with Lyn Mowafy and Joe Stevens) have developed to explore this question involves asking subjects to solve problems about the structure of objects from disconnected, two-dimensional views of the structures. Then, we assess via a surprise recognition test, the nature and extent of information about three-dimensional object structure that is available from the mental representations generated during problem solving. More specifically, the methods of depicting objects that we have used are orthographic projection, in which three different views of an object are displayed at right angles to each other by dropping perpendiculars from each side of the object, and isometric projection, in which three surfaces of an object are displayed in a single plane, with the projectors perpendicular to the picture plane and the three principle axes making equal angles with the plane. Figure 9 shows isometric and orthographic views of two objects, with corresponding sides labelled.

The individuals who participate in these experiments are required to have some facility in understanding these forms of projection. Typically, they are undergraduate students who have completed or are enrolled in a beginning course in mechanical engineering. The problem solving or

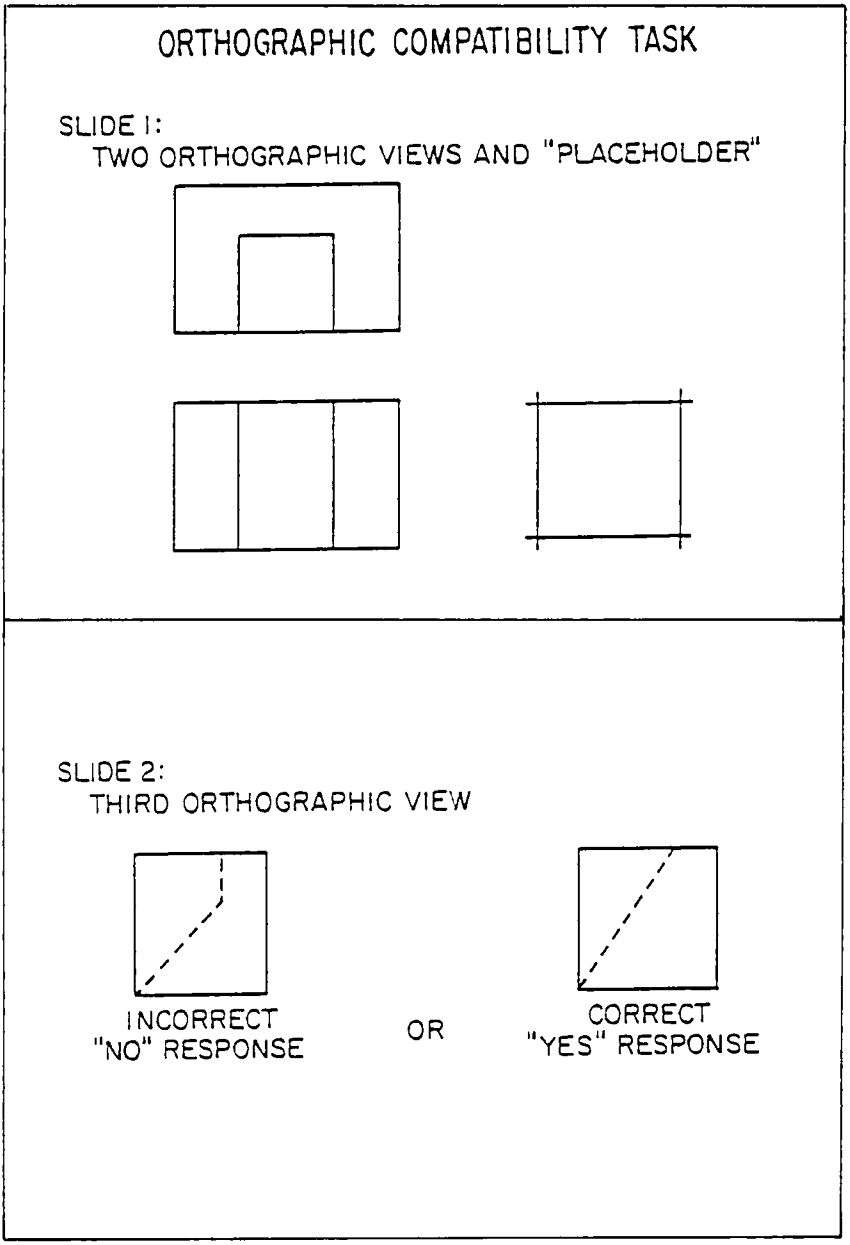

Figure 10. Schematic illustration of the sequence of events on a trial in the orthographic-compatibility problem solving task.

spatial reasoning task in which they engage is illustrated in Figure 10. On each problem, an initial slide containing two orthographic views of an object is presented, along with an empty "placeholder", indicating the location of a third view. A second slide containing a potential third or completion view is then presented, and subjects must judge whether or not the three views will combine to form a possible three-dimensional object. For the orthographic views shown in Figure 10, the completion view on the left will not combine with the two given views to form an object, but the completion view on the right will so combine. (The isometric view corresponding to the three-dimensional object formed by the correct set of orthographic views is shown in the right-hand lower portion of Figure 11, for purposes of confirmation.) *It is important to note that isometric views were never shown to subjects during the problem solving phase of these experiments.*

Immediately following the problem solving phase of the experiment, the observers are requested to engage in a surprise recognition task, and this task is the important one for purposes of the present discussion. In the recognition procedure, subjects are shown pairs of isometric views of objects and are required to indicate which of the two isometrics corresponds to any of the same objects shown during problem solving as sets of orthographic projections. That is, the surprise recognition test asks that subjects discriminate between previously-unseen isometric views of objects, shown before as flat, disconnected orthographic views, and structurally-related distractor objects. An illustration of an incidental recognition trial is given in Figure 11.

The basic finding in this experiment (which has now been replicated several times) is that rate of isometric recognition is very high overall (almost 90%) even though the task is to discriminate between structures that the subjects have never before seen in the surface form provided at the time of recognition (see Cooper, 1988; 1989; 1990, for details). This result suggests that people solve the original "orthographic compatibility" problems by constructing an internal representation of a three-dimensional object, with that mental model being used for subsequent recognition, even though the original problems had been presented as and could have been solved on the basis of separated, flat views of individual sides of objects.

Even more compelling evidence for this interpretation comes from the finding that the probability of correctly recognizing an isometric view of an object (given that the corresponding orthographic compatibility problem has been solved correctly) is significantly greater than the probability of correct recognition, given an incorrect problem solution (.90 and .72, respectively, in an initial experiment, cf., Cooper,

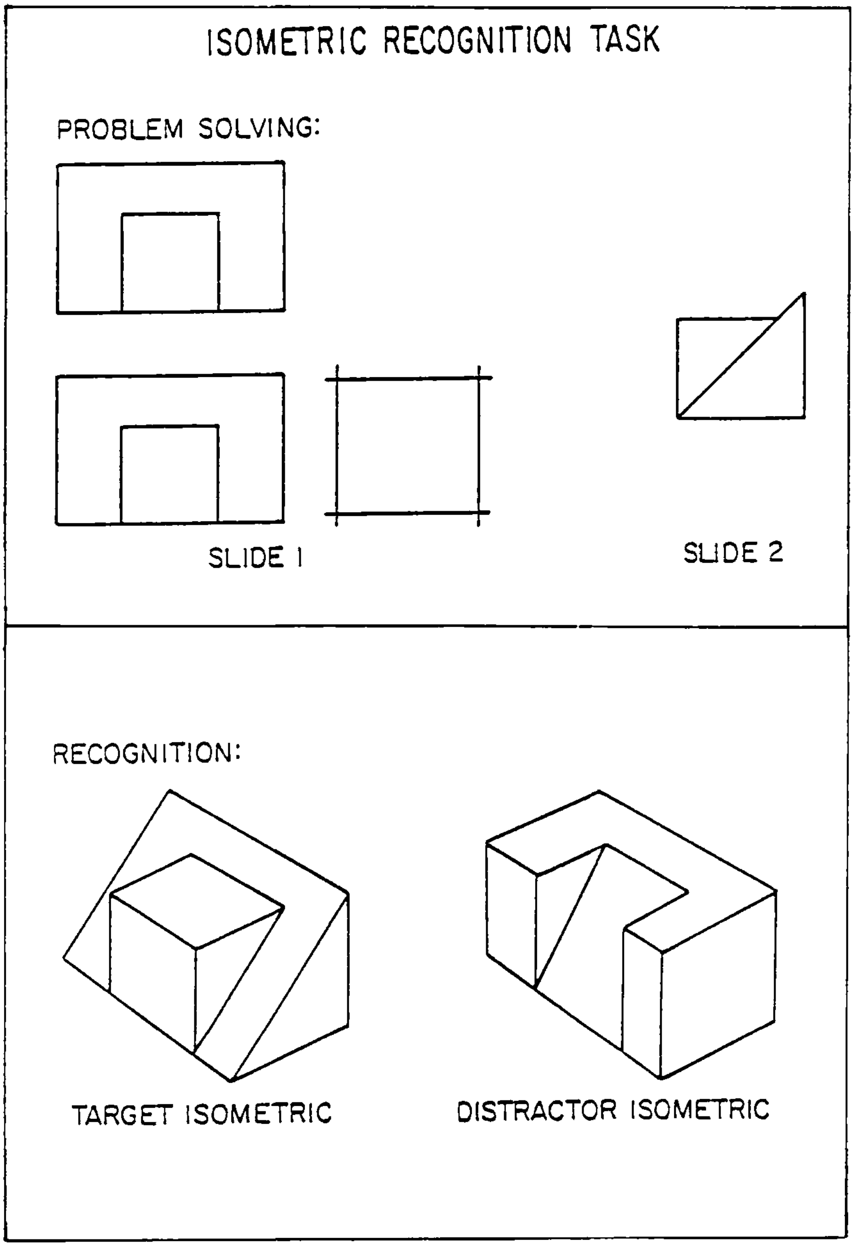

Figure 11. Schematic illustration of the structure of the isometric recognition task.

1990). This last finding strongly supports the claim that the process of problem solution is mediated by a mental model of a three-dimensional object, rather than by the presented orthographic views. This is because retention of information about individual orthographics need not be affected by whether a given problem was solved correctly or incorrectly. However, retention of information about corresponding isometric representations of objects should be strongly affected by problem solving accuracy because, by hypothesis, it is the mental construction of a representation corresponding to a three-dimensional object that underlies the problem solution process.

We have used this "isometric recognition effect" and the same sort of experimental situation to explore how accessible non-depicted or "hidden" surfaces of objects might be. This is accomplished by seeing whether relationships among object surfaces not explicitly presented in orthographic form at the time of problem solving can be induced from the mental representation mediating isometric recognition. In one experiment directed to this issue (see, also, Cooper, Mowafy, & Stevens, 1986), subjects were required to solve orthographic compatibility problems and then to engage in incidental recognition of isometric views. However, the isometric recognition pairs were of eleven different types, divided into four separate conditions. The relationship among the target isometric structures for three of these conditions is shown in Figure 12. The various isometric views all correspond to the same three-dimensional object that would be constructed by combining the orthographic projections shown in the center of the figure. They differ, however, in *which* surfaces of the objects are visible and in *how* the visible surfaces correspond to the three given views displayed in orthographic form. Thus, the isometrics can share all three (the standard isometric view), two, one, or no surfaces in common with the three orthographics explicitly presented in the problem solving task.

In each of the experimental conditions, forced-choice surprise recognition pairs contained either standard isometric views, along with the associated distractors, or isometric/distractor pairs that shared various numbers of views in common (two, one, or zero) with the surfaces presented in orthographic form. In addition to the three conditions comparing performance on standard and viewpoint-transformed isometric recognition items, a fourth condition compared performance on standard isometrics with the same isometric view, but rotated 60, 120, or 180 degrees in the picture plane. The reason for including this condition was to enable comparisons between recognition of isometrics that have undergone structure-revealing transformations around one or more axes in depth and those that have undergone structure-preserving

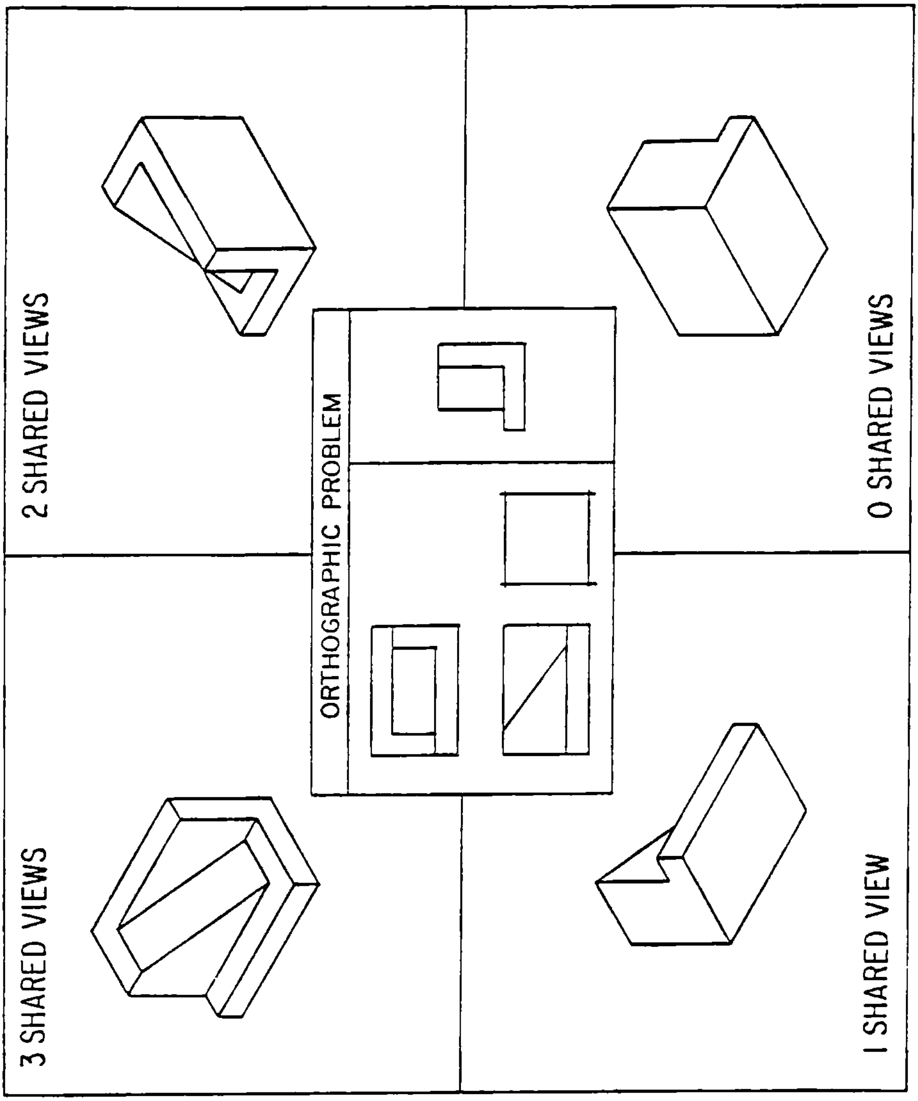

Figure 12. Relationship between the object surfaces displayed in orthographic projection during the problem solving task, and the object surfaces shown in isometric projection on the incidental recognition task. The four types of views define four of the experimental conditions. See text for further explanation.

transformations that do not result in the presentation of previously hidden surfaces. Separate groups of twelve observers participated in each of these four conditions.

Predictions for recognition performance in these experimental conditions are easiest to develop for extreme positions concerning the extent to which constructed mental representations of objects preserve view specificity. On the one hand, if such mental models of objects represent "three-dimensional" structure in a "two-or-two-and-a-half-dimensional" fashion, i.e., from only one point of view, then the various transformed recognition pairs should not differ from each other, and all should exhibit substantially poorer recognition than pairs of isometrics in the standard orientation. On the other hand, if all structural information in a constructed mental representation of an object is equally available, i.e., the representation is "orientation-free" or "object-centered" in Marr's (1982) terminology, then the overall level and pattern of recognition performance should not differ among the various transformed views and the standard orientation of the test isometrics.

The results of this experiment, considering both accuracy and latency of forced-choice recognition, fall somewhere between predicted levels for these two extreme possibilities. Overall recognition performance clearly differed among the five types of target objects, but not in the simple fashion expected under either theoretical account. Specifically, *recognition accuracies* for standard (three), rotated, and two shared views were high and at roughly equal levels. However, recognition of isometrics with no views shared with the standard approached chance, with the one-shared-view condition falling between these performance extremes. The *recognition latencies* showed a somewhat different pattern, with the most substantial decrement in performance occurring between the standard or three-shared-views condition and all of the others.

The conclusion that can be drawn from these patterns in the recognition results is that some degree of non-explicit or "hidden" structural information can be accessed from constructed mental representations of three-dimensional objects. The extent of such accessibility is limited, and the latency data indicate that even when possible (as in the case of picture-plane rotated isometrics) this access required additional time. While the accuracy with which isometrics sharing two views in common with standard isometrics compared well with accuracy for recognizing the standard views themselves, the time needed to make the recognition was significantly longer than for recognizing standard isometric views. In the extreme case of no shared surfaces between standard and test isometrics, latency was long and accuracy was close to chance, indicating that access to object structure which is specified as completely concealed is virtually impossible. Thus, the mental representation of an object formed from separate, flat orthographic projections may contain accessible information about the structure of immediately adjacent surfaces; however, essentially no structural information is available about completely hidden views. These representations, then, are neither strictly viewer-centered nor fully object-centered in a representational scheme like that proposed by Marr (1982).

CONCLUDING REMARKS

The picture emerging from the experimental work described in this paper is of aspects of, or systems for, the representation of objects that appear to act in concert during our ordinary perceptual encounters with or imaginal anticipations of events in the world; however, these representational systems can be dissociated and their separate properties examined more closely by appropriate experimental manipulations with normal adult subjects, and perhaps also in individuals with specific forms of brain injury. Schacter and I (Schacter, Cooper, & Delaney, 1990a; 1990b; Schacter, Cooper, Delaney, Peterson, & Tharan, 1991), along with others (e.g., Riddoch & Humphreys, 1987b), are inclined to refer to the system accessed by implicit memory tasks as the *structural description system.* This system appears to have much in common with other pre-semantic, perceptual representation systems underlying implicit memory phenomena in the verbal domain (see Schacter, 1990, for a discussion).

This structural representation system encodes information concerning the global relations that define the structure of an object, and information accessed from this system supports performance on automatic, non-conscious tests of memory, as revealed by priming effects in the object decision task. It appears relatively insensitive to semantic information afforded by elaborative encodings of objects, and it does not appear to incorporate aspects of visual information (such as size, standard/reflected orientation, and local properties of component line orientations) that are irrelevant to the representation of relatively abstract, structural relations.

The information provided by the structural description system certainly plays some role, however limited, in the explicit recognition of objects. Currently, we can conceptualize the representational system underlying explicit remembering of visual objects and events as being supported by multiple and diverse sources of information. These sources most likely include (in addition to structural information) information about object identity and meaning, contextual knowledge about the circumstances under which objects are encountered, knowledge about the functions that particular objects serve, and other forms of information that elaborate and make distinctive instances of individual objects. And, some aspects of the information that individuate particular objects and events are specifically visual in nature, including information about object size, parity, and feature composition. It is just these sources of information that a considerable amount of research in cognitive psychology (see, in particular, the work of Kosslyn and his associates, e.g., Kosslyn, 1980; Kosslyn, Ball, & Reiser, 1978) has shown are accessed when subjects generate mental images to perform specified tasks. Perhaps, then, the label "mental image" should be reserved for those aspects of representations that underlie explicit tasks like recognizing a previously-seen object or recalling properties of a visual display no longer present.

The dissociable aspects of, or systems for, the representation of visual objects described in the present paper bear at least superficial similarities to distinctions made by Farah, Hammond, Levine, and Calvanio (1988) between "spatial" and "visual" mental imagery systems and by Kosslyn (1987) between processing subsystems common to perception and imagery that represent metric and categorical information. Whether these distinctions will ultimately map onto the present distinction between representational systems for structural information and for the recognition of objects is not yet clear. What *is* clear, however, is that demonstrations of dissociable aspects of mental representations of objects provide a rich source of evidence in the continuing effort to understand the nature of visual cognition.

REFERENCES

Biederman, I., & Cooper, E. E. (1989). *Evidence for complete translational and reflectional invariance in visual object priming*. Unpublished manuscript.

Cooper, L. A. (1988). The role of spatial representations in complex problem solving. *In* S. Steele & S. Schiffer (Eds.), *Cognition and representation*. Boulder, CO: Westview.

Cooper, L. A. (1989). Mental models of the structure of three-dimensional objects. *In* S. Ballesteros & B. Shepp (Eds.), *Object perception: Structure and process*. Hillsdale, NJ: Erlbaum.

Cooper, L. A. (1990). Mental representation of three-dimensional objects in visual problem solving and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **16**, 1097-1106.

Cooper, L. A., Mowafy, L., & Stevens, J. J. (1986, November). *Perceiving hidden structure: View-independent recognition of constructed three-dimensional representations.* Paper presented at the 27th annual meeting of the Psychonomic Society, New Orleans, LA.

Corballis, M. C. (1988). Recognition of disoriented shapes. *Psychological Review* **95**, 115-123.

Farah, M. J., Hammond, K. M., Levine, D. N., & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology* **20**, 439-462.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic patients. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **11**, 501-518.

Graf, P., Squire, L. R., & Mandler, G. (1984). The information that amnesic patients do not forget. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **10**, 164-178.

Jolicoeur, P. (1987). A size-congruency effect in memory for visual shape. *Memory & Cognition* **15**, 531-543.

Jolicoeur, P., & Besner, D. (1987). Additivity and interaction between size ratio and response category in the comparison of size-discrepant shapes. *Journal of Experimental Psychology: Human Perception and Performance* **13**, 478-487.

Kosslyn, S. M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres. *Psychological Review* **94**, 148-175.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance* **4**, 47-60.

Marr, D. (1982). *Vision.* San Francisco, CA: Freeman.

Penrose, L. S., & Penrose, R. (1958). Impossible objects: A special type of visual illusion. *British Journal of Psychology* **49**, 31-33.

Richardson-Klavehn, A., & Bjork, R. A. (1988). Measures of memory. *Annual Review of Psychology* **36**, 475-543.

Riddoch, M. J., & Humphreys, G. W. (1987a). Visual object processing in optic aphasia: A case of semantic access agnosia. *Cognitive Neuropsychology* **4**, 131-186.

Riddoch, M. J., & Humphreys, G. W. (1987b). Picture naming. *In* G. W. Humphreys & M. J. Riddoch (Eds.), *Visual object processing: A* *cognitive neuropsychological approach.* London: Lawrence Erlbaum Associates.

Rock, I., & Di Vita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology* **19**, 280-293.

Schacter, D. L. (1985). Priming of old and new knowledge in amnesic patients and normal subjects. *Annals of the New York Academy of Sciences* **444**, 41-53.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **13**, 501-518.

Schacter, D. L. (1990). Perceptual representation systems and implicit memory: Toward a resolution of the multiple memory systems debate. *In* A. Diamond (Ed.), *Development and neural bases of higher cognition. Annals of the New York Academy of Sciences,* **608**, 543-571.

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990a). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General* **119**, 5-24.

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990b). Implicit memory for visual objects and the structural description system. *Bulletin of the Psychonomic Society* **28**, 367-372.

Schacter, D. L., Cooper, L. A., Delaney, S. M., Peterson, M. A., & Tharan, M. (1991). Implicit memory for possible and impossible objects: Constraints on the construction of structural descriptions. *Journal of Experimental Psychology: Learning, Memory, & Cognition* **17**, 3-19.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, **21**, 233-282.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science* **247**, 301-396.

Warrington, E. K., & Weiskrantz, L. (1974). The effect of prior learning on subsequent retention in amnesic patients. *Neuropsychologia* **12**, 419-428.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 2

Visual imagery in locomotor movement without vision*

Even Loarer
Centre de recherche de l'INETOP, CNAM, Paris, France
and
Alain Savoyant
Université d'Aix-Marseille II, Marseille, France

Since work by Gibson (1966, 1979), the importance of vision, and in particular transformations in the visual scene during movement has been widely recognized for movement control and for perception of the structure of space. Gibson stressed the direct nature of visual perception, and argued that information processing and intermediate representations play little or no role. But what happens when vision is taken away? Reliance upon non-visual information and the implementation of cognitive processes appear to be necessary in this case (Strelow, 1985), but relatively small and immediately perceived spaces may involve the action of processes where there is a functional equivalence between imagery and perception (Finke, 1980; Finke & Shepard, 1986). In this perspective, a visual-type representation may be substituted for visual perception and then be used for both movement control and egocentric perception of the structure of space (i.e. localization of the position of objects with respect to oneself).

Localization of objects in space in non-visual movement

When an individual moves, the relative positions of objects in the subjects environment change. People perceive these changes as apparent movement of these objects in their visual field, forming what is called an optical flow (Gibson, 1950). The localization of objects in space during

*This work was conducted while E. Loarer was member of Cognition et Mouvement movement when vision has been suppressed thus calls for continuous updating, on the part of the subject, of the spatial relations between self and the objects around him/her. Only a few studies have been denoted to this type of task, and have assumed that mental imagery serves as a substitute for the missing visual information. The findings vary as a function of task difficulty. Broadly speaking, two levels of difficulty have been investigated: The most elementary is where the subject is asked to locate a single object and move towards it. In this type of task, commonly termed a "locomotor pointing task", the object is located along the axis of movement of the subject. So, there is no problem of orientation (no change in orientation during movement). The only problem is *distance* to the object, which varies continually. The most complex is where the object is not located along the axis of movement and thus *distance* from the subject and *orientation* both vary during movement.

As we will see, it is likely that mental imagery is used to carry out the first type of task and deal with the problem of distance of self from an object. However, the lack of appropriate methodology has made it much more difficult to obtain clear-cut evidence for the implementation of mental imagery in complex tasks of this kind.

Mental imagery in elementary tasks

In a series of experiments on non-visually guided locomotion, Thomson (1980, 1983) asked subjects to close their eyes and walk up to a target placed on the floor at varying distances straight in front of them, after having looked at the target for 5 seconds while standing at the starting point. The results suggest that subjects form a transitory representation (lasting 8 seconds) of the environment. More specifically, Thomson (1980, 1983, 1986) posits (based on the verbalizations of his subjects, who stated they could "see themselves approaching the target") that such representations are in a visual form, although he does not take a position on whether they are exo- or egocentric ("map-like" or "image-type" representations").

These findings were challenged by Elliott (1986, 1987) and Steenhuis and Goodale (1988) who found that the accuracy of locomotor pointing decreased as the distance from the target increased, and did not obtain any substantial or sharp drop at the 8-second threshold. The observed differences in degree of accuracy appear to be mainly due to the difference in constraints created by the locomotor pointing (Laurent & Thomson, 1988). The 8-second threshold beyond which pointing accuracy decreases sharply and abruptly (Thomson, 1983) in fact appears to originate from a statistical bias (pooling of between- and within-subject variables) mentioned by Elliott (1986) and confirmed by Steenhuis and Goodale (1988). Note that these results do not question the existence of a representation of the environment (the position of the target), but rather its accuracy and high transience.

Another factor, the context of target perception, also appears to affect the accuracy of pointing. Corlett (1986) showed that the presence of landmarks placed before and after the target improved pointing accuracy. This same factor may account for the high precision obtained in situations where subjects had to go around several obstacles before reaching the target (Thomson, 1980; Laurent & Cavallo, 1985). More plentiful spatial information at the onset is thus thought to facilitate locomotor pointing.

Uptake of this initial information before blindfolded movement may be the basis for the programming of locomotion towards a target. The remaining important issue is whether this information is also coded in a representational form which is used during locomotion itself. Simply analyzing the accuracy of locomotor pointing is not sufficient to answer this question. Rather, a more suitable approach would be fine analysis of the behavioral cues of vision-deprived locomotion. Laurent and Thomson (1988) and Loarer (1988) showed that during blindfolded pointing, adjustments in step length are made as the target is approached, as in cases where vision is available (although the pattern is slightly different). Loarer (1988, 1989) also showed that as a target is approached, the head is tilted along the sagittal plane in a similar manner (although to a lesser degree) in the non-vision condition and the vision condition. These results are consistent with the idea that a representation of target location is used during blindfolded locomotion, and serves as a substitute for direct visual perception. In this respect, the representation should be in the form of an egocentric image of the target, and in order to be functional in locomotor pointing, should above all reproduce not only the characteristics of the visually perceived space but also the locomotion-related transformations of that space. It does appear possible to evoke such an image. For example, in a series of experiments by Kosslyn (1978) in which subjects were asked to construct a visual image of a given object (named or perceived beforehand) seen from afar, and then to imagine themselves walking towards it, subjects reported a gradual looming effect of the imagined object. Of course, distortions of such images are natural, and as Steenhuis and Goodale (1988) stated, "the accuracy of any short-term memory or mental image of a target location should be limited by errors and biases inherent to the perceptual system."

Mental imagery in complex tasks

In complex tasks, subjects not only need to form a mental image of the target and its loom as they move towards it, but must also form an image of object trajectories in the optical flow. The basic issue is whether subjects form and use a mental image of the optical flow while they are moving, in vision-deprived situations.

Böök and Gärling conducted a series of experiments concerning "how maintenance of orientation during locomotion in unfamiliar environments is achieved" in which subjects, after having been guided in the dark along a locomotion path, were required to point to a target (Böök & Gärling, 1980, 1981).

In these studies, subjects had no advance information on location of the new point they have to reach, which forced them to process information about their own movement as they moved. This favours continuous transformations of the representation of the initially perceived space. However, for these authors, during the change in point of observation in this type of task, information on the distance and changes in direction of locomotion are centrally processed in a recurrent manner. This information is thought to be converted by computation processes into information on target distance and direction, and then to be decoded later, i.e. "translated into a percept of the location of the target relative to the subject, thereby preparing him/her for making an estimate of respectively the direction and the distance to the target" (Böök & Gärling, 1980).

This perspective thus indeed involves an egocentric, visual image of the target, but only as the resultant of periodic computational processing aimed at compensating for the completed movement, and not as the continuous transformation of the initial visual image. The main factor which seems to make such processing necessary is that subjects only had to orient themselves to a single target, which is relatively easy.

Changes in positions of objects with respect to the subject during locomotion vary as a function of distance from the subject and their deviations from the locomotor axis. These changes in position are governed by a dynamic geometry whose complexity, as soon as there are several objects, is much too great for subjects' computational abilities. This leads to the question of what would happen if the number of objects which had to be located after movement was increased.

Rieser, Guth, and Hill (1986) conducted such an experiment where sighted adults were allowed to view the spatial layout of several objects in a room. They then walked without vision to a new point of observation and judged the structure of their perspective from that new point. Rieser et al. posit that subjects use a "perceptual learning view". "The thesis is that changes in perspective are often visible and that sighted observers learn the predictable covariation of visual flow fields and non-visual cues of locomotion". In other words, "in other situations where the flow fields are either partly or totally occluded from view, the learning generalizes to mediate accurate judgements of changes in perspective" (Rieser et al., 1986). Although the authors did not explicitly say so, such an interpretation should involve a representation of these flow fields. However, unfortunately, unlike the experiments by Böök and Gärling (1980, 1981), the paradigm did not call for continual processing during locomotion. Indeed, subjects were informed in advance of the location and the orientation of the new point of observation, and were asked to examine them from the starting point (the authors did not state for how long). Such conditions allowed for the anticipated cognitive computation of the new perspective based on visual perception, before the change in point of observation. A computation of this sort does not involve taking movement itself into account, and in particular, the continuous processing of locomotor information appears to be uninformative.

Thus although certain results obtained in studies on locomotor pointing tasks without vision (elementary tasks) suggest that a visual image may serve as a substitute for visual perception, this was not found to be true in studies on egocentric perception of the structure of space during a change in point of observation without vision, in what we have termed complex tasks. This appears mainly to be due to the fact that the experimental conditions in these studies made it highly unlikely if not useless to rely on an image (and its continuous transformations) of the initially perceived visual scene.

The aim of the experiment is to test subjects' abilities to process simultaneously changes in the position of several objects in a visually deprived situation (as in Rieser et al. 1986), with no advance information on their final locomotor destination (as in Böök and Gärling, 1980, 1981).

METHOD

Subjects

Six adult subjects (3 men and 3 women between the ages of 22 and 46), all students or members of our laboratory, participated in the experiment.

Experimental space

The experimental space consisted of four target objects located in front of the starting point on either side of the sagittal axis (two objects in each hemi-field thus defined) in a 12 m x 8 m room with no outside view. The final position of the subject (which is termed "the new viewpoint") was located on this same axis at varying distances (4, 5, or 6 metres), always in front of the target objects and always oriented in the same direction (Figure 1). The objects were cylinders 2 metres high and 30 cm in diameter, each of a different colour (red, blue, yellow, and green).

Movement from the starting point to one of the three new viewpoints caused an apparent shift in the objects located in the visual field, which can be expressed in terms of their angular movement (the objects moved away from the locomotor axis): the angular movement of a given object was greater when the object was farther from this axis and closer to the starting point. In addition, for the two objects in the same hemi-field, movement to the new viewpoint caused the angular disparity between the two objects to either increase, decrease or be reduced to zero, or change signs. These three situations are termed opening, closing or alignment, and permutation, respectively (in the latter case, the object seen on the left of the other object at the beginning is seen on its right upon arrival).

Procedure and design

After having practiced associating a given number of steps to each of the three distances, subjects were briefly familiarized with the types of movements the objects would appear to undergo during locomotion.

Two experimental conditions were defined according to whether subjects had advance information or no advance information about the location of the new point they have to reach.

In **condition 1**, subjects had no advance information on where the new viewpoint was located. After having observed the scene from the starting point for 15 seconds, they were deprived of vision (dark room and opaque glasses) and told how many steps they had to take. They walked and were then asked to point to the target objects. In **condition 2**, the new point of observation was formed by a circular target placed on the floor. The new point they had to reach was therefore visible during the 15-second observation phase when subjects were asked to construct a representation of the location of the objects relative to that point of observation. They were then blindfolded and asked to point to the target objects while considering themselves to be located at the new point.

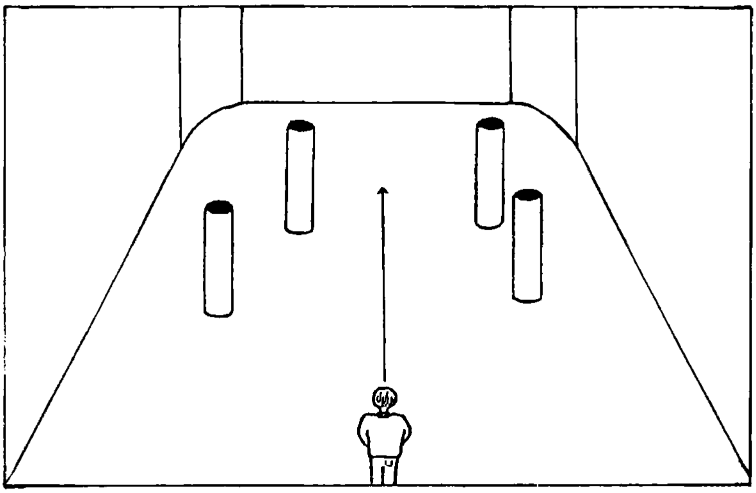

Figure 1: Diagram of the experimental space.

In addition, in a third condition, subjects were actually located at the new viewpoint, and had to point to the objects from it, again, once they had been blindfolded following a 15-second observation period. This condition was basically a **control condition** insofar as it provided us with information about pointing errors without any transformations in the scene brought about by the change in point of observation.

Directional pointing was performed with a rotatable rod fastened to a horizontal dial mounted on a table with castors (Figure 2). The subject placed his/her left hand on the fastening point of the stick presented at the breastbone level, and his/her right hand on the other end of the stick. The right arm extended as such, he/she pointed one by one to the four target objects, in any order, while the experimenter recorded each pointing on a sheet of paper covering the dial.

The pointing device was always presented to the subjects with the stick pointed straight ahead of the subject (in the sagittal plane of the subject's movement), with the same time lapse following blindfolding in each of the conditions. Subjects were pre-trained with sight in how to use this device, and were told to point to the centre of the cylinders. They attained perfect accuracy (±1 degree) during training.

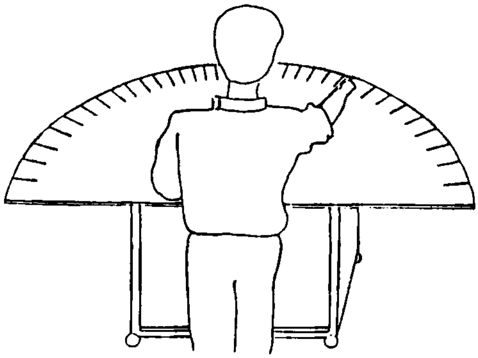

Figure 2: Diagram of the pointing device.

Nine different layouts of the four objects were used (three for each locomotor distance), including six openings, six closings or alignments, and six permutations in the 18 hemi-fields thus defined. All subjects performed first in condition 1, after which they always performed in condition 2 and the control condition, and in that order. In each condition, the order of the nine layouts was randomized.

RESULTS AND DISCUSSION

For each condition, the mean values of intra-individual constant errors (CE), variable errors (VE), and absolute errors (AE) were calculated (see Table 1).

The low mean value of the constant error (CE) in conjunction with the high mean value of the variable error (VE) indicates that pointing error does not correspond to systematic undershooting or overshooting. As a result, mean constant error is not an adequate measure of accuracy because its calculation involves the cancelling of positive and negative errors whose distribution best describes the subjects' performance. For this reason, only absolute (AE) and variable (VE) error will be considered (Figure 3).

The analysis of variance yielded a significant effect of conditions ($F(2,10)=29.69$; $p<0.001$), and a significant interaction between conditions and objects ($F(6,30)=6.7$; $p<0.001$). The layout factor had no significant effect ($F(8,40)=1.225$). Anovas for across-condition comparisons yielded significant differences between the different conditions: C1/C2:

($F(1,5)=15.93$, $p<0.025$); C2/C3: ($F(1,5)=22.94$), $p<0.005$; C1/C3: ($F(1,5)=38.93$), $p<0.005$).

Table 1. Mean values of intra-individual error (in degrees) by condition for absolute error (AE), constant error (CE), and variable error (VE).

	Condition 1	Condition 2	Condition 3 (Control)
Absolute Error	13.5	10.6	5.9
Constant Error	1.8	0.1	-2.8
Variable Error	10.7	8.9	5.1

Performance in the control condition (pointing from the arrival point) was better than in all of the other conditions. The change in point of observation thus does appear to have been a source of error. However, precision was still inferior to that obtained during the pre-training. This probably reflects the temporal delay introduced between blindfolding the subjects and pointing. Performance in condition 2 (knowledge of the arrival point) was better than in condition 1. Prior awareness of the new point of observation thus led to more accurate pointing than in conditions where subjects were unaware of the arrival point before being blindfolded. However, the degree of accuracy obtained in Condition 1 where possibilities for reliable computation were suppressed and where the subjects could, in all likelihood, only use a mental image of the optical flow consecutive to locomotion was remarkably good.

In this experiment, our aim was to test subjects' ability to form and use a mental image of the optical flow created by their own locomotion. Two conditions were devised for this purpose: one in which the location of the arrival point was known ahead of time, the other in which it was not. In the former case (C2), the cognitive processing of the new location of objects was possible. In the latter case (C1), uncertainty of the distance to be covered, and thus of the new point of observation, made calculation of this type extremely difficult or impossible, and subjects were most likely led to form a visual image of the starting scene and to continuously refresh this image during locomotion.

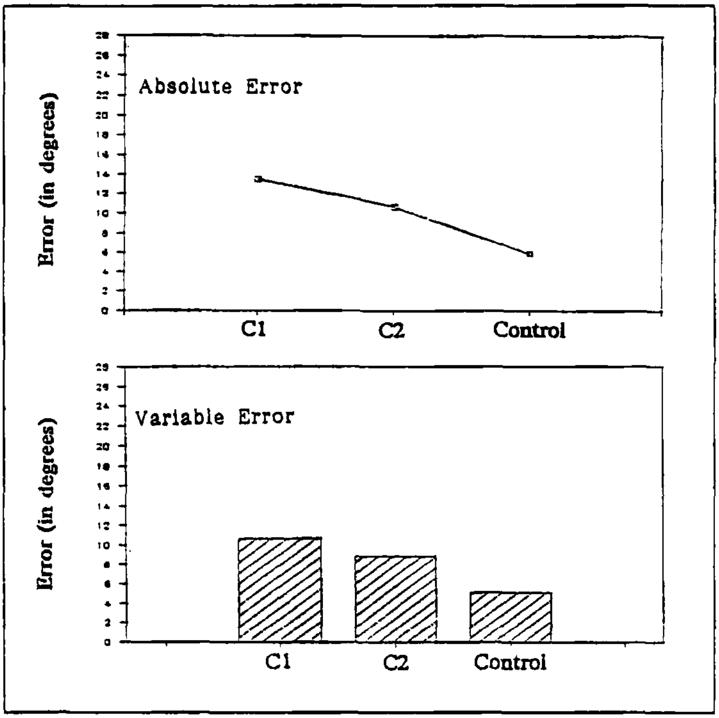

Figure 3: Mean values of intra-individual error (in degrees) by condition for absolute error (AE), and variable error (VE).

The experimental set-up that we chose thus allowed us to demonstrate, first of all, subjects' ability to form a visual image of space which they can use during displacement to locate the targets. In addition, we can compare the accuracy obtained in this condition with that obtained in conditions where the subject knows in advance where the arrival point is located and can therefore do without the mental imagery.

In the condition used in this experiment, resorting to an image of the visual scene and its transformations when the location of the arrival point was not known in advance appears not only to have been the most probable solution, but also the most economical one. Indeed, in addition to the fact that in the familiarization phase the subjects' attention was first directed towards the apparent movement of the objects caused by the displacement, and that in the non-vision condition of the experiment subjects were explicitly invited to form an image of these transformations, reliance upon a computational type of processing such as the one proposed by Böök and Gärling (1980, 1981) appears to have been difficult in these situations where four objects in a layout had to be processed simultaneously.

Such reliance on a visual image is subject to error, so performance was naturally better in the second condition in which calculations could be made on the basis of visual perception. In this respect, there does not appear to be any reason why a subject would prefer to use a visual image when calculation of this sort is possible; Rieser et al.'s (1986) finding thus appears paradoxical: their experimental conditions were particularly favorable to the cognitive computation of the new location of the objects, yet they propose a perceptual learning view based on "the predictable covariation of visual flow fields and nonvisual cues of locomotion."

Finally, it should be stressed that this solution was also fairly reliable since it enabled subjects to identify the new orientations of objects with a good degree of accuracy after locomotion.

REFERENCES

Böök, A. & Gärling, T. (1980). Processing of information about location during locomotion: effects of a concurrent task and locomotion patterns. *Scandinavian Journal of Psychology* **21**, 185-192.

Böök, A. & Gärling, T. (1981). Maintenance of orientation during locomotion in unfamiliar environments. *Journal of Experimental Psychology: Human Perception and Performance* **7**, 995-1006.

Corlett, J.T. (1986). The effect of environmental cues on locomotor distance estimation by children and adults. *Human Movement Science* **5**, 235-248.

Elliott, D. (1986). Continuous visual information may be important after all: a failure to replicate Thomson (1983). *Journal of Experimental Psychology: Human Perception and Performance* **12**, 388-391.

Elliott, D. (1987). The influence of walking speed and prior practice on locomotor distance estimation. *Journal of Motor Behavior* **19**, 476-485.

Finke, R. (1980). Levels of equivalence in imagery and perception. *Psychological Review* **87**, 113-132.

Finke, R., & Shepard, R.N. (1986). Visual functions of mental imagery. *In* K.R. Boff, L. Kaufman & J.P. Thomas (Eds.), *Handbook of Perception* *and Human Performance, Vol. II: Cognitive Processes and performance.* New-York: John Wiley.

Gibson, J.J. (1950). *The perception of the visual world.* Boston, MA: Houghton Mifflin Company.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* Boston, MA: Houghton Mifflin Company.

Gibson, J.J. (1979). *The ecological approach to visual perception.* Boston, MA: Houghton Mifflin Company.

Kosslyn, S.M. (1978). Measuring the visual angle of the mind's eye. *Cognitive Psychology* **10**, 356-390.

Laurent, M., & Cavallo, V. (1985). Rôle des modalités de prise d'information visuelle dans le pointage par positionnement du corps. *L'Année Psychologique* **85**, 41-48.

Laurent, M., & Thomson, J.A. (1988). The role of visual information in the control of a constrained locomotor task. *Journal of Motor Behavior* **20**, 17-37.

Loarer, E. (1988). Can mental image play a part in the control of locomotion? Paper presented at the *Second European Workshop on Imagery and Cognition*, Padova, Italy.

Loarer, E. (1989). Un exemple de coopération vision-représentation : le contrôle visuel intermittent des déplacements locomoteurs. *Psychologie Française* **34** (1), 55-62.

Rieser, J.R., Guth, D.A., & Hill, E.W. (1986) Sensitivity to perspective structure while walking without vision. *Perception* **15**, 173-188.

Steenhuis, R.E., & Goodale, M.E. (1988). The effects of time and distance on accuracy of target-directed locomotion: Does an accurate short-term memory for spatial location exist? *Journal of Motor Behavior* **20**, 399-415.

Strelow, E.R. (1985). What is needed for a theory of mobility: direct perception and cognitive maps. Lessons from the blind. *Psychological Review* **92**, 226-248.

Thomson, J.A. (1980). How do we use visual information to control locomotion? *Trends in Neurosciences*, 247-250.

Thomson, J.A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance* **9** (3), 427-443.

Thomson, J.A. (1986). Cognitive and motor representations of space and their use in human visually guided locomotion. *In* P. Ellen & C. Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man.* Dordrecht, The Netherlands: Martinus Nijhoff Publishers BV.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)
Elsevier Science Publishers B.V., 1991

Chapter 3

Conflicting spatial frames of reference in a locating task

Patrick Péruch and Alain Savoyant
Université d'Aix-Marseille II, France

INTRODUCTION

Defined by Levine (1982) and Levine et al. (1982, 1984) as the non-correspondence between the orientation of a map and of the field it represents, contralignment usually creates some difficulty for the subjects who must orientate themselves in space. This difficulty is interpreted as showing that subjects use a cognitive map with a specific orientation relative to themselves. A similar effect was shown by Shepard and Hurwitz (1984) in a situation in which the problem was not to store spatial information, but to identify the direction of a turn following a rectilinear displacement. Identification became more difficult as the displacement, represented by a segment on a screen, was displayed in a more and more non-canonical orientation (the canonical orientation facing upwards on the screen). In both of these cases difficulties arise from the fact that the same spatial information about the layout has to be used in two different spatial frames of reference (SFR) defined by two reference points:

(1) The subject him/herself watching the map or the screen with his/her own up, down, right and left (front and back are not pertinent here). This defines an egocentric SFR in which the displayed objects can be located.

(2) The subject's represented location on the map (or on the screen) which, insofar as it possesses an orientation, defines a front, back, right and left (up and down are not pertinent here), and thus an object-centered SFR in which the spatial location of the displayed objects can be assessed.

It is precisely the discrepancy between the orientation of these two reference points which defines contralignment. Here the mental processes needed to compensate the contralignment bear mainly on map-like visuo-spatial representations. They involve cognitive operations such as mental rotation. However, in daily orientation behaviour, we use primarily the egocentric information corresponding to the scenes that we see from a specific viewpoint (actual or imagined on the map). Such information is usually neglected in most experimental studies, whereas this experiment aimed to study the contralignment phenomenon in a locating task in which subjects had to find the position and orientation of a viewpoint on a map from an egocentric scene that was shown to them.

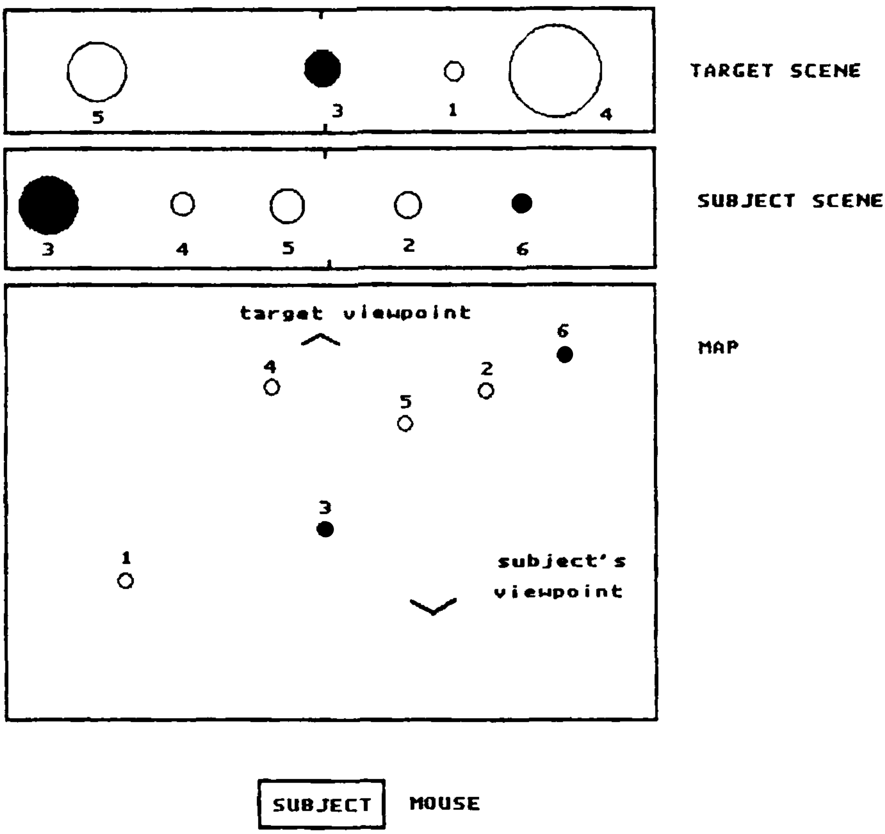

Figure 1. Experimental device (see text for a description).

METHOD

Experimental device

An IBM AT computer linked to a high resolution graphic screen was used to produce maps, target scenes, and subject scenes (see Figure 1).

The subject was seated in front of a screen displaying a map, a target scene and a subject scene. The map was a 20 cm by 36 cm rectangle situated in the lower part of the screen and composed of 6 objects (filled circles 5 mm in diameter shown in different colors). These objects were randomly arranged on the map, thus avoiding obvious patterns such as rows and clusters. The subject scene was presented at the top of the screen in a rectangular 36 cm by 4 cm format and represented what the subject could see from the subject's viewpoint on the map. This 120-degree viewing angle was represented by a graphic cursor, the displacements of which were controlled with a mouse device. Above the subject scene was the target scene, which corresponded to the target viewpoint on the map; the target scene was always composed of four out of six objects on the map. The "egocentric" information about the distance (size) and direction of the environmental objects was given in the subject and target scenes.

Task and conditions

The task consisted in finding the location and orientation of the target viewpoint on the map by moving the cursor using a mouse. The subjects were required to find the solution under six experimental conditions resulting from crossing two-value (the orientation of the target viewpoint) and three-value (the type of available information) factors:

(1) **The orientation of the target viewpoint** with respect to the subject's position in front of the screen (alignment or contralignment). The viewpoint was defined as aligned (condition A1) when it was facing upwards on the screen and contraligned (condition A2) when it was facing downwards (Shepard & Hurwitz, 1984). An example of contralignment is given in Figure 1: the subject's viewpoint is aligned with the subject's position in front of the screen, while the target viewpoint is contraligned.

(2) **The type of available information.** In condition C1, the subject had information about the layout of some objects on the map, and egocentric information about his/her own successively produced scenes. In condition C2, the layout of the objects was not presented, and thus

the subject could only compare the target and subject scenes, both egocentric. In condition C3 the subject was shown the layout on the map but not his/her produced scenes, which had to be mentally represented or computed in order to be compared with the target scene. In each condition, the target scene remained available at the top of the screen during the entire trial, and the subject had on-line information about the current location and orientation of the cursor on the map.

Following Levine and colleagues (1982, 1984) we hypothesized that contralignment would increase the difficulty of the task, in particular because it requires more mental manipulation of visuo-spatial representations than alignment. Moreover, by taking into account the type of available information during the searching procedure, we also aimed to differentiate some of the necessary conditions of the contralignment effect. Indeed, conditions C2 and C3 may correspond to two different problem-solving modes, carried out on the basis of different types of information. The C1 condition, in which these two modes are simultaneously possible, allows their respective use during the searching procedure to be evaluated.

Data processing

A trial lasted until the subject considered his/her response acceptable. The quality of this response was evaluated by comparing the target scene with the subject scene of the final viewpoint reached by the subject. To do so three criteria were used following the procedure advanced by Péruch et al. (1986):

(1) A cardinal criterion which assessed the number of correct (i.e. present in the target scene) and incorrect (i.e. absent from the target scene) objects present in the subject scene.

(2) An ordinal criterion which assessed the order of appearance (from left to right or right to left) of the correct objects present in the subject scene.

(3) An angular criterion which assessed the angular distance of the correct objects in the subject scene.

Cardinal, then ordinal, and then angular criteria were usually satisfied within a given trial in that order (Péruch et al., 1986). On this basis, two main phases in the course of a trial were distinguished:

(1) A first phase at the end of which the subject scene was composed of the 4 correct objects (and possibly 1 or 2 incorrect objects), with only one order inversion between 2 adjacent correct objects. The solution

reached at this time entirely satisfied the cardinal criterion, roughly the ordinal criterion, and only partially the angular criterion.

(2) A second phase lasting until the end of the trial, in which the intermediate solution obtained at the end of the first phase was refined until satisfaction of all three criteria.

The cognitive processes involved in these two phases could be different, so the contralignment effect (in terms of duration) was measured in each of them.

Subjects and experimental design

The subjects were 18 female and male researchers or graduate students, from 25 to 45 years old. The two alignment/contralignment and three information conditions were crossed in a completely between-subject design, and three trials of each were presented in random order.

RESULTS

Results concerning the total time-to-location, that is, the time between the start and the end of a trial, are first reported. Then, a summary of the contralignment effect in each phase is given.

Total time

The results concerning the total time-to-location, that is, the time between the start and the end of a trial, are given in Figure 2.

The type of available information had a main significant effect [$F(2,34) = 21.99$, $p<.001$]. The time-to-location was longer in the C2 condition, in which only information about the subject scenes was presented, than in the C1 condition, in which the map was also presented [$F(1,17)=25.53$, $p<.001$]. It was also greater than in the C3 condition, in which only the map was presented [$F(1,17)=25.93$, $p<.001$]. Conditions C1 and C3 did not differ in time-to-location [$F(1,17)=.03$, ns].

Orientation of the target viewpoint had an overall significant effect: the time-to-location was longer in contraligned conditions [$F(1,17)=15.58$, $p<.005$]. There was no interaction between orientation of the target viewpoint and type of available information [$F(2,34)=2.81$, ns]. However, an effect of the orientation of the target viewpoint was observed respectively in condition C1 [$F(1,17)=14.4$, $p<.005$] and condition C2 [$F(1,17)=8.00$, $p<.025$], but not in condition C3 [$F(1,17)=1.53$, ns].

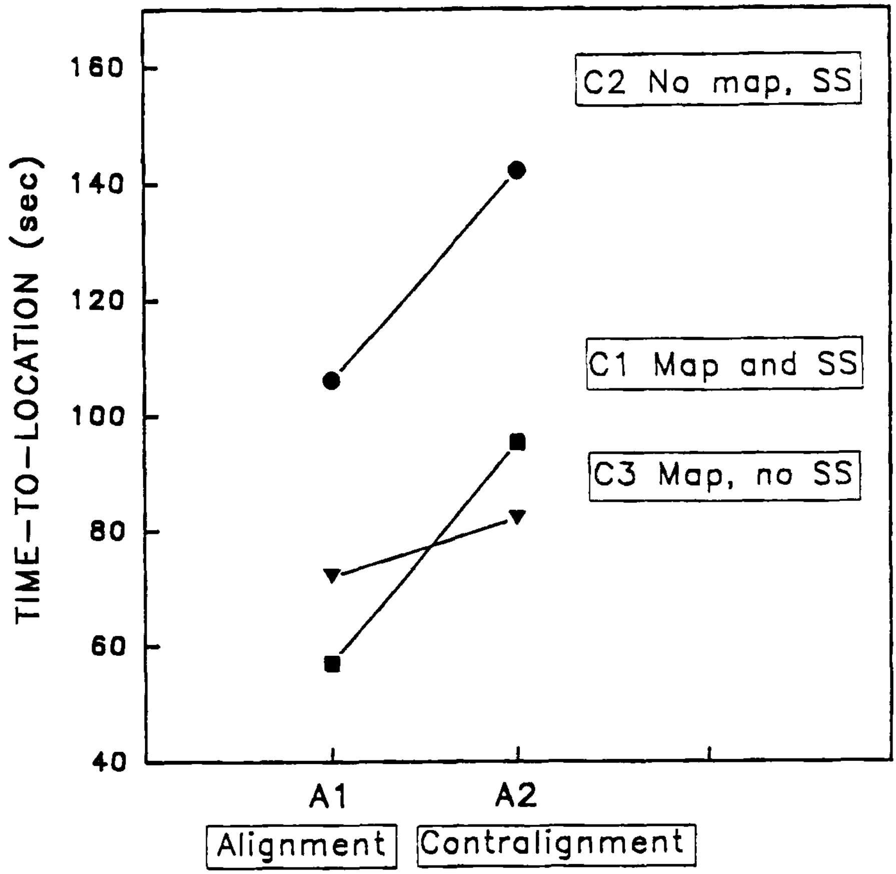

Figure 2. Mean total time-to-location (sec) for alignment/contralignment and information conditions (SS = Subject Scene).

Contralignment effect in the two phases

Table 1 sums up the presence or absence of a contralignment effect, in terms of duration, by phase and available information.

In the C1 condition there was always a contralignment effect: in both phases the time was longer when the target viewpoint was contraligned than when it was aligned. However, such an effect was found only in the first phase of the C2 condition, and only in the second phase of the C3 condition.

Table 1. Presence (+) or absence (-) of a contralignment effect by phase and available information (SS = Subject's Scene).

	first phase	second phase
C1 (Map & SS)	+	+
C2 (no Map, SS)	-	+
C3 (Map, no SS)	+	-

DISCUSSION AND CONCLUSION

In this experiment, following Levine et al. (1982) and Shepard and Hurwitz (1984), contralignment was defined by the fact that the point of view the subject had to find was facing downwards on the screen. In such a situation the subjects' difficulty was interpreted as resulting from the non-correspondence between the SFR of the subject watching the screen and the SFR of the viewpoint representing the subject on the map. Indeed, we found an overall significant effect (negative as regards the time to solution) of this contralignment. However this effect depended on the phase and the available information.

In the first phase the goal was to find an intermediate solution characterized by a relatively good orientation (the four correct objects with at least three in the right order had to be included) and a relatively poor position (there were also one or two incorrect objects in the subject scene). In this phase the contralignment effect appeared to be linked to the use of the information about the spatial layout of the objects on the map (since this effect was found in C1 and C3, i.e. when the map was displayed on the screen, and not in C2 when only the subject's scenes were available). Thus, it appears that the difficulty brought on by contralignment results from the discrepancy between the egocentric SFR of the subject watching the screen and the object-centered SFR of the viewpoint on the map, which involves a right-left inversion of the spatial location of the displayed objects on the map. The absence of a contralignment effect in C2 suggests that there was no need to use the same spatial information in different a SFR. We can think that the location of the viewpoint on the screen at the end of this first phase was reached by means of random or systematic exploration aimed at producing a satisfactory match between the target scene and the subject scene. In such a procedure it was not necessary for the subject to try to imagine the object layout (this was confirmed by the subject's post-experimental verbal reports).

In the second phase the task consisted essentially of refining the solution. This could involve some shifts of the viewpoint (the cursor) on the screen in order to improve the distance cues, and possibly remove some incorrect objects and eliminate one inversion still present in the subject scene. In this phase a contralignment effect was found in conditions C1 and C2 (in which subject scene information was available), and not in C3 (in which there was only map information). In the latter condition it can be assumed that the subject did not try to imagine the information about his/her produced (subject) scenes, but solved the problem by means of geometrical computations on the map without involvement of his/her own SFR. For the other two conditions, there was no strong discrepancy between the target scene and the subject scene since at the beginning of this phase the subject's viewpoint was globally in the correct orientation (and thus, the correct objects in the subject scene were nearly in the right order, with only one possible inversion between two of them). Thus, the contralignment effect did not result from a purely perceptual conflict between different SFR's in which the same information about the (target and subject) scenes had to be used. Here the movement of the mouse controlled by the subject, the corresponding displacement of the cursor (the viewpoint) on the map, and the consequences of these moves on the subject's scene must all be taken into account. In a contraligned situation (the viewpoint facing downwards on the screen) a displacement of the mouse to the right, when referred to the egocentric SFR of the subject facing the screen, results in a displacement to the right of the viewpoint on the screen if it is referred to the same SFR, but in a displacement to the left if it is referred to the object-centered SFR of the viewpoint on the screen. Moreover, such a displacement involves a shift to the right of the objects in the subject's scene. Thus, the conflict between the different SFR's may be rather perceptuo-motor, in the sense that subjects have to take into account the orientation of mouse displacements.

In conclusion, when the map information was used in the first phase of the locating task we found the same type of contralignment effect as Levine and colleagues (1982, 1984) and Shepard and Hurwitz (1984).

Here the contralignment problem seemed to be mainly a perceptual one. In the second phase, when the location process mainly dealt with the transformation of the egocentric information due to the mouse displacement, we identified another type of contralignment problem, this time a perceptuo-motor one. A more detailed analysis of the translation and/or rotation movements produced during the searching procedure should allow us to gain a better knowledge of the cognitive processes involved in such a locating task.

REFERENCES

Levine, M. (1982). You-Are-Here maps: Psychological considerations. *Environment and Behavior* **14**, 2, 221-237.

Levine, M., Jankovic, I.N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General* **111**, 2, 157-175.

Levine, M., Marchon, I., & Hanley, G. (1984). The placement and misplacement of You-Are-Here maps. *Environment and Behavior* **16**, 139-157.

Péruch, P., Pailhous, J., & Deutsch, C. (1986). How do we locate ourselves on a map: A method for analyzing self-location processes. *Acta Psychologica* **61**, 71-88.

Shepard, R.N., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition* **18**, 161-193.

This Page Intentionally Left Blank

Part 2

Imagery and Working Memory

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 4

Auditory imagery and inner speech

Daniel Reisberg
Reed College, Portland, Oregon, USA

Meg Wilson
University of California, Berkeley, USA
and
J. David Smith
New School for Social Research, New York, NY, USA

INTRODUCTION

The study of mental imagery is surely one of the great success stories of modern psychology. We know a vast amount about mental imagery, knowledge gained with impressive quantitative methods of experimentation, testing sophisticated theories. But this wealth of information is almost exclusively concerned with *visual* imagery, and remarkably little is known about all the remaining imagery modalities.

There are several reasons why we should work to correct this information imbalance. At the very least, we need to ask whether claims about visual imagery generalize to imagery in other modalities. Moreover, we believe that imagery in other modalities is interesting in its own right, and, in particular, auditory imagery seems to invite our close attention. Auditory imagery occupies an intriguing position, standing at the intersection of a number of phenomena, and a number of research domains. For example, auditory imagery seems implicated in memory, specifically for rehearsal processes. There has also been much discussion of the role of imagery in music, both for the composer and the listener. Likewise, there continues to be discussion of the role of silent speech in reading, again a potential role for auditory imagery. Consider also the discussions of the role of inner speech in guiding thought, discussions that appear prominently in Piaget, or Vygotsky.

Last but not least, there is an extensive literature on the delusional voices heard both by alcoholics and by schizophrenics, and we can ask how these are related to more mundane forms of auditory imagery.

These suggestions and phenomena obviously invite several questions: Is auditory imagery involved in each of these domains? And if imagery is involved, is it the same species in each domain, with a similar mode of representation, drawing on the same processes? Or might there be several species of auditory imagery? We do not want to merge our discussions of imagery in these various domains if this means marrying together phenomena that are truly diverse. At the same time, we do not want to proliferate forms of imagery, potentially one for each of the domains just mentioned. As will soon be clear, these questions were very much in our minds as we designed the research described here.

In this chapter, we will describe a series of studies, organized roughly historically, but also organized to make plain the evolution of our thinking about auditory imagery. We began thinking about auditory imagery merely as visual imagery's poor relation, a place to test the generality of claims made about visual imagery. This research has led us, however, to consider auditory imagery to be an important and intriguing entity, drawn on by many tasks. We hope that this chapter will persuade the reader that the latter view is indeed correct.

GENERALIZING CLAIMS ABOUT VISUAL IMAGERY

How do we come to understand the contents of our own mental images? On one view, images are much like stimuli, and so need to be interpreted via processes related to perception (Finke, 1980; Kosslyn, 1980, 1983). Alternatively, one can claim that images are best thought of as mental representations, and not stimuli (Casey, 1976; Chambers & Reisberg, 1985; Fodor, 1981; Kolers & Smythe, 1984; Reisberg & Chambers, 1991). On this latter view, images can still have depictive qualities; that is the sort of representations they are. But, as mental representations, images only exist in the context of the imager's understanding and knowledge. Therefore one does not have to interpret a mental image to find out what it represents; instead, the interpretation is there from the start.

These two conceptions differ in many ways, including their implications for the possibility of *ambiguity* within imagery. According to the view that 'images are like stimuli,' images can be ambiguous, just as pictures can be. According to the view that 'images are representations,' images come into being as representations of some particular content, and so are not, and cannot be, ambiguous.

Chambers and Reisberg (1985) evaluated these contrasting claims for imagery in the visual modality. Subjects were asked to memorize certain forms which included some of the classical ambiguous figures; the duck/rabbit, the Necker cube, and so on. Crucially, subjects were not familiar with these forms, so that we can ask whether subjects can discover the alternate reading of each form from imagery. Subjects were shown the figures briefly, then the pictures were removed, and subjects were asked to create an image of the form they had just seen, and to reinterpret that image if they could. Subjects were given a variety of hints and suggestions to help them out, and then, if all else failed, they were given a blank piece of paper, asked to draw the form they had just been imaging, and then to try reinterpreting their own drawing.

The results are easy to describe. Subjects universally failed to reinterpret their own images. They routinely succeeded, though, just a moment later, in reinterpreting their own drawings. These results indeed suggest that visual images are unambiguous. (For further discussion, see Chambers & Reisberg, 1985; Finke, Pinker & Farah, 1989; Reisberg & Chambers, 1991.)

Note that the claims at stake here are general ones, and should apply to imagery in other modalities. In particular, we can ask whether *auditory* images are unambiguous, just as visual images seem to be. This question was pursued in a series of studies reported by Reisberg, Smith, Baxter and Sonenshine (1989). As their ambiguous inputs, Reisberg et al. exploited the fact that certain words, if repeated over and over, yield a soundstream compatible with more than one segmentation (Warren, 1961, 1982; Warren & Gregory, 1958). For example, rapid repetitions of the word "life" produce a physical soundstream fully compatible with segmentations appropriate to repetitions of "life" or of "fly." These repetitions are usually perceived first as one of these words, then the other, then the first, changing in phenomenal form just as the Necker cube or duck/rabbit do. This allows us to ask if imagined repetitions produce verbal transformations, just as heard repetitions do.

In Reisberg et al.'s experiments, some subjects *heard* repetitions of the word "stress." Some of these subjects pronounced the repetitions themselves; others heard the experimenter repeat the word; still others heard a tape loop. In all of these conditions, 100% of the subjects quickly reported hearing the soundstream turn to repetitions of "dress." In contrast, another group of subjects was asked simply to *guess* what this word would turn into, if repeated over and over. These subjects rarely guessed the stress-to-dress shift, making this shift a good means of detecting bonafide reversals; easily (even universally) perceived, but difficult to guess.

Other subjects *imaged* a friend's voice repeating the test stimulus over and over; about half of them (46%) detected the stress-to-dress reversal. This is a lower level than when subjects heard the stimuli, but a higher level than when subjects guessed. This suggests a striking contrast between visual and auditory imagery: Subjects in this experiment were succeeding in reconstruing their auditory images, while subjects in the visual experiments routinely failed.

This result implies that auditory imagery does involve some neutral stimulus, capable of supporting reinterpretation, while visual imagery does not. It seems plausible that this neutral stimulus may be supplied by subvocalization. That is, perhaps the imagery subjects were talking to themselves, and so providing themselves with a covert stimulus, capable of supporting reconstrual. To explore this possibility, other subjects imaged the repetitions while chewing on a large piece of candy, to block use of subvocalization. With this chewing manipulation in place, the rate of stress-to-dress reversals dropped to that obtained by guessing subjects. Various control activities (such as asking subjects to imagine the relevant event while concurrently doing a rhythmic tapping task) did not have this effect (Reisberg et al. 1989). In another procedure, the number of image reconstruals was also sharply reduced, from 73% to 27%, by having subjects clamp their jaws shut, press their lips together and push their tongues up against the roof of their mouths (Wilson, Smith & Reisberg, unpublished). Thus, once again, with subvocalization blocked, the rate of image reconstruals was reduced to guessing levels.

WHAT KIND OF STIMULUS DOES SUBVOCALIZATION CREATE?

Apparently, image reconstruals do rely on use of the inner voice. But what is the role of the covert speech in this task? We have suggested that covert speech provides a kind of internal stimulus for the subject, an icon neutral to any interpretation, that can be reconsulted and reconstrued. But what sort of stimulus is this? One possibility is that subvocalization provides a kinesthetic stimulus, so that subjects note that repetitions of "stress" *feel* just like repetitions of "dress." Another possibility is that covert speech provides an implicit 'acoustic' stimulus, which is then processed through the channels of auditory perception.

We can distinguish these alternatives by simply presenting subjects with an irrelevant auditory input, while they are trying to do the stress/dress task. There is reason to believe that this use of the outer ear (i.e, actual hearing) prevents use of the "inner ear" (Baddeley, 1986; Salamé & Baddeley, 1982; Segal & Fusella, 1970). Thus, if subvocalization provides some sort of input to the processes of audition, then performance will be disrupted if these processes are occupied with other simultaneous inputs. If, on the other hand, subvocalization merely provides a kinesthetic cue, then there will be no need for the inner ear, and so irrelevant sounds played during the task will have no effect.

In asking this question, our experiments take on the familiar logic of a selective interference design, as depicted in Figure 1. That is, we can ask subjects to perform the stress/dress (or any other) task with no other requirements or manipulations. This is the top-left cell in the five patterns shown in the figure. Alternatively, we can ask subjects to perform the task while we take steps to block subvocalization, i.e., to disrupt use of the inner voice. (This would be the top-right cell in these patterns.) Or we can block use of the mechanisms of hearing, by means of an irrelevant auditory input, and in this way block use of the inner ear (the bottom-left cell). Or, finally, we can apply both kinds of interference simultaneously (the bottom-right cell in the patterns).

Which of these patterns describes performance in the stress/dress task? We already know that blocking subvocalization impairs performance in this task. What about blocking the inner ear? To determine this, we asked subjects to perform the stress/dress task while simultaneously hearing, through headphones, irrelevant auditory inputs (a voice reading to them); subjects were instructed to ignore this voice as best they could. With this manipulation, performance was again reduced to guessing levels. In this procedure, 73% of the subjects detected a reversal if merely instructed to image these repetitions; in contrast, only 13% heard the reversals if imaging in the presence of distracting sounds. In short, it seems that the inner ear is required by this task.

The stress/dress task thus produces results like those in Figure 1-D, a pattern we will refer to as the "partnership" pattern. Both the inner voice and the inner ear are needed for this task. If either is disrupted, performance drops to guessing levels. In essence, subjects seem to perform this task by pronouncing the stimuli to themselves, and then in some sense "hear" the results of this pronunciation.

How should one understand the idea of "hearing" in this context, since, to be sure, subjects produced no overt sounds? A number of investigators have proposed that the pathways used for speech production might be connected directly to central mechanisms of perception. These connections are indicated, for example, by evidence implicating covert speech production in speech perception. (Liberman & Mattingly, 1985, provide a recent review of this evidence.) Likewise, the

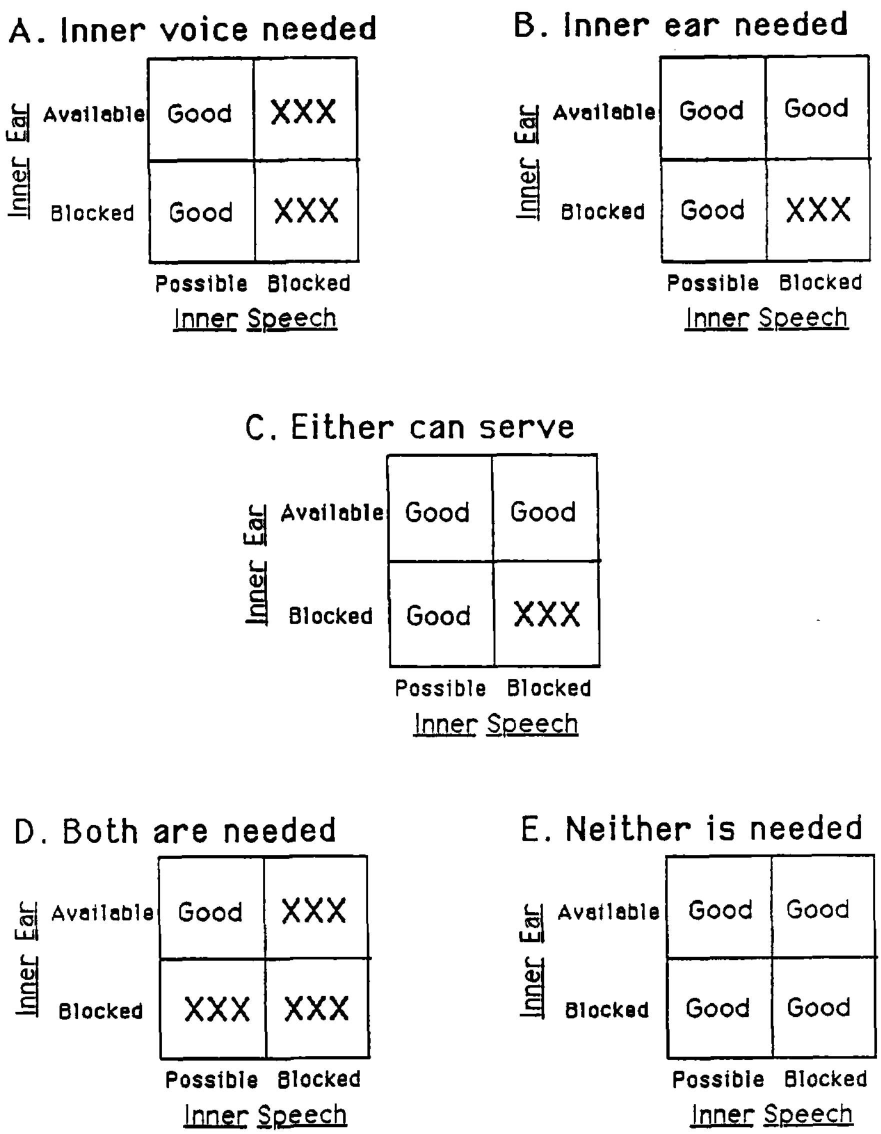

Figure 1: Possible performance patterns resulting from manipulations designed to block use of the inner voice or inner ear.

mechanisms for *planning* speech may also be linked to perceptual channels (either audition or proprioception); these linkages are hypothesized as means through which the speech produced can be checked for correctness (e.g., Zivin, 1986). In these ways, subvocalization (speech acts or plans) could load the registers for speech perception. This provides a plausible description of the kind of hearing involved in our imagery tasks, and would produce the "partnership" pattern we have described.[1]

HOW WIDESPREAD IS THE INNER-EAR/INNER-VOICE PARTNERSHIP?

Our early studies were designed in search of generalization data for visual imagery claims. And, in a sense, we have gotten those data; auditory images seem just as resistant to reinterpretation as visual images, once subvocalization is removed from the scene. But we have also gleaned something more interesting than mere generalization: a strong indication of a partnership between the inner voice and the inner ear. This pattern itself demands study. How widespread is this partnership? What kinds of tasks require this use of the inner voice to support imagery, and what kinds of tasks do not?

These questions set the research agenda for us over the next series of experiments. We began by examining another case of imagery for speech sounds. In one experiment, for example, we presented subjects with visual strings like these: "D2R" or "NCQR" or "DK." We asked subjects to imagine what these strings would sound like if pronounced aloud, and to report the familiar words thus produced: "detour", "insecure", "decay", and so on.

It seemed likely that subjects would use auditory imagery to perform this task. But to ask exactly how this task is performed, we returned to the selective interference logic already sketched. Subjects performed this task in one of four conditions. For some subjects, deciphering these visual strings was their only task. Other subjects attempted to decipher these strings while hearing irrelevant sounds through headphones (blocking use of the inner ear). A third group did this primary task while doing a concurrent articulation task (blocking use of the inner voice). Finally, a fourth group of subjects sought to decipher these

---

[1]We emphasize, however, that the connections between covert speech and perception are likely to be relatively central - i.e., not to require use of the articulatory muscles themselves (e.g., Baddeley & Wilson, 1985; Cole & Young, 1975).

strings while simultaneously hearing irrelevant sounds through headphones and also doing the concurrent articulation task.

Performance in decoding strings like "NCQR" or "DK"

Inner Ear		Inner Speech: Possible	Inner Speech: Blocked
	Available	73%	21%
	Blocked	40%	19%

Performance in pitch task

Inner Ear		Inner Speech: Possible	Inner Speech: Blocked
	Available	83%	66%
	Blocked	69%	68%

Figure 2: The top panel shows selective interference results when subjects try to decode visual strings such as "D2R" or "NCQR." The bottom panel shows selective-interference results when subjects are asked to make judgments about melody. Note that chance performance in this task would be 50% correct.

The data are clear-cut (Figure 2, top panel). We again find the "partnership" pattern. Subjects need to subvocalize to decipher these strings. If subvocalization is blocked, performance is crippled. But subjects are not merely reading kinesthetic information from this inner speech. If they were, then occupying the ears should not be disruptive, but clearly it is. Once again subjects seem to be talking to themselves, and then listening to hear what it is they have just said.

Is this pattern unique to tasks involving imagery for *speech*? To find out, we turned to the domain of music. Subjects were presented with the names of familiar tunes: *Happy Birthday, Mary Had A Little Lamb,* and so on. Subjects were asked, for each tune, whether the melody *rose* or *fell* from the second note to the third. Here a different result seems possible, namely the pattern shown in Figure 1-C. This is the pattern we would get if subjects had two different options in performing this task; either singing to themselves, and making judgments based on kinesthetic information (e.g., felt vocal tensions), *or* imagining the melody, perhaps in the voice of a friend, and making judgments based on this. If either one of these routes was blocked, subjects could use the other. Performance would then suffer only if both of these strategies were blocked.

The results, though, are quite different, and, in fact, quite familiar (Figure 2, bottom panel). This task, like the others we have described, yields the 'partnership' pattern. Subjects need both the inner voice and the inner ear, and blocking either substantially impairs performance. (Though performance seems still robust, even with our manipulations present, note that chance performance in this task would be 50% correct. Hence our manipulations did succeed in driving performance much of the distance toward chance levels.)

IMAGERY FOR TIMBRE

The results so far indicate that the inner voice plays a role in diverse auditory imagery tasks. However, it seemed to us implausible that subvocalization is *always* required for auditory imagery. After all, what about imagery for events one cannot vocalize, sounds outside of the human vocal repertoire, such as the sound of glass breaking, or automobile brakes squealing? We thought it plain that subjects can image these events, but it seems unlikely that this imagery is supported by subvocalization. The problem, though, is to test these claims in the laboratory.

In an ongoing series of experiments, we are exploring subvocalization's role in imagery for specific instrument timbres, building on recent work by Robert Crowder (Crowder, 1989; Crowder & Pitt, 1991). Crowder's elegant demonstration of imagery for timbre proceeded in two steps. First, subjects were *perceptually* presented with two instrument tones, say a trumpet playing an A, and a flute playing a B. Subjects' task was to say whether the two instruments were playing the same pitch, or a different one. Strictly speaking, timbre was irrelevant to the task. However, it turns out that there was a strong interaction between pitch and timbre. Crowder found reaction-time facilitation whenever the relation between the timbres of the two tones was congruent with the required pitch comparison (the same timbres facilitated same-pitch trials; different timbres facilitated different-pitch trials).

With this perceptual result in hand, Crowder turned to imagery. Subjects were presented with a sine wave playing at a particular frequency, and then were visually presented with the name of an instrument. Subjects were instructed to imagine what that instrument would sound like if playing the sine wave's pitch. Once subjects had formed the image, a test tone was presented, and subjects had to judge whether the test tone was or was not playing the same pitch as the one they were imaging. Again, timbre is, strictly speaking, irrelevant to this task. Nonetheless, the data revealed an interaction between pitch and imagined timbre that paralleled that between pitch and perceived timbre. This clearly suggests that timbre is somehow depicted, presumably in a quasi-perceptual form, in the image itself.

How are timbre images produced? Crowder has argued that the inner voice could not be involved in this task. This is because subjects cannot make trumpet noises, or flute noises. And if they cannot vocalize these sounds, then presumably they cannot subvocalize them either. Hence subvocalization should be irrelevant to performance. However, it seems likely that the inner ear is involved in this task: Given the parallel findings with imagery and perception, it seems plausible that subjects are somehow 'listening' to these imaged tones, via processes resembling those of audition.

We believe, however, that these claims must be evaluated empirically. A brief explanation of this seems worthwhile, since it will illuminate several points about the nature and (potential) function of subvocalization. Crowder is surely correct in asserting that subjects cannot vocalize convincing flute or trumpet noises. However, this by itself does not indicate the irrelevance of subvocalization to timbre imagery. As one possibility, note that subjects' vocalizations need not be accurate to support performance in this task. In order for subvocalization to influence performance, all one need assume is that a (subvocalized) trumpet noise resembles a trumpet more than it resembles a flute, and so forth.

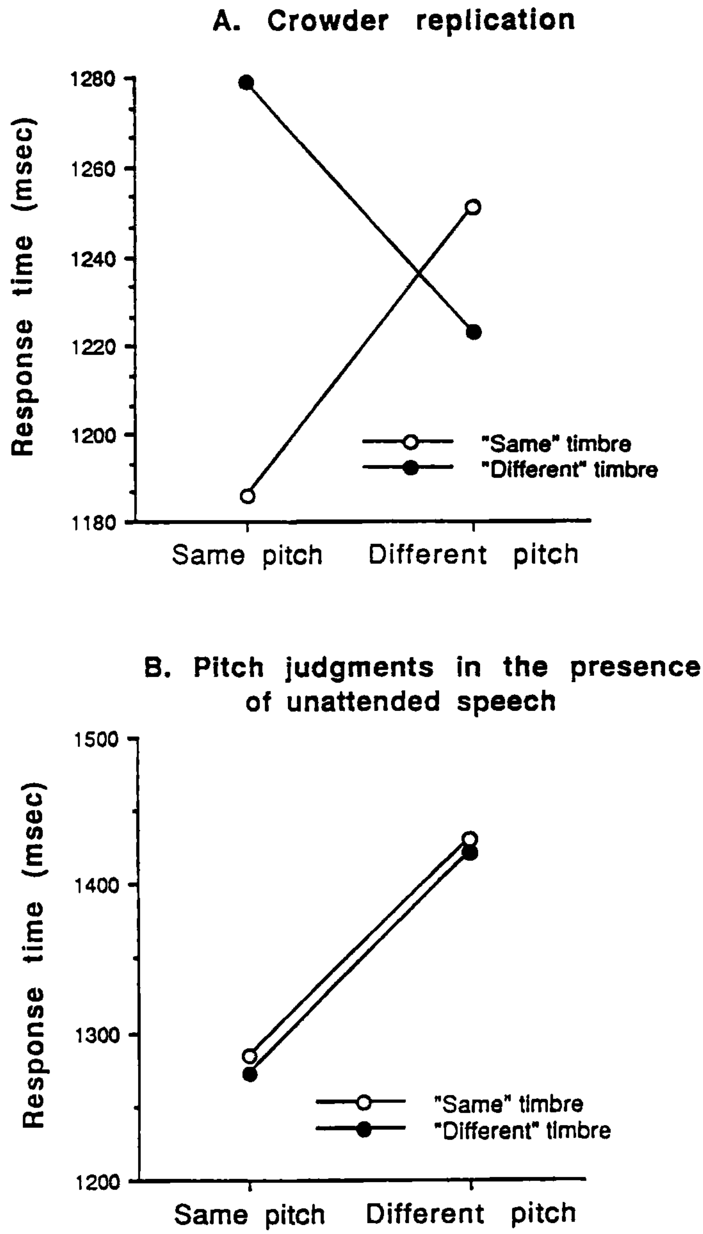

Figure 3: Reaction time (ms) for subjects to compare the pitch of imaged and actual tones. Panel A shows a replication of Crowder's (1989) findings. Panel B shows results from the same procedure run in the presence of unattended speech.

As an alternative, one might propose that the limitations on overt vocalization are greater than those on subvocalization. That is, perhaps one can subvocalize sounds (or at least *plan* to make these sounds) that one cannot vocalize. By analogy, one cannot actually lift an automobile, but one could at least plan the muscle sequence that would be involved in this action. For that matter, perhaps one could even pantomime some rough version of this action. In these ways, one can plan, or covertly produce, motor sequences outside of one's actual motoric repertoire. By the same token, perhaps subjects can plan, or covertly produce, sounds outside of their vocal repertoires.

Thus subvocalization *might* be implicated in imagery for instrument sounds, despite the fact that these sounds are, in fact, outside of subjects' vocal repertoire. Hence data are required to examine subvocalization's role in timbre imagery. All of this led Hespos (1989) to examine these issues, by combining the Crowder procedure already described with the selective interference logic. Hespos' first study was designed simply to replicate the Crowder result. Subjects heard a sine wave, then saw an instrument name, and imaged that instrument playing the pitch of the sine wave. They then heard a test tone, and had to judge whether the pitch of the test tone matched that of their image. The test tone was the same pitch on half the trials, different on half; the test tone matched the imaged timbre on half the trials, so that (for example) subjects were imaging a trumpet and the test tone was a trumpet, and did not match on half the trials. As Figure 3A shows, Hespos obtained a robust interaction between imaged timbre and pitch.

What happens to the data pattern if use of the inner ear is blocked? In this procedure, subjects heard, through headphones, a voice reading from an anthropology textbook. The voice was silent during the presentation of the sine-wave, then resumed, then was silent again during presentation of the test tone. Thus we can be sure that the voice did not disrupt perception of the test stimuli themselves, but the voice was otherwise continuously present. As Figure 3B shows, Hespos found that this manipulation obliterated any effect of imagined timbre. Apparently, the imagery revealed in this effect is dependent on the inner ear, and so is disrupted if use of the inner ear is blocked.

What about the inner voice? Hespos had subjects do Crowder's task while saying 'tah-tah-tah' aloud, to block use of the inner voice. Subjects in the control group did Crowder's task while tapping with their fingers. (This allowed Hespos to control for any general distraction effects from the concurrent task.) Summing across these two groups, the data replicate Crowder's finding of a pitch x timbre interaction; that is, the data show the effects of imaged timbre. Critically, this pattern is visible in both the 'tah-tah' and tapping groups. (That is, there is no interference type x pitch x timbre interaction.) Hence it seems that the inner voice is *not* needed for this task.

We note in passing that a different conclusion was implied by some of Hespos' earlier results (including results described in Smith, Reisberg & Wilson, 1991). In these early procedures, manipulations designed to block use of the inner voice eliminated the pitch x timbre interaction, suggesting that the inner voice was needed to support timbre imagery. However, the more recent data indicate that we should attribute these early findings to some sort of generalized distraction effect, rather than to specific blocking of the inner voice.

Given these early findings, though, we would urge caution in interpreting our timbre imagery results. (In addition, the current studies are still underway, again a reason for caution.) Nonetheless, the data do indicate, overall, that timbre imagery requires the inner ear, but not the inner voice. And this raises a question for us: If we set the cautions to the side, it seems that one can retrieve information from memory (e.g., information about timbres), and use this information to create an auditory image directly, without use of the inner voice. That is, subjects do not *need* the inner voice to create auditory images. But then why do our other data reveal such an important role for the inner voice? We will approach this question in three steps: First, we consider further evidence that the inner-voice/inner-ear partnership does indeed appear in a broad range of tasks. Second, we present some further tasks that do *not* show this partnership pattern. That puts us, finally, in a position to address why some tasks show this pattern, while others do not.

CONVERGING EVIDENCE FOR THE INNER-VOICE/INNER-EAR PARTNERSHIP

Auditory imagery experiments reveal an important role for the inner-voice/inner-ear partnership. As it turns out, evidence from rather different domains also indicates the same pattern. First, there is clearly a close convergence between our language, in describing auditory imagery, and language already in the literature for describing rehearsal processes in working memory. Indeed, our thinking about these issues, and, for that matter, our terminology, have been heavily influenced by studies of working memory, largely Baddeley's work (e.g., many of the studies reviewed in Baddeley, 1986). Ordinary memory rehearsal shows exactly the partnership pattern we have been describing. Rehearsal is disrupted if use of the inner voice is blocked, and also disrupted if use of the inner ear is blocked. In addition, blocking both of these does no more harm than blocking just one, precisely the pattern that signals the inner voice and inner ear working in concert with each other. This convergence between our work and Baddeley's is clearly telling us something important about the interrelation between working memory and auditory imagery, or between rehearsal and imagery, but what is crucial here is that, once again, we are seeing the partnership pattern. To put it in rough terms, one rehearses material in working memory by talking to oneself, and then listening to what one has said.

Surprisingly, the hallucinations of schizophrenics may require a similar explanation (Smith, 1991). Many theories have taken these auditory hallucinations to be sensory/perceptual phenomena, received onto the sensory platforms much as sensory stimulation is. However, these accounts fail to explain several observations. Why do voices so dominate the schizophrenic's hallucinatory experience, rather than visions, or other kinds of sounds? Why do the voices seem to the schizophrenic to reside "inside" the mind or head? Why do schizophrenics never wonder where the voices come from, and never try to approach or find them?

In explaining these observations, suppose that schizophrenics produced hallucinated voices by subvocalizing them, and then 'heard' the hallucinatory speech images that resulted. This would, of course, imply something just like the inner-voice/inner-ear partnership we have been discussing throughout, although a partnership in this case with bizarre consequences. And this partnership would readily explain all of the observations about schizophrenic hallucinations just mentioned.

In fact, electromyography often seems to reveal articulatory activity correlated with hallucinations. Moreover, various therapeutic interventions (and even controlled studies) have reduced or eliminated these hallucinations by having schizophrenics clamp the articulators much as we do, in our experiments, to block the inner voice. Thus it seems likely that schizophrenics employ the inner-voice/inner-ear partnership to create their own hallucinations. As one schizophrenic summed up this process, and his hallucinations, "Somebody is attacking me and my lips move."

WHAT TASKS DO NOT SHOW THE PARTNERSHIP?

The partnership between the inner voice and inner ear is widespread, but not universal. We therefore need to ask when this partnership is needed, and, correspondingly, what function the partnership serves. In moving us toward answers to these questions, we next present two types of tasks for which the partnership is not necessary.

Performance in judging whether words spelled with "s" are pronounced with "s" or "z"

Inner Ear		Inner Speech: Possible	Inner Speech: Blocked
	Available	90%	78%
	Blocked	94%	85%

Performance in judging whether words spelled with "ed" are pronounced with "ed" or "t"

Inner Ear		Inner Speech: Possible	Inner Speech: Blocked
	Available	84%	70%
	Blocked	85%	77%

Figure 4: Selective interference results when subjects judge how the ending of (visually) presented words would be pronounced - whether with an 's' or 'z' (top panel), or whether with an 'ed' or 't' (bottom panel).

In one kind of task, we asked subjects to make judgments about word endings. For example, subjects were instructed that: "Some words that end with the letter 's' sound as if they end with a 'z'. In the following list, imagine hearing each word spoken aloud, and mark the ones that sound to you like they end with 'z'." Subjects were then given a list of words, visually presented, such as larks, dogs, tabs, and so on, and rendered their judgments. In a parallel procedure, subjects were told that some words that are spelled with an "ed" at the end sound as if they have a "t," and, again, subjects were given a visually-presented series of words, and had to judge how these would be pronounced, were they to be said aloud.

Figure 4 shows the results for these judgments. Irrelevant auditory inputs did not disrupt performance, indicating that the inner ear is not involved here. The data indicate, though, that the inner voice *is* involved, although does not seem crucial. That is, performance dropped when subvocalization was blocked, but nonetheless remained quite high.

Thus these judgments do seem to draw on stimulus support from the inner voice, but the stimulus provided seems to be a kinesthetic one. That is, subjects pronounce these stimuli to themselves, and then use the kinesthetic feedback from this pronunciation to make their judgments. This would explain why subjects do not need to listen to this pronunciation with the inner ear, hence the lack of any effect of unattended speech. Having explained the s/z and t/d tasks in this way, one might even suggest that they are not tasks, or rather need not be tasks, of auditory imagery proper. They are more like judgments about the behavior of one's articulators. Perhaps kinesthesis alone is sufficient for these judgments since they rely on detecting a single simple property of the vocal chords (which is the main difference between pronunciation of final s and z, or t and d). This is a point however on which further research is needed.

A different kind of task in which the inner-ear/inner-voice partnership breaks down is linked to the processes of reading, and to the decoding of print into sound. There has been much debate about the role of "inner speech" in reading. In particular, does one read by somehow translating visual strings into semantic representations, going directly "from sight to meaning?" Or must one first translate the visual strings into phonological representations, a covert version of what one does in reading aloud? These phonological representations would then be processed through mechanisms related to those of ordinary perception.

Evidence indicates that both of these channels are available for readers, but what is critical for our purposes is this: Readers are able to translate print into some mental representation of sound. Consider the fact that we can make judgments about rhyme, with visually presented stimuli, such as the judgment that these words would rhyme: "BLUE" and "TOO". These judgments obviously hinge on phonological representations more than on visual form. Perhaps most persuasively, we can read pronounceable nonsense words; we can tell, for example that the string, "PHLAIM", shares its sound with an ordinary English word. Likewise, we can judge that the following two strings would sound alike if pronounced aloud: "HEJ" and "HEDGE". These judgments cannot rely on familiarity with specific visual strings, since these strings are, after all, made up for the occasion. Instead, the judgments must hinge on some mental representation of sound.

But what is this mental representation of sound? What is the relation between it and the kinds of imagery we have been discussing? One obvious way to address these questions is to use, once again, the selective interference logic. Relevant data are already in the literature, published by a number of authors, and can be easily summarized: Articulatory suppression, that is blocking use of the inner voice, has some impact on judgments about rhyme, but essentially no impact on judgments about homophones, with visually presented materials (Baddeley & Lewis, 1981; Besner, 1987; Besner & Davelaar, 1982; Besner, Davies & Daniels, 1981; Richardson, 1988; Wilding & White, 1985). This has been demonstrated in a number of contexts, and even with less-skilled readers, that is: children (Mitterer, 1982).

An obvious reading of this is that homophone judgments rely on "pure" auditory images, that is, representations residing in the inner ear without a contribution from the inner voice. But this suggestion turns out to be wrong. If it were correct, homophone judgments should be disrupted if we disrupt use of the inner ear, by the standard move of presenting irrelevant auditory distractors. But this experiment fails. Baddeley and Salamé (1986) found that this manipulation left homophone judgments untouched, implying that the inner ear is not required for these judgments.

Perhaps, though, these homophone judgments can be made in more than one way. Perhaps subjects can make homophone judgments using *either* the inner voice *or* the inner ear, and simply switch strategies depending on which variety of interference is on the scene. Then their pattern would be that of Panel C in Figure 1. In this case, subjects would be able to cope with either kind of interference alone. Performance would only fall if both interference tasks were delivered simultaneously, since then strategy switching would not be possible.

The literature tells us that judgments of homophony survive manipulations of either the inner ear or the inner voice. However, to our knowledge, subjects had never been run under the double phonological jeopardy relevant to distinguishing Panel C in Figure 1 from Panel E. We have recently therefore run the relevant experiment. As Figure 5 shows, latency was unaffected by either interference task. Error rates were unaffected by irrelevant speech heard, but were increased slightly by blocking subvocalization. What is critical for our purposes, however, is that performance stayed robust even given both kinds of interference simultaneously; subjects performed quickly in that condition, and with 82% accuracy, only 7% below their accuracy with no interference.

Inner Ear	Inner Speech: Possible	Inner Speech: Blocked
Available	30.5	30.8
Blocked	29.8	31.6

Time to complete 24 judgments (seconds)

Inner Ear	Inner Speech: Possible	Inner Speech: Blocked
Available	89%	83%
Blocked	89%	82%

Percent Correct

Figure 5: Selective interference results with a homophone task. The top panel shows response latencies (sec); the bottom panel shows accuracy data (percent correct).

Homophone judgments, therefore, fit the pattern of Figure 1-E, requiring neither the inner voice nor the inner ear. Yet these tasks do seem to involve a representation of sound. And, interestingly enough, subjects routinely claim that they make these judgments by "hearing," in the "mind's ear," what the strings sound like. On these grounds, it does seem that homophone judgements involve auditory imagery of some sort, but apparently not the sort evident in all of our other tasks!

THE INNER EAR, THE INNER VOICE, AND THE LEXICAL EAR

What shall we make of all this? The beginnings of relevant theory are in the literature, largely thanks to Besner (1987). The idea, roughly, is this: We have been discussing the relationship between the inner voice and the inner ear. Let us add to this duet a third member: a "lexical ear," with the following properties. The lexical ear, like the inner ear, is a specialized means of representing information about sound. Unlike the inner ear, the lexical ear does *not* draw on processes closely allied to those of hearing, and so is not disrupted by auditory inputs. Finally, and critically, the lexical ear is rather limited in its use: It can access whole-word units from long-term memory. Apparently, it can also hold representations created based on phonological knowledge. (E.g., consider the fact that this store can hold phonological representations of nonsense sounds, such as HEJ or PHLAIM.) However, any post-assembly operations performed on these units, such as segmentation, or even maintenance, will not be possible. If a task requires these operations, then the task will require the inner ear, and not the lexical ear. In this case, given the apparent relation between the inner ear and inner voice, performance is likely to show the partnership pattern: blocked by concurrent articulation, as well as by irrelevant auditory inputs.

This conception merges readily with the data surveyed here. Homophone judgments require no analysis, only the activation of a single logogen in LTM. Thus they require only the lexical ear, and are not disrupted by manipulations of inner ear or inner voice. In contrast, judgments about rhyme require analysis, as one must strip away part of a word to make a judgment about the remainder. Therefore, rhyme judgments cannot be supported by a representation in the lexical ear, and instead need the inner-ear/inner-voice partnership. Thus rhyme tasks are hurt by manipulations of inner ear or inner voice. This conception is also consistent with the finding (e.g., Kimura & Bryant, 1983) that articulatory suppression disrupts children's spelling, even though it does not disrupt their reading. This is sensible on the assumption that the former, but not the latter, requires analysis of phonological codes, and so the former, but not the latter, requires the inner-ear/inner-voice partnership.

Likewise, consider the detection of auditory ambiguity, as in the "stress/dress" task we mentioned at the very outset. This task requires that a phonological entity be created, but then the subject must set aside his or her initial intent in this creation, in order to resegment the sound stream, precisely the sort of thing that cannot be done with representations in the lexical ear. This is a case in which the subject needs something like an actual stimulus, existing independently of how the subject understands it. And this is exactly what subvocalization (that is, the inner voice) provides. It is not surprising, therefore, that this task is reduced to guessing levels by manipulations of the inner voice or inner ear.

Thus there is diversity within the set of tasks requiring judgments about sound. All these tasks require that a mental representation of sound be created. The tasks differ, though, in how much they require that one "detach" oneself from this representation. For some tasks, one creates a representation of a sound, and then uses it intact and unchanged. For these tasks, one does not need any sort of "stimulus support," and so one can use the lexical ear. For other tasks, one creates a representation, but then must set one's understanding of it to the side; one must "disown" the representation, so to speak, to hear it as a stranger might. In this case, one must create an actual stimulus, and then observe it. The inner-voice/inner-ear partnership is one means through which this "stimulus-support" can be provided, but there are others. For example, in our s/z and t/d tasks, one creates a *kinesthetic* stimulus, and then perceives it through the channels of proprioception. We suspect the same general principle, and the importance of self-produced stimulus support, also applies to imagery in other modalities (cf. Reisberg, 1987), but this is clearly a point in need of further research.

As another direction for future research, we note that all extant demonstrations of the lexical ear involve imagery for speech sounds, imagery drawing on phonological knowledge, and, in some cases, on knowledge of spelling-sound correspondence rules. But, on our view, this storage may also be useful when other sorts of "whole sounds" are to be activated, provided that no post-assembly analysis of these representations is required. In this case, *lexical* ear may be too narrow a label for this store. For example, perhaps musicians can activate the sounds of single chords, just as readers activate whole-word logogens; the musicians could then make judgments about these representations. (Note, though, that the inner ear would be required if the judgment demanded analysis of the chords into component notes.)

To our knowledge, no firm evidence speaks to this suggestion, although consider the following anecdotal information: If music can be represented in the lexical ear, then the inner ear will not be needed to access or support this representation. And, in this case, disruption of the inner ear (via presentation of extraneous sounds) should have no impact on this representation. (This follows from how we have defined the lexical ear.) Consistent with this, some skilled musicians report being able to contemplate a piece of music even in relatively noisy environments, indeed even while overtly hearing some other piece of music. This seems to be a case of auditory imagery drawing on mechanisms separate from those of hearing, and, consequently, of auditory imagery immune to manipulations of the inner ear. This is, of course, the profile one would expect if it is the lexical ear, not the inner ear, in use.

Clearly experiments are needed before we can make anything of such anecdotes. Be that as it may, the framework we have proposed apparently does provide an interesting way to think about musical expertise (or expertise in any auditory realm), and so seems an attractive avenue for further research.

In this chapter, we have surveyed a range of phenomena, consistent with our initial observation that auditory imagery is a particularly rich domain, with ties to many research areas. These ties have clearly shaped our research, inasmuch as we have been led to consider data from memory, schizophrenia, speech perception, music, and reading processes. It is striking that a reasonably consistent pattern emerges across these domains. We take this as a strong indication that auditory imagery may be a mental resource with broad application, and therefore with considerable importance.

REFERENCES

Baddeley, A. & Lewis, V. (1981). Inner active processes in reading: The inner voice, the inner ear and the inner eye. *In* A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* pp. 107-129. Hillsdale, NJ: Erlbaum Associates.

Baddeley, A. & Salamé, P. (1986). The unattended speech effect: Perception or memory? *Journal of Experimental Psychology: Learning, Memory and Cognition* **12**, 525-529.

Baddeley, A. & Wilson, B. (1985). Phonological coding and short-term memory in patients without speech. *Journal of Memory and Language* **24**, 490-502.

Baddeley, A.D. (1986). *Working Memory*. Oxford: Clarendon Press.

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *Quarterly Journal of Experimental Psychology* **39A**, 467-478.

Besner, D. & Davelaar, E. (1982). Basic processes in reading: Two phonological codes. *Canadian Journal of Psychology* **36**, 701-711.

Besner, D., Davies, J. & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *Quarterly Journal of Experimental Psychology* **33**, 415-437.

Casey, E. (1976). *Imagining: A phenomenological study*. Bloomington, IN: Indiana University Press.

Chambers, D. & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance* **11**, 317-328.

Cole, R. & Young, M. (1975). Effect of subvocalization on memory for speech sounds. *Journal of Experimental Psychology: Human Learning and Memory* **1**, 772-779.

Crowder, R. (1989). Imagery for musical timbre. *Journal of Experimental Psychology: Human Perception and Performance* **15**, 472-478.

Crowder, R. & Pitt, M. (1991). Research in memory/imagery for musical timbre. *In* D. Reisberg (Ed.), *Auditory imagery*. Hillsdale, NJ: Erlbaum Associates.

Finke, R. A. (1980). Levels of equivalence in imagery and perception. *Psychological Review* **87**, 113-132.

Finke, R., Pinker, S. & Farah, M. (1989). Reinterpreting visual patterns in mental imagery. *Cognitive Science* **13**, 51-78.

Fodor, J. (1981). Imagistic representation. *In* N. Block (Ed.), *Imagery* pp. 63-86. Cambridge, MA: MIT Press.

Hespos, S. (1989). The characteristics of pitch, timbre and loudness in auditory imagery. Unpublished bachelor's thesis, Reed College, Portland, OR, USA.

Kimura, Y. & Bryant, P. (1983). Reading and writing in English and Japanese: A cross-cultural study of young children. *British Journal of Developmental Psychology* **1**, 143-154.

Kolers, P. & Smythe, W. (1984). Symbol manipulation: Alternatives to the computational view of mind. *Journal of Verbal Learning and Verbal Behavior* **23**, 289-314.

Kosslyn, S.M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.

Kosslyn, S.M. (1983). *Ghosts in the Mind's Machine*. New York: W.W. Norton & Company.

Liberman, A. & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition* **21**, 1-36.

Mitterer, J. (1982). There are at least two kinds of poor readers: Whole-word poor readers and recoding poor readers. *Canadian Journal of Psychology* **36**, 445-461.

Reisberg, D. (1987). External representations and the advantages of externalizing one's thought. *In* E. Hunt (Ed.), *The Ninth Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum Associates.

Reisberg, D. & Chambers, D. (1991). Neither pictures nor propositions: What can we learn from a mental image? *Canadian Journal of Psychology.*

Reisberg, D., Smith, J.D., Baxter, D.A. & Sonenshine, M. (1989). "Enacted" auditory images are ambiguous; "Pure" auditory images are not. *The Quarterly Journal of Experimental Psychology* **41A**, 619-641.

Richardson, J.T.E. (1988). Phonology and reading: The effects of articulatory suppression upon homophony and rhyme judgments. *Language and Cognitive Processes* **2**, 229-244.

Salamé, P. & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior* **21**, 150-164.

Segal, S.J. & Fusella, V. (1970). Influence of imaged pictures and sounds in detection of visual and auditory signals. *Journal of Experimental Psychology* **83**, 458-474.

Smith, J.D. (1991). The auditory hallucinations of schizophrenia. *In* D. Reisberg (Ed.), *Auditory imagery.* Hillsdale, NJ: Erlbaum Associates.

Smith, J.D., Reisberg, D., & Wilson, M. (1991). The role of inner speech in auditory imagery. *In* D. Reisberg (Ed.), *Auditory imagery.* Hillsdale, NJ: Erlbaum Associates.

Warren, R. (1961). Illusory changes of distinct speech upon repetition - the verbal transformation effect. *British Journal of Psychology* **52**, 249.

Warren, R. (1982). *Auditory Perception.* N.Y.: Pergamon Press.

Warren, R. & Gregory, R. (1958). An auditory analogue of the visual reversible figure. *American Journal of Psychology* **71**, 612-613.

Wilding, J. & White, W. (1985). Impairment of rhyme judgments by silent and overt articulatory suppression. *Quarterly Journal of Experimental Psychology* **37A**, 95-107.

Zivin, G. (1986). Image or neural coding of inner speech and agency? *The Behavioral and Brain Sciences* **9**, 534-535.

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 5

Visuo-spatial interference and apperception in chess

Pertti Saariluoma
University of Helsinki, Finland

INTRODUCTION

Chess players' thinking is essentially "seeing" (Chase & Simon, 1973a,b; de Groot, 1966). Good chess players' are able to calculate longer variations than less skilled chess players, though they do not always need to do so (Charness, 1981; de Groot, 1965; Holding, 1989; Saariluoma, 1990). As the sizes of the basic problem spaces in chess usually comprise millions of paths, a difference of a move or two in the length of an episode can, however, explain very little of the skill differences and the mechanisms of search. Grandmasters are not good players because they calculate like computers, but because they "see" better moves in the large and complex jungles of alternative paths (de Groot, 1965, 1966; Chase & Simon, 1973a,b; Newell & Simon, 1972).

The psychological interpretation of skilled chess players' search process has proven to be difficult and controversial (Chase & Simon, 1973; Hartstone & Wason, 1983; Holding, 1985, 1989; Newell & Simon, 1972). Saariluoma (1990) has suggested a model in which the recognition of familiar configurations of pieces, is thought to abstract a small hypothetical subspace of the total problem space and the search process takes place in this subspace. The main motivation for search is to prove or disprove the subspace. Because the abstracted problem spaces are based on learned conceptual structures and they control the chess players' perceptual world up to the point of fixation, the model suggests that chess players' "see" good moves by a process of apperception.

By apperception I refer to conceptual or concept based perception, or second order perception (Leibniz, 1704; Neisser, 1987; Saariluoma, 1990; Wundt, 1920). Every human being with normal vision is able to perceive chess pieces, but only skilled chess players are able find those beautiful and often emotionally highly satisfying moves and problem subspaces which are characteristic of high level chess and this is why it is very important to make a distinction between object perception and more complex conceptual perception, i.e. between perception and apperception.

Apperception is an elusive process and to study it we must exceed the limits of single level planning research. The traditional concepts of planning, such as plan, search, operation, heuristic rule etc. are not alone sufficient in analysing subspace abstraction, because they do not provide information about the cognitive conditions of apperception. They do not help in answering the question how apperception is possible. How is it that skilled chess players are able to construct a visual world which is so totally different from the one directly provided by the simple object level perception?

This question can only be answered if the role of underlying processing system is known in sufficient detail. This presupposes multifunctional analysis of chess players' information processing. It is necessary to analyse the cognitive conditions of planning, such as attention, long term memory, working memory and visual imagery to understand observable behaviour during thinking tasks.

With this kind of approach it is also necessary to divide a total process into a number of subcomponent skills and consider the underlying processing resources of these subcomponent skills. This can be done either by simultaneous measurement, as was done by Logie, Baddeley, Mane, Donchin and Sheptak (1989) or by analysing the subcomponent skills in separate experiments (Saariluoma, 1990). In chess the latter approach is more practical though it does demand more experimental work and does not provide such direct comparisons as simultaneous measurement.

Chess players' calculation cannot be based on perception alone, because chess players generate moves in chess protocols that are never made on the board. This is not possible without some mediating representation. Direct object perception is not sufficient but the visible positions must be transformed into some format which allows players to distance themselves from current visual input. What this format could be and what is the processing system that supports these cognitive operations, is the basic question of this paper.

It has been argued that the mediating representation must be basically propositional (Holding 1985; Pfau & Murphy, 1988) but recent research provides also some evidence for imagery processing within the framework of working memory (e.g. Baddeley, 1986; Baddeley & Hitch 1974). Bradley, Hudson, Robbins and Baddeley (1987) have found that an articulatory suppression task, (i.e. repeating aloud an irrelevant word), does not interfere with the memory for chess positions, while the visual and central executive tasks decrease substantially the percentage of recalled pieces (See Logie & Marchetti, and Quinn, in this volume for description of working memory). These results also can be generalized to chess players' tactical play and strategic evaluation processes (Robbins, personal communication). The time to find a two ply, i.e. one move by black and one move by white, mate continuation is clearly longer under a continuous visual task (Brooks, 1968) than in control conditions, while articulatory suppression causes no impairment in performance (Saariluoma, in press).

These results provide the first indirect evidence for subspace processing in visuo-spatial working memory. There is not yet, however, any more direct evidence on the issues. The experiments on visual search for chess pieces chess pieces show only that information intake is impaired by concurrent visuo-spatial suppression, but such evidence does not provide direct proof that visuo-spatial processing is involved in thought processes. Only protocol analysis can provide direct evidence for the involvement of visuo-spatial working memory in thinking in chess.

EXPERIMENT 1

Experiment 1 was designed to investigate the effects of imaging and articulatory suppression tasks on chess players' problem solving. A modified version of the Brooks' (1968) task was used and the effects were measured by standard protocol analysis (Ericsson and Simon, 1984). The main measure was the fluency of thought. By fluency of thought I refer to the average time a subject needs to generate one move in the protocol. The more fluent he/she is the less time he/she needs per move. If subjects have greater difficulty in calculating variations in moves in visuo-spatial secondary task conditions than in other conditions, this disruption should be reflected in the fluency of thought.

Method

Subjects

Eight male subjects were used, whose native language was Finnish. Four of them were experienced chess players with 2000 SELO points at least and the rest were moderately skilled players with less than 1750 SELO-points or no rating. The SELO-system is calibrated with the ELO system; in SELO-system the theoretical mean is 1700 points and the standard deviation 200 points.

The ELO system calculates for each registered player an expected strength value on the basis of the tournament results. The principle is very simple; the stronger a player one is able to beat, the more points are awarded, and the weaker a player one loses to, the more points are deducted. Thus, each player will find his/her level within a few games and the skill differences are very objective (Elo, 1978).

Stimuli and design

Nine chess positions entailing a combination of pieces were presented to the subjects. Three of them were involved no secondary task, two involved concurrent articulatory suppression and two involved an imaging task. The order of presentation was random, but it was counterbalanced over subjects. With the articulatory tasks subjects were asked to repeat a two-syllable nonsense word. The word was different for each position: TIHTOH, KIKLEK, and SOHPAH. With imaging tasks subjects were asked mentally to walk around a letter pair and classify turning left or right at the corners. Again there was one pair for each position and they were HA, KV, and LE.

Normal protocol analysis techniques and the thinking aloud method could not be used, because the secondary task conditions required overt articulation. This is why subjects were asked to analyse the positions by moving the pieces on the board. Of course, this is a very common form of behaviour for chess players. They often analyse opening variations, games and adjourned games in this way.

The subjects' hand movements were videotaped. When analysing the results only the real moves were counted. The hand movements used to put the pieces back on their original squares were not registered. The number of generated moves were counted in each protocol and divided by the solution time. This proportion was termed the fluency of thought. The maximum time allowed was five minutes per position.

Results and Discussion

The tasks proved to be quite difficult. The fluency of thought measurements are presented in Figure 1.

Skill differences were statistically significant ($F(1,6)=10.09$, $p<.05$) in that times per generated move for experienced players were shorter than for moderately skilled players. This result is interesting because it sheds light on one very important question in the psychology of skilled chess.

Since de Groot's (1965) studies it has been known that skilled chess players do not calculate essentially more moves than the less skilled, unless they have to do so. The results of the first experiment show that skilled chess players are, however, much faster in generating moves and the fluency of their thought is essentially higher that of the less skilled players. The interaction between skill and the conditions was not significant (F(2, 12)<1).

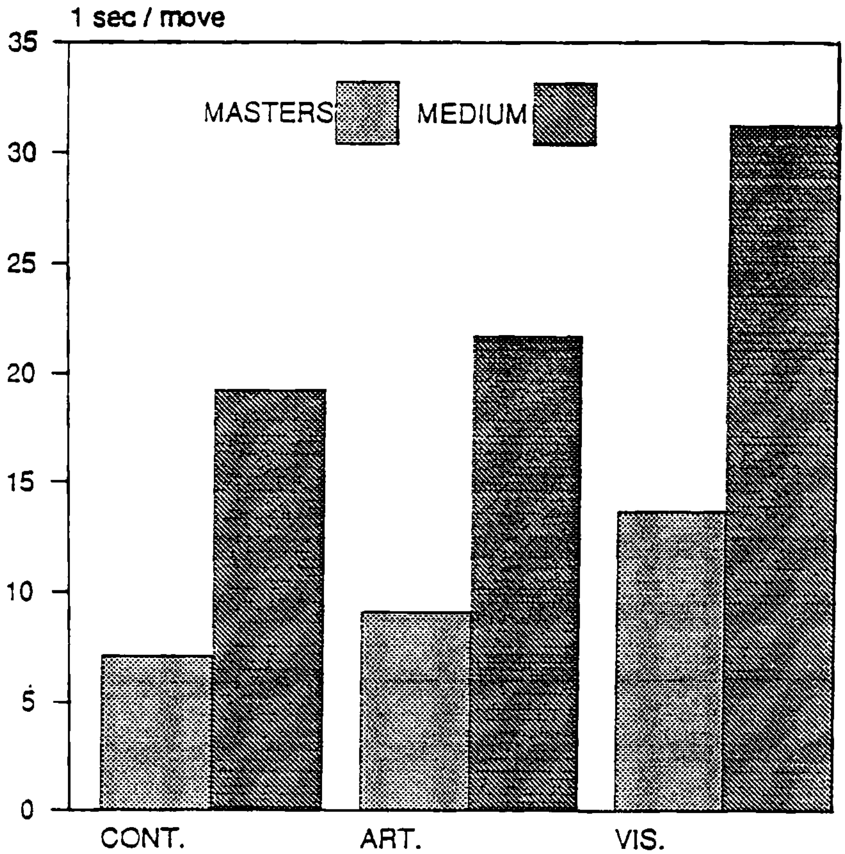

Figure 1. The influence of visuo-spatial (vis.) and articulatory (art) secondary tasks compared to control (cont.) on chess players' generation of moves.

The overall effect of the secondary task conditions was significant ($F(1,12)=4.60$, $p<.05$). The comparison of means showed that imaging suppression impaired subjects performance most. A Newman-Keuls test showed that the statistical difference between imaging and control was significant ($CR=4.99$, $p<0.01$) as was the difference between imaging and articulatory suppression ($CR=3.49$, $p<0.01$). Also the difference between articulatory suppression and control was significant ($CR=3.50$, $p<0.01$). The results provides evidence for Holding's (1989) thesis that articulatory loop is also significant in chess players' calculation of variations. The interference in Holding's (1989) experiment perhaps is somewhat stronger than here, which suggests that backwards counting also affects the central executive component of working memory (See Holding 1989 for discussion).

EXPERIMENT 2

The first experiment showed that the imaging suppression task causes clear interference with chess thinking. In contrast, concurrent articulatory suppression results in negligible interference. A subsequent question is whether the interference can be eliminated by changing the visuo-spatial secondary task into a more spatial task than Brooks' (1968) letter task, which requires vivid visual imagery. Baddeley and Lieberman (1980) have clearly shown that a predominantly visual and a predominantly spatial suppression task may lead to qualitatively different types of interference. In chess one might imagine that only the visual component may be really essential, since the pieces, board and the locations of the pieces, i.e. all the relevant components of the game are highly visual. For this reason a covered formboard filler task was selected as the visuo-spatial secondary task for experiment 2. The alternative suppression task was this time a tapping task. Subjects were asked to tap each time they heard the words white (or black) in a long random series of "white" and "black".

Method

Subjects

Six subjects participated in the experiment. The three skilled subjects were strong chess masters, all having a rating of well over 2200 SELO points. The three other subjects had ratings less than 1800 SELO points.

Stimuli and design

Six chess positions in three conditions were presented to the subjects. Two of the positions were presented without any secondary task, two of the positions were given with an auditory secondary task in which subjects heard a series of the words "white", and "black" through headphones at the speed of one word every two seconds. The task was to tap the table with their preferred hand table each time they heard the selected keyword (either "white" or "black"). In the visuo-spatial conditions subjects were asked to place the pieces of a formboard in correct slots. The subject's preferred hand was out of sight in a box.

A formboard was covered by the box and about five to ten centimeters from the formboard in the box there was a cup with the form pieces. Subjects took one piece at a time and searched for a suitable slot. When they found the right slot for the first piece they took the next and so on until they had solved the primary task. Verbal thinking aloud protocols were collected with a tape recorder.

Results and Discussion

The change to a more spatial secondary task did not have a real effect for the results mainly replicate the first experiment (See Figure 2).

Skill differences again were clear. The masters use significantly less time per generated move than the less skilled chess players in all conditions ($F(1,14)=19.88$, $p<0.05$). The second experiment thus provides further confirmatory evidence for the greater fluency of thought with masters compared to less skilled players. The interaction between the type of secondary task and skill was not significant ($F(2,8)<1$).

The overall effect of the secondary task conditions was significant ($F(2,8)=4.70$). Imaging suppression impaired subjects performance most, and the effect was consistent over all the experiments with chess players. A Newman-Keuls test showed that the statistical differences between imaging and control were significant at the 0.01 level ($CR=4.54$). The difference between articulatory suppression and control was however, not significant ($CR=3.24$, $p>0.05$).

The absence of an articulation effect is also interesting. This time the task was relatively "difficult" because it required semantic level processing. Subjects had to discriminate the word "white" from the word "black" to be able to respond. This should be a more difficult task than the routine "the".."the".., because it should involve the central executive. However, in the results there was no sign of an effect obtained by Holding (1989) in experimenting with a backwards counting task, which presumably required far more central executive involvement than the black-white classification task.

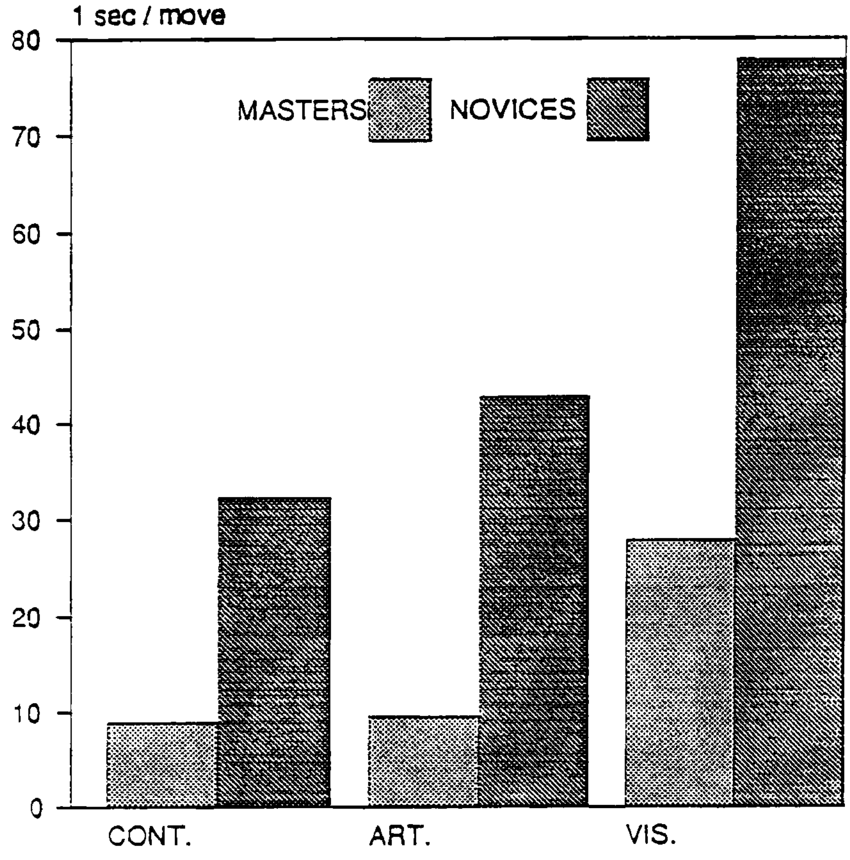

Figure 2. The results of experiment two. The influence of tapping (art.) and form board filling secondary tasks on chess players' move generation compared to control (cont.).

The situation was not, however, this simple. There was an articulation effect in the first experiment. The first experiment and Holding's (1989) experiment both required active production of words. It may be that the articulatory interference in the second experiment reflects the subjects' ability to build a direct pathway from the ears to the fingers, and therefore the use of the articulatory loop is at a minimum (Allport, Antonis & Reynolds, 1972). The articulatory involvement is obviously a very interesting and complicated issue. It is, however, outside the main theme of this paper.

GENERAL DISCUSSION

In both experiments visuo-spatial suppression impaired the fluency of chess players' thinking, while articulatory tasks had a much smaller effect or no effect at all. The impairment is clear in both predominantly visual and predominantly spatial tasks. The results of the two experiments thus confirm that visuo-spatial suppression causes delay in the stream of thought and not only in visual information intake (Bradley et. al, Robbins, personal communication, Saariluoma, in press). Therefore it is possible to argue that visuo-spatial working memory is a very central processing system in chess players' thinking and thus also in apperception.

In competitive games chess players' are not allowed to touch pieces before they make a move. Therefore, they must carry out all planning and calculations in their minds, and it would be quite pointless to argue that chess players' "see" the imagined positions in the same sense as the see positions that are physically present. The present experiments suggest that visuo-spatial working memory is the locus of chess players' advanced calculations. Presumably, the calculations involve processes which are similar to those involved in mental rotation and transformation experiments (Bundesen & Larsen, 1975; Church & Church, 1977; Cooper & Shepard 1973; Milojkovic, 1982). This extensive use of mental imagery in chess is not idiosyncratic, because people use images in reasoning (Huttenlocher 1968), decision making (Beach & Mitchell 1987), in mathematics (Hayes 1973) and mental abacus (Hatano, Miyake & Binks 1977; Hatano & Osawa, 1983; Hishitani, 1989; Stigler, 1984), and in scientific problem solving (Shepard, 1978).

Chess players' active use of their visuo-spatial working memory in searching for a move is an important piece of knowledge in trying to understand apperception. It shows that apperception is not direct perception but is mediated by visuo-spatial working memory. Chess players do not directly see the continuations they work with but they use this low level analogical information storage to manipulate their representation and to study the consequences of the hypotheses they make to find good moves. Since "seeing" an imaginary position cannot be equated with the perception of a physically present position, apperception cannot be reduced to perception.

Characteristic of multifunctional thinking is the fact that the contents of the apperceived spaces cannot be understood by analysing visuo-spatial working memory alone. It is necessary to postulate higher level conceptual control systems in order to explain the contents of the thought. To take a simple analogy: In mental rotation the speed of transformation can be measured, but if the rotation is made to the left or to the right, it is no longer an issue of transformation level. Higher level processes are needed to explain the direction of rotation and analogously the contents of the variations in chess. Possibly, the first step in this direction are the experiments suggesting the systematic central executive interference effects and superior verbal knowledge of skilled chess players (Bradley et al., 1987; Holding, 1989; Pfau & Murphy, 1988; Robbins, personal communication).

REFERENCES

Allport, D., Antonis, B. & Reynolds, P. (1972). On division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology* **24**, 225-235.

Baddeley, A. (1986). *Working memory*. Cambridge: Cambridge University Press.

Baddeley, A. & Hitch, G. (1974). Working memory. *In*: G. Bower (ed.). *The psychology of learning and motivation*, vol.8. New York: Academic Press.

Baddeley, A. & Lieberman, K. (1980). Spatial working memory. *In* R. S. Nickerson (ed.), *Attention and performance VIII*. Hillsdale, N. J.: Erlbaum.

Beach, L. & Mitchell, T.R. (1987). Image theory: Principles, goals, and plans in decision making. *Acta Psychologica* **66**, 201-220.

Bradley, A. Hudson, S.R., Robbins, T., & Baddeley, A. (1987). Working memory in chess. Unpublished report: University of Cambridge, May 1987.

Brooks, L. (1968). Spatial and verbal components in the act of recall. *Canadian Journal of Psychology* **22**, 349-368.

Bundesen, C. & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance* **1**, 214-220.

Charness, N. (1981). Search in chess: Age and skill differences. *Journal of Experimental Psychology: Human Perception and Performance* **7**, 467-476.

Chase, W.G. & Simon, H.A. (1973a). Perception in chess. *Cognitive Psychology* **4**, 55-81.

Chase, W.G. & Simon, H.A. (1973b). The mind's eye in chess. *In* W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press.

Church, R. & Church, K. (1977). Plans, goals and search strategies for the selection of a move in chess. *In* P. Frey (Ed.). *Chess skill in man and machine.* New York: Springer. pp 54-81.

Cooper, L. & Shepard, R. (1973). Chronometric studies of the rotation of mental images. *In* W. Chase (Ed.), *Visual information processing.* New York: Academic press.

Elo, A. (1978). *The ratings of chess players: Past and present.* London: Batsford.

Ericsson, K.A. & Simon, H. (1984). *Protocol analysis.* Cambridge, Mass.: MIT-press.

de Groot, A. (1965). *Thought and choice in chess.* The Hague: Mouton.

de Groot, A.D. (1966). Perception and memory versus thought: Some old ideas and recent findings. *In* B. Kleinmuntz (Ed.), *Problem solving.* New York: Wiley.

Hartstone, W. & Wason, P.C. (1985). *The psychology of chess.* London: Batsford.

Hatano, G., Miyake, Y. & M. Binks, M. (1977). Performance of expert abacus operators. *Cognition* 5, 47-55.

Hatano, G. & Osawa, K. (1983). Digit memory of grand experts in abacus derived mental calculation. *Cognition* **15**, 95-110.

Hishitani, S, 1989. The usefulness of studies of imagery experts in imagery research. *Journal of Mental Imagery* **13**, 119-134.

Holding, D.H. (1985). *The psychology of chess skill.* Hillsdale, N.J.: Erlbaum.

Holding, D.H. (1989). Counting backwards during chess move choice. *Bulletin of the Psychonomic Society* **27**, 421-424.

Holding, D.H. & Pfau, H.D. (1985). Thinking ahead in chess. *American Journal of Psychology* **98**, 271-282.

Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review* **75**, 286-298.

Leibniz, G.W. (1704/1979). *New essays on human understanding.* Cambridge: Cambridge University Press.

Logie, R., Baddeley, A., Mane, A., Donchin, E. & Sheptak, R. (1989). Working memory in the acquisition of cognitive skills. *Acta Psychologica* **71**, 53-87.

Milojkovic, J. (1982). Chess imagery in novice and master. *Journal of Mental Imagery* **6**, 125-144.

Neisser, U. (1987). From direct perception to conceptual structure. *In* U. Neisser (Ed.) *Concepts and conceptual development: Ecological and intellectual factors in categorization.* Cambridge: Cambridge University Press.

Newell, A. & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, N.J.: Prentice-Hall.

Pfau, H.D. & Murphy, M.D. (1988). The role of verbal knowledge in chess. *American Journal of Psychology* **101**, 73-86.

Saariluoma, P. (1984). Coding problem spaces in chess. *Commentationes scientiarum socialium vol 23.* Turku: Societas scientiarum fennica.

Saariluoma, P. (1990). Apperception and restructuring in chess players' problem solving. *In* K. J. Gilhooly, M. Keane, R.H. Logie & G. Erdos (Eds.) *Lines of thought: Reflections on the psychology of thinking,* Vol 2. Chichester: Wiley, pp 41-57.

Saariluoma, P., (in press). Visuo-spatial and articulatory interference in chess players' information intake. *Applied Cognitive Psychology.*

Shepard, R., 1978. Externalization of mental images and the act of creation. *In* B. Randhawa & W. Coffman, (Eds.), *Visual learning, thinking and communication.* New York: Academic Press.

Stigler, J. (1984). "Mental abacus": The effect of abacus training on childrens' mental calculation. *Cognitive Psychology* **16**, 146-176.

Wundt, W. (1920). *Grundriss der Psychologie.* [Outlines of Psychology, in German] Stuttgart: Kröner.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 6

Encoding and maintenance of information in visual working memory

Gerry Quinn
University of St. Andrews, United Kingdom

INTRODUCTION

Over the last fifteen years or so, the Working Memory (WM) model (Baddeley and Hitch, 1974) has proved a profitable way to conceptualise the short term storage of information. The model involves the development of a number of modality-specific systems, often referred to as "slave systems", capable of dealing with modality-specific information and which are controlled by an amodal Central Executive (CE) system. While the precise details of the mechanisms involved in the model are still subject to empirical debate, at least one of the slave systems, the articulatory loop, has been subject to intense investigation. This system is responsible for dealing with verbal information and, after its original formulation in 1974, has been significantly modified to account for a variety of phenomena associated with short term verbal memory (Baddeley, 1983; Baddeley, Lewis & Vallar, 1984).

While the model envisages that a number of other slave systems will be developed and their characteristics investigated, only one other has so far been considered. The visual-spatial sketchpad (VSSP) is assumed to be responsible for the maintenance and manipulation of visuo-spatial information. The theoretical development of this slave system has lagged behind the development of the articulatory loop, however it has increasingly attracted the attention of researchers over the last ten years (e.g. Baddeley & Lieberman, 1980; Logie, 1986; Farmer, Berman & Fletcher, 1986) and is now the subject of several, specific controversies.

After some early debate about the status of the VSSP (Phillips & Christie, 1977) there now seems little doubt that it exists as a functionally separable system within the WM model (Farmer et al., 1986;

Baddeley, 1988). In addition, insofar as controversies capable of empirically tractable solutions are indicative of progress, it seems clear that the characteristics of the VSSP will be increasingly understood.

Some early controversies

An early concern was with the nature of the coding in the VSSP. Baddeley and Lieberman (1980) investigated whether information was held in a visual or a spatial form. Within an interference paradigm, they used the popular Brooks (1967) technique as the primary task. This technique involves the presentation of a series of eight sentences which refer to the placing of digits within the adjacent component squares of a 4x4 mental matrix. With the first sentence designating a given square, the subjects' task is to visualise the digits as a path through the matrix and to use this visual image to recall the sentences verbatim. The processing of the sentences is assumed to rely heavily on the VSSP. A set of sentences would be as follows:

In the starting square put a one
In the next square to the left put a two
In the next square up put a three
In the next square to the right put a four
In the next square to the right put a five
In the next square down put a six
In the next square to the right put a seven
In the next square down put an eight

In the control task, the adjectives up, down left and right are replaced with the non-spatial alternatives good, bad, quick and slow. The alternatives are assumed to involve verbal rather than spatial processing. Baddeley and Lieberman paired the Brooks primary tasks with secondary tasks assumed to involve spatial and visual processing. The secondary tasks were investigated in two separate experiments. The spatial task required subjects to keep a flashlight beam on the bob of a pendulum. Subjects were blindfolded with auditory feedback on their accuracy given by means of a photo-sensitive cell on the pendulum. The visual task required subjects to judge the relative brightness of successive grey slides. The main outcome of the experiments was clear: The spatial task alone disrupted the Brooks spatial sentences while the brightness judgment task had no effect.

While Baddeley and Lieberman concluded that the VSSP was spatial in nature, it was not long before evidence was provided suggesting that a visual secondary task could also cause disruption to a primary task involving the VSSP. Using a technique that has obvious parallels to the "unattended speech" effect in verbal working memory (Salamé and Baddeley, 1982), Logie (1986) showed that merely looking at a series of coloured squares could cause disruption of the recall of a list of words if the words were processed using the imaged-based pegword mnemonic. There was no disruption if the words were remembered under rote learning instructions.

It is currently considered, therefore, that information in the VSSP may be disrupted by visual or spatial information with the likelihood of the disruption from either source dependent on the tasks used; the Brooks task involves a definite spatial component and is susceptible to spatial disruption whereas the pegword mnemonic is more visual and is consequently more susceptible to visual interference. While an either/or result of this sort is in itself unsatisfactory when the goal is the further delineation of the attributes of the VSSP, the investigative techniques are sufficiently sound and the phenomenon of the VSSP sufficiently well established to suggest that a further separation of the VSSP into distinct visual and spatial components may be profitable. Within the WM model itself, however, little research has so far been focussed on this.

The debate on the visual or spatial nature of WM has led some research to look for a more explicit definition of spatial processing. Much of the work on spatial processing, such as that of Baddeley and Lieberman, implicitly defines spatial processing in an operational manner: It is the processing that underlies subjects' ability to track a pendulum while blindfolded. As this involves the movement to relative positions in space Quinn and Ralston (1986) and Quinn (1988) have looked at the involvement of movement in the VSSP. In addition, Smyth and her colleagues (e.g. Smyth and Pendleton, 1989) have looked at the sorts of movement that cause interference in the VSSP. Quinn and Ralston paired performance on a slightly modified version of the Brooks' task with the requirement that subjects move their hand in directions that were either compatible or incompatible with the directions indicated by the Brooks' spatial sentences. Incompatible movement caused interference with recall. Moreover, to minimise the contribution of attention (the CE in the WM formulation) a further condition involved passive movement of subjects' hands where the subjects had to concentrate on the sentences while their hands were moved by the experimenter. Under these circumstances interference was again found suggesting explicitly that movement itself, rather than attention to the movement (Idzikowski, Baddeley, Dimbleby and Park, 1983), is involved in spatial processing.

The foregoing selective review has two purposes: First, to give an overview of some of the research into visual memory within the Working Memory framework; second, to illustrate the techniques that are being developed which will enable further elucidation of the phenomenon. Both the Brooks' task and the technique of passive movement are used in the following empirical report which is concerned with a single and very straightforward problem: Is the information in the VSSP particularly susceptible to interference at encoding compared to maintenance or is the information equally susceptible to interference during encoding and maintenance?

Effects of interference during encoding, retrieval and maintenance of information in the VSSP

Three papers (Idzikowski at al, 1983; Morris, 1987; Quinn, 1988) have researched this problem within the WM framework. All have used some form of movement as the interference task and have used either the Brooks' task (Idzikowski et al and Quinn) or memory for the location of circles (Morris) as the primary task.

In their series of experiments, Idzikowski et al required their subjects to follow a bell shape moving in a sinusoidal pattern on a stationary background screen while being presented with the Brooks sentences. Subjects had to be ready to react by pressing a button whenever the bell shape changed. Other conditions included both the bell and the background being stationary, the background but not the bell moving and both the bell and the background moving. Their results showed that only the conditions which involved eye movements caused a decrement in the recall of the spatial sentences. These conditions had no effect on the control sentences.

A further experiment showed that interference occurred when eye movements were required at either presentation or recall of the spatial sentences. Since these results did not differ from a condition where eye movements were required throughout the experiment, the authors concluded that the locus of the interference was most plausibly with the processes involved in the maintenance of the image. However, as interference with maintenance of the spatial sentences was never exclusively tested, the conclusion may be unsafe.

Morris (1987) and Quinn(1988) have reported experiments similar to each other and to that of Idzikowski et al.. Morris' primary task involved the memory for the random location of five circles on an otherwise blank screen with a capacity for up to 81 circles. The circles were serially presented and only one appeared on the screen at any one time. The interference tasks, presented concurrently with the circles, involved articulatory suppression or spatial movement where subjects had to press the buttons of a 5x5 array of buttons. The array was hidden from view and the subjects were trained to press the buttons in a horizontal-boustrophedal sequence (a pattern akin to the path of oxen ploughing a field).

In a separate experiment the interference tasks occurred during a 10 second delay between presentation and recall.

Morris' results showed that only the spatial movement task caused interference and only caused interference when it was concurrent with the presentation of the circles.

Quinn's experiment investigated two factors: The relative effects on the Brooks' sentences of visual and spatial interference tasks; and the timing of the interference tasks. The following summary refers only to the timing manipulation. Quinn's primary task was a modified Brooks task. The interference tasks again included the spatial movement task used by Morris. Unlike Morris, Quinn ran the delayed interference conditions and the concurrent interference conditions in the same experiment. In addition, the Brooks' non-spatial sentences were included as a control.

Quinn's results were consistent with those of Idzikowski et al and with those of Morris: An interaction of sentence type and timing of interference showed that while the spatial sentences recall actually improved when interference was during the delay between presentation and recall compared with concurrent interference, the performance with the control sentences fell when interference was delayed rather than concurrent.

While these sets of results encouraged both Morris and Quinn to isolate the encoding stage as the locus of the interference, and would allow this interpretation of the results of Idzikowski et al., all the papers contain a confounding factor. The interfering tasks used were effortful and therefore are likely to have involved the CE. Indeed, while Idzikowski et al. acknowledge the confounding and Quinn proffers an explanation implicating the CE, Morris (without the benefit of a control condition) cites the disruption of attention as the most obvious explanation. Clearly, any experiments which have as their goal a more precise delineation of interference effects in the VSSP must attempt to minimise the contribution of attention.

The following experiment sought to clarify this issue by using the technique of passive movement to minimise CE involvement. It tested the hypothesis that interference in the VSSP by a spatial movement task is confined to the active encoding processes rather than to the processes involved in maintenance.

METHOD

The technique used is similar to that reported by Quinn and Ralston (1986). A modified version of the Brooks' task is used where nine rather than eight sentences are used and the subjects are required to recall by writing down the digits in the appropriate formation on blank pages of a booklet. Written recall of the control sentences was also required.

Subjects

Eighteen right-handed undergraduates served in a within subject design

Materials

Each subject was given two blank booklets, one containing six pages to be used during the practice session and one containing eighteen pages to be used during the experimental session. The experimenter held sheets of paper containing either the spatial sentences or the control sentences. There were eighteen sets of nine spatial sentences in addition to six sets of nine practice sets and the same number of control sets although the control sets contained eight sentences each. In common with others who have used the Brooks sentences, pilot studies indicated that performance on the two sentence sets was equalised when the control sets contained a smaller number of sentences. The spatial sentences were in the form of the Brooks sentences illustrated above, as were the controls. All the squares designated were within a 5x5 matrix and there were never more than three squares sequentially designated within one row or column. Within these constraints, the designation of the squares was random. A stop watch was used to time the presentation of the sentences. A fixation point was present throughout the experiment and placed approximately 122cm in front of the subject.

A 5x5 matrix was taped onto the right-hand half of a 122x76cm table. The squares of the matrix were 6.4x6.4cm with the masking tape an additional 2.5cm thick. The matrix was covered by a box, open at one end to allow the experimenter to see the matrix and covered by a flap at the other end to allow the subjects to move their hands over the matrix without being able to see the matrix at the same time.

Design and Procedure

Subjects were run individually. In the spatial condition the fixation point and the table matrix were shown to them. They were told they would hear so many sets of nine sentences of the sort, "In the starting square put a 1, in the next square to the left...." and were shown how the sentences related to the matrix. The starting square was always the second row of the second column. The subjects had to keep their eyes on the fixation point and to attempt to build up an image of the digits-in-position so they could write the digits down in the appropriate formation in the booklet.

In the control condition, the subjects were told they would hear so many sets of eight sentences of the sort, "In the starting square put a 1, in the next square to the quick....". Of course, the sentences were not related to the matrix although the subjects were required to fixate. The instructions were to remember the sentences as a verbal list. To equate performance levels, subjects were told to write down only the adjective and the digit ("start 1, quick 2") in a vertical column; "Start 1" had always to be written down first with the other sentences written in any order although they had to appear in the correct spatial order down the page. The adjective and digit had to be written together as a pair. In both sentence conditions, no time limit was put on the response period.

All subjects served under three delayed passive movement conditions under each set of sentences. The movement conditions were as follows:

1. *No movement* where the subject's right hand was held still on the starting square by the experimenter throughout the presentation of the sentences and the delay.
2. *Concurrent movement* where the subject's hand was moved in time with the sentence presentation, beginning as the spatial/control adjective was mentioned, but not during the delay.
3. *Delayed movement* where the hand was moved only during the delay following sentence presentation.

In all conditions, the box was removed from the matrix to allow subjects to place their hand on the starting square. It was then replaced to allow the experimenter who was seated in three quarters profile at the opposite side of the table to the subjects' right to move the hands and make frequent checks that the subjects maintained fixation. The set of sentences was presented verbally at a rate of one sentence every 4 seconds. Presentation time was therefore 36 seconds for the spatial sentences and 32 seconds for the control sentences. The delays which followed presentation lasted one half the time of the presentation, 18 seconds in the spatial condition and 16 seconds in the control condition.

Recall followed the delay. The 18 test trials were made up of three blocks of 6 trials sets, one block under each of the movement conditions. The three blocks were presented under a Latin-square design to ensure that each block appeared equally often in each serial position. The 18 sets appeared equally often under each movement condition, ensuring that any peculiarities of a particular set were not associated with a particular movement condition. Two practice sets preceded each block. Periodically throughout the experiment, subjects were reminded of the instructions and to fixate.

RESULTS

One point was given for each digit written in the appropriate relative position or each adjective/digit pair correctly written. The scores were converted into percentages to allow direct comparison. They were analysed using a 2 (sentence type) x 3 (movement condition) within subject ANOVA. The movement condition and the interaction between movement and sentence type were significant, though the former was marginally so (respectively, $F(2,34)=3.24$, $(0.05<p<0.1)$; $F(2,34)=4.66$, $p<0.025$. Table 1 illustrates the performance levels.

Table 1. Percent performance under each of the two sentence conditions and the three movement conditions

	Movement condition		
	None	Concurrent	Delayed
Spatial	86	79.7	84.67
Control	83.6	80.5	81.8

Newman Keuls tests showed that while there was a marginal decrease in performance in both sentence conditions under concurrent movement compared to no movement, the decrease was statistically significant only with the spatial sentences ($p<0.05$). There were no other significant effects.

DISCUSSION

The passive movement condition was used to minimise the effect of an attentional component confounding the results. Plausibly, little attention has to be devoted to hand movements when the movement is neither initiated nor monitored by the subjects. This is borne out by the marginal effect the movement had on the control sentences. In terms of Working Memory theory, the control sentences would be handled in the articulatory loop and would not be affected by a spatial movement task unless there was a general effect mediated by the CE. In contrast, the VSSP would be specifically susceptible to the spatial interference. The results show that this interference is limited to the active encoding stage. There was no effect during maintenance of the information.

It is clear that the encoding of relative direction is vulnerable to movement in incompatible directions. This finding replicates that of Quinn and Ralston (1986). It is more difficult to explain why there is less interference once the directions have been encoded. One explanation, consistent with the work of Smyth and Pendleton (1989), is that once formed the directions are remembered as an overall pattern rather than a sequence of discrete relative directions and are therefore not susceptible to interference by discrete movement. If this is the case, it should be possible to design an interference task which would selectively interfere with a static complex pattern and not with spatially defined sequences and so gain a better insight into the form of the representation. An alternative explanation is that maintenance involves a more conceptual long term storage requiring an interplay of both spatial and semantic features. In this case, such a complex representation would be less likely to yield to an interference paradigm although the complexity itself could be shown to have tractable memorial consequences.

REFERENCES

Baddeley, A,D. (1983). Working Memory. *Philosophical Transactions of the Royal Society of London* **B 302**, 311-324.

Baddeley, A. (1988). Imagery and working memory. *In* M. Denis J. Engelkamp and J.T.E. Richardson (Eds.) *Cognitive and Neuropsychological Approaches to Mental Imagery.* pp 169-180. Martinus Nijhoff Publishers

Baddeley, A.D. & Hitch, G.J. (1974). Working Memory. In G. H. Bower (Ed.) *The Psychology of Learning and Motivation, vol 8.* pp 47-90. New York: Academic Press

Baddeley, A.D., Lewis, V.J. & Vallar, G. (1984). Exploring the articulatory loop. *Quarterly Journal of Experimental Psychology* **36**, 233-252.

Baddeley, A.D. & Lieberman, K. (1980). Spatial working memory. In R. Nickerson (Ed.) *Attention and Performance* **8**. pp 521-539. Hillsdale, NJ: Erlbaum.

Brooks, L.R. (1967). The suppression of visualization by reading. *Quarterly Journal of Experimental Psychology* **19**, 289-299.

Farmer, E.W., Berman, J.V.F. & Fletcher, Y.L. (1986). Evidence for a visuo-spatial scratchpad in working memory. *Quarterly Journal of Experimental Psychology* **38A**, 675-688.

Idzikowski, C., Baddeley, A.D., Dimbleby, R. D. & Park, S. (1983). Eye movements and imagery. Paper presented to the *Experimental Psychology Society*, April.

Logie, R.H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology* **38A**, 229-247.

Morris, N. (1987). Exploring the visuo-spatial scratchpad. *Quarterly Journal of Experimental Psychology* **39A**, 409-430.

Phillips, W.A. & Christie, D.F.M. (1977). Interference with visualisation. *Quarterly Journal of Experimental Psychology* **29**, 637-650.

Quinn, G. (1988). Interference effects in the visuo-spatial sketchpad. *In* M. Denis, J. Engelkamp and J.T.E. Richardson (Eds.) *Cognitive and Neuropsychological Approaches to Mental Imagery*. pp 181-189. Martinus Nijhoff Publishers.

Quinn, J.G. & Ralston, G.E. (1986). Movement and attention in visual working memory. *Quarterly Journal of Experimental Psychology* **38A**, 689-702.

Salamé, P. & Baddeley, A.D. (1982). Disruption of short term memory by unattended speech: implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior* **21**, 150-164.

Smyth, M.M. & Pendleton, L.R. (1989). Working memory for movements. *Quarterly Journal of Experimental Psychology* **41A**, 235-250.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 7

Visuo-spatial working memory: Visual, spatial or central executive?

Robert H. Logie,
University of Aberdeen, United Kingdom
and
Clelia Marchetti,
Clinica del Lavoro, Veruno (Novara), Italy

INTRODUCTION

The concept of Working Memory as proposed by Baddeley and Hitch (1974) refers to the temporary storage and manipulation of information. Working memory comprises a central executive, responsible for reasoning, decision making, and coordinating the activities of two subsidiary systems. One of the subsystems, the articulatory loop, is responsible for retention of verbal material, while the other subsystem, the visuo-spatial scratch pad, is responsible for retention of visual and/or spatial information.

The significant theoretical developments in this literature have centred on the articulatory loop (see Baddeley, 1986). The concept of an articulatory loop is now relatively sophisticated (Vallar & Baddeley, 1984; Salamé & Baddeley, 1982). It is thought to comprise a passive phonological store and an active subvocal articulatory rehearsal loop. It has been shown to be useful in counting (Logie & Baddeley, 1987), in arithmetic (Hitch, 1978), in aspects of normal reading (Lewis & Baddeley, 1979), and in the acquisition of language (Gathercole and Baddeley, 1989).

The concept of the visuo-spatial scratch-pad is less well developed both in theory and in application, although this trend is beginning to change (eg. Morris, 1987; Logie, 1989; in press; Quinn, 1988; Quinn, this volume). There is for example evidence that visual confusions occur when subjects attempt to remember visually presented letters or characters that are visually similar to one another (Frick, 1988; Hue & Ericsson, 1988; Wolford & Hollingsworth, 1974). Saariluoma (this volume) presents evidence supporting the usefulness of the visuo-spatial scratch-pad in chess expertise. Certainly, work in this area is sufficiently advanced to have generated a number of debates.

One debate has been whether there is any need for a specialised visuo-spatial system separate from a central executive (eg. Phillips, 1983). This issue will be addressed to a limited extent in this paper, but a more detailed discussion is provided elsewhere (Logie, Zucco & Baddeley, 1990).

A second debate is as follows. If we consider the cognitive functions involved in temporary visuo-spatial storage and processing, do they reflect just one mechanism, or are there separate mechanisms for visual and for spatial material (Baddeley & Lieberman, 1980; Logie, 1986; Quinn & Ralston, 1986)? Baddeley and Lieberman (1980), demonstrated that use of a spatial mnemonic (the method of loci; see e.g. Paivio, 1971), is disrupted by the subject being asked to concurrently track a moving target. Moreover, the tracking task causes disruption even when it is performed with auditory feedback and no visual input. Concurrent tracking has a much less dramatic effect on retention of visual, rather than spatial information. These results led Baddeley and Lieberman to conclude that the visuo-spatial scratch-pad is probably a *spatial* rather than a visual scratch pad. Logie (1986) further demonstrated that use of a visual imagery mnemonic (rhyming peg-word mnemonic; e.g. Paivio, 1971) *is* disrupted by concurrent presentation of irrelevant visual material that the subject is requested to ignore. This pattern of results suggests that there may well be two separate scratch-pads, one for spatial material, the other for visual material. This discussion echoes a similar distinction current in the literature on visual imagery, (e.g. Farah, Hammond, Levine & Calvanio, 1988).

Finally, there is now a growing literature on visuo-spatial storage and processing as studied during encoding and recall of visuo-spatial information (For reviews see Baddeley, 1986; Logie, 1989; 1991). However, there is very little if any literature concerning the retention interval in visuo-spatial memory tasks. Morris, (1987) has shown that when retention of a pattern of different sized circles is coupled with concurrent arm movement performed by the subject, there is mutual disruption of performance, but only if the concurrent movement occurs during the presentation (encoding) of the pattern of circles. If the concurrent arm movement occurs during a retention interval, there is no disruptive effect on retention of the circle patterns.

Quinn (this volume) reports a similar result when concurrent movement is performed in concert with retention of a square matrix pattern (Brooks, 1967). The 'Brooks task' involves asking subjects to imagine a blank square matrix pattern and then to place consecutive numbers in adjacent squares in a path around the matrix. Subjects are later required to recall the position of the numbers in the matrix (see Quinn's chapter for details). Movement during encoding disrupts recall of the positions of the numbers, but movement during a retention interval has little or no effect.

In terms of the visual versus spatial debate outlined above, these results suggest that whatever cognitive systems are involved in encoding are also involved in the control of active movement. They also suggest that the system responsible for retaining the information, once encoded, is most likely not associated with movement control.

In terms of the working memory model, such findings present something of an anomaly, since it seems reasonable to assume that a visuo-spatial scratch-pad would be involved in retention of visuo-spatial information on a temporary basis. Moreover, the tasks used by Quinn and by Morris do appear to require visuo-spatial processing. One possible interpretation is that the *encoding* of such material is the prerogative of the central executive system of working memory, that movement control also requires this system, and that the disruption observed is a general distraction effect, rather than reflecting competition for specialised visuo-spatial resources. This makes intuitive sense since encoding requires attention to both the task and to the material in hand, and the central executive is thought to play a key role in directing and controlling attention.

What then is the role of a visuo-spatial scratch-pad?. If indeed it is involved in retention rather than encoding, how might it best be studied? One possible approach is to assume that there are two separate specialised systems, one spatial and one visual, and both operating more or less independently of a central executive. At first blush, this distinction seems to have some difficulty with the results described above. If there is indeed a system specialised for dealing with spatial material, should it not be involved also in retention of spatial material? That is, should not concurrent movement disrupt both encoding and retention of spatial material if the same system is involved in each case?

An important issue to consider here is that in the tasks used both by Morris and by Quinn, only the initial encoding of the material need be a spatial process. In the Brooks matrix task for example, subjects are instructed to imagine moving through an image of a matrix in response to verbal instructions, thus encoding the positions of particular numbers in particular squares of the matrix. This would qualify as a spatial process by most criteria. Once encoding is complete, however, the subject need retain only a static pattern of numbers in an imaged matrix. They do not have to retain a sequence of imagined movements. This imagined pattern can then be used as a mnemonic for later recall. The recall process may also be spatial, if we assume that subjects 'scan' their visual image in order to report the numbers. Thus a spatial process may be involved during encoding and retrieval of a sequence of verbal instructions that describe spatial positions. A visual temporary memory system could be responsible for retaining the imaged pattern of numbers in specific locations in a matrix. In contrast, if the subject is genuinely required to retain a sequence of movements, rather than a static pattern, they would be more likely to rely on a spatial temporary memory system. The experiment reported here was designed to investigate this possibility.

In the experiment to be described, subjects were required to retain information from one of two kinds of visually presented displays, either colour hues or the sequential order in which a series of squares was presented at different locations on the screen. The colour hue memory task was intended to be a purely visual temporary memory task. The sequential squares task was intended to involve retention of spatial material, possibly with subjects mentally rehearsing the sequence in which the squares appeared. A retention interval between presentation and recognition, was filled either by a concurrent movement task, or by presenting irrelevant pictures. If the disruptive effects reported by Quinn and by Morris were primarily due to general distraction during stimulus encoding, then we would anticipate no disruptive effect of either secondary task during the retention interval. However, if separate systems are involved in retaining spatial and visual material, then we would expect that the movement task would disrupt memory for presentation sequence, while the irrelevant pictures (Logie, 1986) should disrupt retention of the colour hues.

METHOD

Subjects

Sixteen female and eight male subjects took part in this experiment. Their mean age was 25.7 years (sd=6.08). Twelve of the subjects were drawn from among undergraduate and postgraduate students at the

University of Aberdeen. The remainder were volunteers from among the medical staff at a hospital in Veruno, Italy.

Materials and Procedure

Memory for colour shades

The subject sat about 600 mm away from a 210 x 280 mm EGA computer screen on which were displayed four 23 mm squares in a rectangular arrangement with a horizontal separation of 75 mm and a vertical separation of 55 mm (edge to edge). On any one trial, each square was shown in a different shade of a single basic colour. Three basic colours were used, namely blue, green and purple. The squares, appeared one after another at a rate of one per second until all four squares were displayed on the screen. One second after the fourth square was shown, the subject heard a warning tone and the screen was blanked for 10 seconds. After the retention interval, the subjects again heard a warning tone and four squares again appeared on the screen, in the same spatial locations as before. On half of the trials, one of the squares was changed to a different shade of the same basic colour.

The subject's task was to indicate vocally whether or not the shades of the squares were identical to those that they had seen previously.

For each subject there were 40 trials in each of three conditions as follows:

Control: during the retention interval the subject was requested to do nothing other than to try to remember the shades of the presented squares.

Active movement: the subject was asked to place his or her right hand in the second square of the second row of a 5 x 5 matrix taped on a table situated to the right of the computer screen. The squares of the matrix were 6 cm x 6 cm. The matrix was covered by a box, open at one end to allow the experimenter to see the matrix and covered by a flap at the other end to allow the subject to move their right hand over the matrix out of their view. As soon as they heard the warning signal indicating the beginning of the retention interval, the subject was to move their hand a square at a time, from the starting square, three squares to the right (the end of that row), go down to the square below, and move left through the squares until arriving at the left end of the row. They were then to move down to the next row, followed by a movement to the right.

Movement was at a rate of one per second, and this timing was such that the retention interval was completed before the subject moved their hand to the end of the matrix at the bottom right hand corner. Prior to the start of the experimental trials, subjects were given sufficient practice to allow them to make an error free unseen sequence of movements.

Irrelevant pictures: In this condition, as soon as the retention interval commenced, the subject was required to look up at a white screen placed in clear view behind the computer screen, and about 2 metres from the subject. During the retention interval five slides were projected onto the screen depicting black and white line drawings of common objects and animals, chosen at random from Snodgrass and Vanderwart (1980). Presentation of the slides was at a rate of one slide every two seconds. The subjects were requested to keep their eyes open and directed toward the screen but to ignore the presented line drawings and to concentrate on remembering the shades of the squares. At the end of the retention interval, the subjects looked back at the computer display for the visual or spatial recognition test.

The slide projector was used in all three conditions, but with the projector light blanked out. This was to ensure that any effect of general distraction due to the noise of the projector would be present in all three conditions.

The order of the three conditions was counterbalanced across subjects.

Memory for spatial sequences

Subjects taking part in this condition were shown six 23 mm squares of the same colour and hue (light blue). The squares had a horizontal separation of 25 mm and a vertical separation of 15 mm (edge to edge) on the screen. They were presented respectively in six different locations on the computer screen at a rate of one per second, until all six squares were present on the screen. One second after the sixth square was shown, there was a warning tone, and the screen became blank for 10 seconds, following which was another warning tone, and again the six squares appeared in sequence. On half of the trials, the sequence in which the squares was presented was identical to that shown prior to the retention interval. On the remaining trials, the order in which two of the squares appeared was changed. The subject was required to indicate vocally whether or not the two sequences were identical.

As for the colour shade memory group, recognition memory for the sequence of presented squares was tested under three conditions; in a control condition with a blank retention interval, or where the retention interval was filled either with irrelevant pictures or with unseen arm movement. Also as before the subject performed 40 trials under each of the three conditions, which were presented in a counterbalanced order across subjects.

The 24 subjects were divided into two groups, with one group tested for colour memory, and the other group tested for spatial sequence memory. Half of the Aberdeen sample of subjects, and half of the Italian subjects were in the group tested for colour recognition memory. The remaining subjects from both sites were allocated to the spatial sequences group.

RESULTS

The number of false recognitions for the Italian and the Scottish subjects were first analysed to investigate whether there were any national group differences that might be confounded with experimental manipulations. There was no difference in overall levels of performance on the basis of their origin ($F(1,20)=1.84$). Thus for subsequent analyses, the source of the subjects was ignored. Results are shown in Table 1 in terms of the mean number of false recognitions for each memory task, and each distractor task.

Table 1. Mean incorrect recognitions (max=40) for spatial and visual displays, with a retention interval either unfilled (control), filled with active movement, or filled with irrelevant pictures

	Control	Movement	Pictures
Spatial	7.58	12.33	7.17
Visual	8.08	7.92	10.08

Analysis of variance on these data revealed that there was no overall difference in the levels of performance with the visual and the spatial versions of the task, but there was a small overall disruption of performance with the presence of a secondary task ($F(2,44)=4.58; p<0.016$). There was, however, a highly significant interaction, between the type of memory task and the type of secondary task ($F(2,44)=11.79; p<0.001$). As can be seen from Table 1, spatial recognition memory appears to have been disrupted by the secondary movement task, but was unaffected by irrelevant pictures. In contrast, performance on the visual recognition memory task was impaired by irrelevant pictures, but was unaffected by intervening arm movement in the retention interval.

DISCUSSION

This experiment was carried out in an attempt to cast further light on the characteristics of visuo-spatial temporary memory. It appears that when subjects are required to retain for about ten seconds, a sequence of movements on the screen, their ability to do so was impaired by unseen arm movement during the retention interval. Retention of a set of colour hues, was disrupted by interpolated irrelevant pictures. This latter result replicates and extends a finding reported elsewhere (Logie; 1986; Matthews, 1983).

The suggestion was that the cognitive function responsible for temporary retention of visual and spatial material might best be conceptualised as comprising separate systems, one responsible for temporary retention of visual material, the other providing a similar function for spatial material. The data obtained in this experiment are consistent with two such separate cognitive systems being reflected in performance on the spatial and the visual tasks respectively. Also, they are consistent with neurophysiological and neuropsychological evidence. Mishkin, Ungerleider and Macko (1983) have shown that there are distinct cortical pathways involved in dealing with respectively object identity and object location. Farah et al (1988) reported a patient who appeared to have a deficit in the performance of visual imagery tasks, but had spared function for spatial imagery tasks. In a more recent paper Hanley et al (in press) have reported a patient who appears to have the opposite deficit, namely a sparing of visual imagery but with a deficit in spatial processing.

One caveat is that the data from Hanley et al also appear to reveal deficits in other cognitive functions, such as verbal serial recall. Other data which we have collected appears to suggest that spatial *encoding,* such as in the Brooks matrix task, involves a very high central executive involvement (Salway, 1991). Thus it would be interesting to investigate the so-called 'central executive' functions in these patients. This central executive load in encoding supports our earlier interpretation of the data reported by Quinn (this volume) and by Morris (1987). That is, the central executive is involved in encoding the Brooks matrix material in some form of mental image, but is less important for maintenance of the information in that image, or the image itself.

How then is such maintenance accomplished? Logie (1989) suggested that visuo-spatial short-term memory might comprise two functions; a passive visual store, and an active rehearsal process, with the latter function related to the control of movement. The rehearsal process would also be akin to some form of mental scanning of the visual representation. Watkins, Peynircioglu and Brems (1984) have provided complementary evidence for a pictorial rehearsal function, and Johnson (1982) has provided independent support for the relationship between visual imagery and the retention of movements.

The data we have described in this paper are consistent with such an hypothesis. According to this interpretation, retention of the colour hues was primarily the responsibility of a passive visual store. Irrelevant visual input is thought to have direct access to this store, thus causing disruption of its contents during a retention interval. Retention of a series of movements would be accomplished by the rehearsal mechanism. Since the rehearsal mechanism also is involved in the control of movement, a requirement to generate a series of irrelevant movements would disrupt this rehearsal mechanism, leading to poorer recall of the original movement sequence.

This interpretation has some considerable explanatory power, and has a highly seductive symmetry with the verbal temporary memory system, the articulatory loop (e.g. Baddeley, 1986). Also, it generates a number of questions as to the role of the rehearsal function in retaining complex static visual images, as well as retaining spatial and movement information (e.g. Smyth & Pendleton, 1990). In so doing, it provides an extremely useful heuristic for the exploration of the relationships between movement control, visual short-term memory and the human capacity to generate and manipulate visual images.

REFERENCES

Baddeley, A.D., (1986), *Working Memory*, Oxford University Press.

Baddeley, A.D., and Hitch, G.J., (1974), Working Memory. In G. Bower (ed.), *The Psychology of Learning and Motivation, vol. VIII*, 47-90, New York: Academic Press.

Baddeley, A.D. & Lewis, V.J. (1981). Inner active processes in reading: The inner voice, the inner ear and the inner eye. In A.M.

Lesgold & C.A. Perfetti (Eds.). *Interactive processes in reading*, 107-129. Hillsdale NJ: Erlbaum.

Baddeley, A.D., and Lieberman, K., (1980), Spatial Working Memory. In R.S. Nickerson (Ed.), *Attention and Performance VIII*, 521-539, Hillsdale, N.J: Lawrence Erlbaum Associates.

Farah, M.J., Hammond, K.M., Levine, D.N. & Calvanio, R. (1988). Visual and spatial mental imagery: Dissociable systems of representation. *Cognitive Psychology* **20**, 439-462.

Gathercole, S. & Baddeley, A.D. (1989). Development of vocabulary in children and short-term phonological memory. *Journal of Memory and Language* **28**, 200-213.

Hanley, J.R., Young, A.W. & Pearson, N.A. (in press). Impairment of the visuo-spatial sketch pad. *Quarterly Journal of Experimental Psychology.*

Hitch, G.J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology* **10**, 302-323.

Johnson, P. (1982). The functional equivalence of imagery and movement. *Quarterly Journal of Experimental Psychology* **34A**, 349-365.

Logie, R.H., (1986), Visuo-spatial processing in working memory, *The Quarterly Journal of Experimental Psychology*, **38A**, 229-247.

Logie, R.H. (1989). Characteristics of visual short-term memory. *European Journal of Cognitive Psychology* **1**, 275-284.

Logie, R.H. (1991). Visuo-spatial short-term memory: Visual working memory or visual buffer? In C. Cornoldi & M. McDaniel *Imagery and Cognition*, 77-102. New York: Springer-Verlag.

Logie, R.H. & Baddeley, A.D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology* **13**, 310-326.

Logie, R.H., Zucco, G.M. & Baddeley, A.D. (1990). Interference with visual short-term memory. *Acta Psychologica* **75**, 55-74.

Matthews, W.A. (1983). The effects of concurrent secondary tasks on the use of imagery in a free recall task. *Acta Psychologica* **53**, 231-241.

Mishkin, M. Ungerleider, L.G. & Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neuroscience* **6**, 414-417.

Morris, N., (1987), Exploring the visuo-spatial scratch pad, *The Quarterly Journal of Experimental Psychology*, **39A**, 409-430.

Phillips, W.A. (1983). Short-term visual memory. *Philosophical Transactions of the Royal Society of London* **B302**, 295-309.

Quinn G (1988). Interference effects in the visuo-spatial sketchpad. *In* M. Denis, J. Engelkamp & J.T.E. Richardson (Eds.). *Cognitive and*

Neuropsychological Approaches to Mental Imagery, 181-189. Martinus Nijhoff.

Quinn, J.G., and Ralston, G.E., (1986), Movement and attention in visual working memory. *The Quarterly Journal of Experimental Psychology,* **38A**, 689-703.

Salamé, P. & Baddeley, A.D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior* **21**, 150-164.

Salway, A.F.S. (1991). Random generation in the working memory dual-task paradigm. Unpublished PhD thesis, University of Aberdeen.

Smyth, M.M. & Pendleton, L.R. (1990). Space and movement in working memory. *Quarterly Journal of Experimental Psychology* **42A**, 291-304.

Snodgrass, J.G. & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity.

Vallar, G. & Baddeley, A.D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of verbal Learning and Verbal Behavior* **23**, 151-161.

Watkins, M.J., Peynircioglu, Z.F. & Brems, D.J. (1984). Pictorial rehearsal. *Memory and Cognition* **12**, 553-557.

Wolford, G. & Hollingsworth, S. (1974). Evidence that short-term memory is not the limiting factor in the tachistoscopic full-report procedure. *Memory and Cognition* **2**, 796-800.

This Page Intentionally Left Blank

Part 3

Imagery and Verbal Processes

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 8

Imagery and enactment in paired-associate learning

Johannes Engelkamp
Universität des Saarlandes, Germany

INTRODUCTION

It is well known that cued recall (CR) in paired-associate learning (PAL) of nouns is better after imagery than after standard verbal learning instructions, and that CR is particularly good after interactive imagery instructions (Begg, 1973, 1978; Bower, 1972). In contrast to this pattern is that of free recall (FR). Free recall is often not better under imagery than under standard learning instructions, and FR with interactive imagery instructions is not better than under separate imagery instructions. This holds true for PAL as well as for list learning (e.g. Begg, 1973, 1978; Denis, 1975; Hasher, Riebman & Wren, 1976). Until recently these findings were explained within the framework of Paivio's dual coding theory (e.g. Begg 1978, 1983; Paivio, 1986). The main explanation for the findings has been that imagery is a very efficient means of integrating information in holistic images; being provided with part of the image in CR reactivates the whole image.

Recently Marschark (Marschark & Hunt 1989; Marschark & Surian, 1989; Marschark & Cornoldi, 1990) repeatedly has argued that it is not imagery per se that is responsible for the observed findings. According to Marschark the results are explained better by the constructs of distinctive (or item-specific) and of integrative (or relational) encoding, which are more general than imagery. In PAL, relational encoding takes place on two levels. Item-pairs are integrated and inter-pair relations are encoded (Hirshman & Bjork, 1989; McDaniel, Wadill & Einstein, 1988). Cued recall, according to this conceptualization, is essentially dependent on pair integration, while FR is dependent on relational information in general and on item-specific information. According to this position it is decisive that encoding processes are manipulated, not whether encoding is done by imagery. Nevertheless the findings mentioned above show that interactive imagery is one appropriate means of integrating concrete noun pairs, and that imagery does not contribute much to the item-specific information of concrete nouns. It is my goal here to show that enactment with verb pairs, in contrast to imagery, hinders pair integration in PAL instead of facilitating it, and that on the other hand it enhances item-specific information. The latter phenomenon has been demonstrated repeatedly for list learning, where enactment clearly improves FR as opposed to standard encoding and imagery (Engelkamp, Zimmer & Mohr, 1990; Zimmer & Engelkamp, 1989). Therefore, the focus here will be on studying the influence of enactment on pair integration in comparison with that of imagery.

Since the verbs used in these experiments denote actions, two forms of imagery must be distinguished. On the one hand it is possible to imagine somebody else performing an action (other-imagery), on the other hand one can imagine oneself performing an action (self-imagery). This difference is important theoretically. It is assumed that other-imagery is comparable to object-imagery and based on the visual-imaginal system in so far as the imaginal processes refer to external stimulus events. Self-imagery, on the other hand, is more comparable to overt enactment and is based on the motor system in so far as the imaginal processes in this case refer to internal motor processes (Johnson, 1982). Pair-integration should therefore be easier with other-imagery than with self-imagery. In other words, under other-imagery similar effects are expected as under object-imagery, and under self-imagery similar effects as under overt enactment. To demonstrate this is a second goal.

A third goal is to show that the effects of imaginal and motor encoding generalize over word class, that is, that they cannot be reduced to word class. In other words, it is assumed that visual imagery allows pair integration for nouns as well as for verbs, and motor imagery hinders pair integration of nouns and verbs, alike.

A final goal is to explore whether enactment hinders pair integration fundamentally or whether it is only that it is not done spontaneously under enactment but can be done intentionally. These four questions will be dealt with in turn.

To study these questions, the PAL paradigm was used and memory was tested by FR and CR. It is assumed that pair-relational encoding is reflected in CR and item-specific encoding in FR. In all experiments the FR preceded the CR. It might be objected that in CR, subjects are recalling material from their earlier FR, rather than from the originally presented list. There are, however, several experiments which used different lists for FR and CR and which found the same pattern of results as are reported below. In the studies of Denis, Engelkamp and Mohr (in press) and of Marschark and Hunt (1989), for instance, it was observed with different lists for FR and CR that instructions which intended to improve pair-relational encoding led to an increase in CR, but left FR unaffected. Since this is exactly the finding that we observed in our studies in which we tested FR and CR successively using one identical list, it is unlikely that our results were determined by testing CR after FR. Since in PAL there are twice as many items to be recalled in FR as in CR, relative frequencies instead of absolute recall scores were used to make FR and CR comparable (if nothing else is said about the scorings). This procedure is appropriate because it has been repeatedly observed that the first and second elements of pairs are recalled equally often in FR (Denis, Engelkamp & Mohr, in press; Engelkamp, Zimmer & Denis, 1989). Finally, it should be mentioned that in all experiments students attending the University of the Saarland served as subjects.

OTHER-IMAGERY AND ENACTMENT

That enactment, as opposed to imagery, facilitates item-specific encoding of action verbs and hinders pair-relational encoding was shown by Engelkamp (1986). One group of 20 subjects learned 12 verb pairs under the instruction to imagine somebody else performing the actions denoted by the verbs (e.g. to hammer, to nod) and one group of another 20 subjects under the instruction to overtly perform the actions. After one presentation of the list there was a FR, and after a second presentation there was a CR. There were additional aspects of the experiment which are of no interest in this context. The results of this experiment are shown in Table 1.

There was a clear interaction between type of recall and type of encoding. As expected and also confirmed in other experiments (e.g. Engelkamp, Zimmer & Mohr, 1990), FR was better following enactment instructions than following other-imagery instructions. This is interpreted to mean that enactment enhances item-specific encoding of action verbs as compared to other-imagery. Cued recall, in contrast, was worse with enactment instructions than with other-imagery instructions. This is explained by the assumption that enactment hinders pair-relational encoding as compared to other-imagery.

Table 1: Mean proportional free and cued recall (FR, CR) of verbs as a function of type of encoding (other-imagery and enactment) (Exp. 2 from Engelkamp, 1986)

	Free Recall	Cued Recall
Other-imagery	.19	.43
Enactment	.37	.23

OTHER-IMAGERY AND SELF-IMAGERY

Whether self-imagery triggers processes comparable to overt enactment was studied by Engelkamp, Zimmer & Denis (1989). If this holds true an interaction between other-imagery and self-imagery comparable to that between other-imagery and overt enactment is to be expected.

Subjects in the experiments of Engelkamp et al. (1989) learned 10 pairs of action verbs under either other-imagery or self-imagery conditions. In the experiments a CR test followed the FR test with no second presentation of the lists. The combined results of Experiments 2 and 3 of Engelkamp et al. (1989) are presented in Table 2. Sixteen subjects participated in Experiment 2 and 18 in Experiment 3. Both experiments used within subject designs. To control for sequence effects, half the subjects were given a list with other-imagery instructions first, followed by a list with self-imagery instructions. This order of presentation was reversed for the remaining subjects. In these experiments only recall of second elements was scored.

The analysis of variance yielded a significant interaction between type of recall and type of encoding. The same interaction was observed when memory of both pair elements was analysed; it was replicated in a third experiment.

The CR performance was as expected. Pair-relational encoding was more difficult under self-imagery than under other-imagery, underlining the assumption that similar processes take place under overt enactment and self-imagery. Unexpectedly, however, FR did not differ between self-imagery and other-imagery. This finding could mean that it is more difficult to concretize an action and thereby to provide good item-specific information with self-imagery than with overt enactment. Further studies are required to clarify this point.

Table 2: Mean proportion of correct recall of the second pair-elements in free and cued recall depending on encoding condition (Exp. 2 and 3 from Engelkamp, Zimmer & Denis, 1989)

	Free Recall	Cued Recall
Other-imagery	.46	.46
Self-imagery	.49	.37

OTHER-IMAGERY AND SELF-IMAGERY WITH NOUNS AND VERBS

The study of Denis, Engelkamp & Mohr (in press) primarily was planned to explore whether the findings concerning other-imagery and self-imagery generalize to nouns. At the same time, however, it contains a replication of the experiments of Engelkamp et al. (1989). If the effects observed by Engelkamp et al. (1989) are due to the encoding processes and not to the word class, they should also occur when actions are applied to objects denoted by concrete nouns.

Denis et al. (1990) therefore presented their subjects in Experiment 2 with lists of 14 noun pairs and 14 verb pairs. Both lists were followed by FR and without second presentation, by CR. Instructions with the verb lists were the same as in the experiments of Engelkamp et al. (1989). With the noun lists subjects were instructed to imagine themselves or somebody else performing a typical action with the objects denoted by the nouns. Sixteen subjects learned the lists under self-imagery instructions, and 15 under other-imagery instructions. The results of this experiment are given in Table 3.

For verbs the same interaction as by Engelkamp et al. (1989) was observed. CR was worse after self-imagery than after other-imagery. This finding supports again the assumption that self-imagery triggers processes like overt enactment. Also the finding that FR does not differ under the two encoding instructions was replicated. This supports the above-mentioned explanation. Although self-imagery, like overt enactment, hinders pair-integration, it does not provide item-specific information to the same degree as does overt enactment. This could mean that the decision to act instead of the decision to simulate an action is critical for generating the good item-specific information associated with overt enactment. In this context we might be reminded of the finding that mental practice is less efficient than physical practice, but better than performance under control conditions (Mendoza & Wichman, 1978). For nouns there was also an interaction between type of recall and type of encoding, but here only the difference of CR between other-imagery and self-imagery was significant. The difference between FR and CR under self-imagery did not yield the usual level of significance in this analysis. Thus, although there are hints that type of encoding influences memory for nouns and verbs in a comparable way, the findings are not very clear-cut.

Table 3: Mean proportional free and cued recall depending on word class and encoding conditions (Exp. 2 from Denis, Engelkamp & Mohr, in press)

	Nouns		Verbs	
	Free Recall	Cued Recall	Free Recall	Cued Recall
Other-imagery	.54	.62	.38	.34
Self-imagery	.53	.46	.34	.20

ENACTMENT AND EXPLICIT INSTRUCTIONS TO INTEGRATE PAIRS

Engelkamp, Mohr & Zimmer (in press) ran a further experiment in which verbs and nouns were compared. In this study, nouns and verbs were to be learned under overt enactment instructions, so that the most extreme and clear case could be examined. The main goal of this experiment was, however, to study the question of whether pair-relational encoding can be enhanced by overt enactment when it is explicitly required.

Subjects in this experiment learned lists of 14 noun pairs and 14 verb pairs. Four groups of subjects received different encoding instructions. Group 1 (20 subjects) was simply instructed to overtly enact typical actions with nouns and the denoted actions with verbs. Group 2 (19 subjects) was additionally instructed to think of a goal for each action. Group 3 (15 subjects) was instructed to think of a common goal for both actions of each pair and Group 4 (16 subjects) to rate how easy it was to find a common goal. In Table 4 the data for Groups 1 and 2 and for Groups 3 and 4 are collapsed because they did not differ. The two resulting groups, which we called "separate" and "common", represent conditions which reflect the spontaneous tendency to integrate pairs under enactment (separate) and the intentional integration of pairs (common). Again CR is tested after FR without further list presentation. The results for these two groups are given in Table 4.

Table 4: Mean proportional free and cued recall depending on word class and encoding conditions (after Engelkamp, Mohr & Zimmer, in press)

	Nouns		Verbs	
	Free Recall	Cued Recall	Free Recall	Cued Recall
Separate	.55	.29	.50	.18
Common	.57	.55	.50	.33

There was an interaction between encoding condition and type of recall for nouns showing that CR was worse under separate encoding than under common encoding and that there was no such difference for FR. Furthermore, performance in CR and FR was equally good in the common encoding group where subjects were explicitly required to integrate pairs.

The factors "encoding condition" and "type of recall" also interacted with verbs. Again CR was worse under separate encoding than under common encoding and again FR did not differ between these encoding conditions. Deviating from the findings for nouns, however, CR was worse than FR when subjects were instructed explicitly to integrate pairs of verbs.

Thus there are four important findings in this experiment. First, enactment hinders spontaneous pair-relational encoding for both nouns and verbs. Second, pair-relational encoding can be induced by an explicit instruction, for both nouns and verbs. Third, pair-relational encoding is generally more difficult for verbs than it is for nouns. Fourth, free recall, that is item-specific encoding, is independent of encoding instruction and comparable for verbs and nouns.

The latter effect means that enactment provides such good item-specific information that further encoding operations do not enhance it, and that it compensates for the usually observed noun-superiority effect (Engelkamp, Zimmer & Mohr, 1990). That enactment under "separate" conditions produces worse CR than FR confirms once more that it hampers spontaneous pair-relational encoding. That an explicit integration instruction under "common" conditions enhances CR shows, however, that enactment does not hinder pair-relational encoding fundamentally. That both effects (worse CR than FR under "separate" conditions and better CR under "common" than under "separate" conditions) hold true for nouns as well as for verbs demonstrates that encoding processes and not word class are decisive for these effects. However, this statement has to be qualified. Although enactment provides good item-specific information and hinders spontaneous pair-relational encoding, but allows nevertheless intentional pair integration, the pair-relational encoding processes are generally easier with nouns than with verbs and the item-specific encoding is needed less for nouns than it is for verbs.

CONCLUSION

Taken together, the results reported here confirm the position of Marschark (e.g. Marschark & Hunt, 1989). They provide evidence that the distinction between item-specific and relational encoding is appropriately applied to PAL if one distinguishes between inter-pair and intra-pair relational encoding, and they demonstrate again that different orienting tasks are appropriate for inducing these different encoding processes. The results reported here, however, also make it clear that modality-specific encoding processes such as imagery and enactment impose constraints on the use of these encoding processes. Visual imagery seems to be particularly suited to integrate pairs, whereas with motor-performance, pair-integration seems particularly difficult and item-specific encoding is particularly effective. Moreover, these constraints are reinforced by word class. The latter finding might mean that actions are inherently more closely connected with verbs while outer events are inherently more closely connected with concrete nouns. In other words, not all encoding processes are equally compatible with all modalities. It is worth exploring further the limits which imagery and enactment impose.

REFERENCES

Begg, I. (1973). Imagery and integration in the recall of words. *Canadian Journal of Psychology* **27**, 159-167.

Begg, I. (1978). Imagery and organization in memory: Instructional effects. *Memory & Cognition* **6**, 174-183.

Begg, I. (1983). Imagery instructions and the organization of memory. *In* J.C. Yuille (Ed.), *Imagery, memory and cognition*. Hillsdale, N. J.: Lawrence Erlbaum.

Bower, G.H. (1972). Mental Imagery and associative learning. *In* L.W. Gregg (Ed.), *Cognition in learning and memory*. New York: Wiley.

Denis, M. (1975). *Représentation imagée et activité de mémorisation*. Paris: Editions du C.N.R.S..

Denis, M., Engelkamp, J. & Mohr, G. (in press). Memory for imagined actions: Imaging oneself versus other person. *Psychological Research*.

Engelkamp, J. (1986). Nouns and verbs in paired-associate learning: Instructional effects. *Psychological Research* **48**, 153-159.

Engelkamp, J., Mohr, G. & Zimmer, H.D. (in press). Pair-relational encoding of performed nouns and verbs. *Psychological Research*.

Engelkamp, J., Zimmer, H.D. & Denis, M. (1989). Paired associate learning of action verbs with visual- or motor-imaginal encoding instructions. *Psychological Research* **50**, 257-263.

Engelkamp, J., Zimmer, H.D. & Mohr, G. (1990). Differential memory effects of concrete nouns and action verbs. *Zeitschrift ffr Psychologie* **198**, 189-216.

Hasher, R., Riebman, B. & Wren, F. (1976). Imagery and the retention of free-recall learning. *Journal of Experimental Psychology: Human Learning and Memory* **2**, 172-181.

Hirshman, E. & Bjork, R. A. (1988). The generation effect: Support for a two-factor theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**, 484-494.

Johnson, P. (1982). The functional equivalence of imagery and movement. *Quarterly Journal of Experimental Psychology* **34A**, 349-365.

Marschark, M. & Cornoldi, C. (1990). Imagery and verbal memory. *In* C. Cornoldi & M.A. McDaniel (Eds.), *Imagery and cognition*. New

York: Springer, 133-182.

Marschark, M. & Hunt, R.R. (1989). A reexamination of the role of imagery in learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15**, 710-720.

Marschark, M. & Surian, L. (1989). Why does imagery improve memory? *European Journal of Cognitive Psychology* **3**, 251-263.

McDaniel, M.A., Waddill, P.J. & Einstein, G. O. (1988). A contextual account of the generation effect: A three-factor theory. *Journal of Memory & Language* **27**, 521-536.

Mendoza, D. & Wichman, H. (1978). Inner darts: Effects of mental practice on performance of dart throwing. *Perceptual and Motor Skills* **47**, 1195-1199.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Zimmer, H.D. & Engelkamp, J. (1989). Does motor encoding enhance relational information? *Psychological Research* **51**, 158-167.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 9

Size comparison with verbal and pictorial material*

Gilbert Mohr and Johannes Engelkamp
Universität des Saarlandes, Germany

INTRODUCTION

Storing and retrieving size information is a cognitive ability used often, at least implicitly, in daily life. Searching for a 5-inch female screw within a workshop, for instance, would be a time-consuming affair, if the only category used to decide were "female screw". Relatively precise knowledge about the wanted screw's size can save a lot of time otherwise spent on the inspection of possibly dozens of very large or very small screws (Cave & Kosslyn, 1989; Larsen & Bundesen, 1978).

Memory for size information is primarily studied within two paradigms. In the first paradigm, episodic memory for transitive orders is considered. Subjects are provided with verbal information in the following form:

D is larger than C. C is larger than B, and B is larger than A.

After having encoded such information, which is usually embedded in meaningful contexts (Potts, 1972), subjects have to verify sentences like the following:

D is larger than A. (Yes)
B is larger than D. (No)
D is larger than B. (Yes)

*We thank Michel Denis and Hubert Zimmer for thoughtful discussions during the preparation of the experiment and the paper. The research was supported by a PROCOPE grant provided by the German Office for Academic Exchange (DAAD). Correspondence concerning this article should be addressed to Gilbert Mohr, FR-Psychologie, Universität des Saarlandes, FRG-6600 Saarbrücken

Verification latencies are most often used as dependent variable in this paradigm.

In the second paradigm, long-term, semantic memory for pre-experimentally acquired size information is investigated. Without any study phase, subjects are directly tested mostly by a binary choice task, where they have to indicate which of two alternatives is the larger one (Banks & Flora, 1977; Cech, 1989; Paivio, 1975; Shoben, Cech, Schwanenflugel & Sailor, 1989). Also in this paradigm, reaction time is the dependent variable usually used.

In episodic memory for serial orders, three results are consistently reported: the distance effect (Griggs & Shea, 1977; Henderson & Well, 1985; Pohl, 1990; Potts, 1972, 1974; Woocher, Glass, & Holyoak, 1978), the end-term anchoring effect (Potts, 1972, 1974; Trabasso & Riley, 1975) and no effect of explicitness (Potts, 1972, 1974). "Distance effect" refers to the finding that reaction times decline with increasing distance between the two elements of a test pair (e.g. rt(BA)>rt(DA) for the order ABCD). "End-term anchoring effect" means that judgements on pairs where an end-term of an order is involved are faster than judgements on inner pairs (like BC in the order ABCD). "Explicitness" labels an often used manipulation of the study material: either only the adjacent pairs of an order are given (AB,BC,CD) or all possible pairs, adjacent and remote (AB,BC,CD,AC,BD,AD), are presented. It was possible to show that the explicit presentation of the remote pairs did not lead to latencies different from those produced by a condition where the remote pairs were not shown but had to be inferred (Potts, 1972, 1974). This shows that subjects build a coherent order during encoding, even if some relations are not given explicitly.

The distance effect is also observed in the semantic, long-term memory paradigm. Here, however, a further finding is reported: the picture superiority effect (Paivio, 1975, Banks & Flora, 1977). Judgements on the basis of pictorial input are considerably faster than judgements on the basis of verbal input.

In both paradigms two major classes of explanations are discussed, claiming different representational formats as the basis of the findings. Advocates of the first class argue in favour of an abstract amodal conceptual representation, whereas advocates of the second class prefer an explanation in terms of a visuo-spatial, analogue representation.

The distance effect as well as the picture superiority effect are taken as pieces of evidence supporting the assumption of a visuo-spatial representation as being accessed in tasks in which physical attributes of objects are to be compared (c.f Paivio, 1986, p.180 ff.). The picture superiority effect is, however, well documented only in the semantic, long-term memory paradigm. Thus the question arises whether a picture superiority effect can be demonstrated in the case of the episodic memory paradigm too, suggesting that the representations and processes, which are used in both tasks are similar or even partially identical.

Richardson (1987) tackled this question, but failed to demonstrate a conclusively interpretable picture superiority effect within the episodic memory paradigm. One possible reason for his results could be that he did not present the relevant information in an analogue, visuo-spatial format. The size information had to be encoded always from verbal input. The pictorial information in his case consisted of portraits, photographs, which had to be associated with the verbal information in a pre-experimental study phase, which possibly could cause additional retrieval problems.

The present study differs from that of Richardson (1987) in two respects: first, the relevant size information is given directly by the pictorial input; second, there is no pre-experimental study phase in which picture-name associations have to be encoded. The study should answer three questions. First, is there a picture superiority effect if the size information is given pictorially? Second, are the distance effect and the end-term anchoring effect identical for verbal and pictorial input? Third, is there an explicitness effect for pictures, but not for verbal input?

METHOD

Subjects

38 students of the University of the Saarland participated in the experiment. They were paid for their participation. 20 subjects learned with pictorial input. Verbal material was given to 18 subjects.

Materials

Nine four-term series of natural objects were chosen as stimuli. The modal natural length of the four stimuli within one set was about equal as in the following examples: safety pin, screw, nail, key / pipe, cigarette, candle, torch. The actual lengths for the experimental stimuli were determined arbitrarily around the modal length with a fixed relative difference between two adjacent elements. The longer element was always 10% longer than the shorter one, which is well above the "just noticeable difference" in the pictorial condition. The nine four-term sets covered a range from 4.0 to 62.5 cm. With one exception the lengths within different sets did not overlap.

Design and Procedure

A three-factorial design with the between-subjects factor "modality" (verbal vs. pictorial input) and the within-subjects factors "presentation form" (remote pairs implicit vs. remote pairs explicit vs. remote pairs only) and "test-pairs" was realized (2*(3*6)-design).

In the study phase, the elements of a set were presented by pairs to the subjects at a rate of 8 sec/pair. Dependent on the presentation form, either the 3 adjacent pairs (AB,BC,CD) or all 6 pairs (AB,BC,CD,AC,BD,AD) or only the remote pairs (AC,BD,AD) were given. Under the pictorial condition, the elements of a pair were given side by side within a reference frame. They stood vertically on the basis of the frame so that the larger one could easily be identified. Under the verbal condition, sentence pairs were offered at the same rate ("The nail has a length of 4.0 cm." "The screw has a length of 4.4 cm"). The sentences were typed one below the other. After the study phase, there was a fixed retention interval of 10 sec. The following test-phase was announced by an acoustic signal 500 ms before the first test item appeared.

For each of the nine experimental sets, 12 test pairs were given. Under the pictorial condition, pairs of same-sized objects (15 cm) were presented with the layout analogue to the study phase. Under the verbal condition, the names of the objects were presented, with one on the left, the other on the right side of the display. All six possible pairings were given twice: once with the larger element on the left and once with the larger element on the right. 27 different random orders of the test-pairs were used.

For each test-pair, subjects had 3000 ms to indicate, by pushing the respective button, on which side the larger object was depicted or denoted. After the response or after the maximum of 3000 ms a blank delay of 1500 ms started, which was interrupted by a short acoustic signal 500 ms before the next test-pair appeared. After having finished a test sequence, subjects could take a short rest before they started the next study-test block.

Reaction time was measured from the onset of a test-pair until the subject pressed a key. Each subject completed 108 trials, 36 under each presentation form.

Before starting the actual experimental session, subjects could work through a practice item several times, until they felt familiar with the procedure.

RESULTS

Proportion of errors

First, error rates for the different presentation forms, the two modalities and the six test pairs are reported. Table 1 provides the percentages of errors for the full design.

Table 1. Percentages of errors as a function of modality, presentation form and test pair

	AB	BC	CD	AC	BD	AD	Mean
pictorial: mean:18							
adjacent only	14	20	23	18	13	8	16
adjacent+remote	14	18	24	23	9	8	16
remote only	28	43	26	22	13	6	23
verbal: mean:13							
adjacent only	10	17	11	8	6	6	10
adjacent+remote	14	13	8	8	4	2	8
remote only	19	48	14	20	9	7	20
Overall Mean	17	27	18	17	9	6	16

The overall error rate was in the same range as that reported by Potts (1972). A 3-way ANOVA over subjects as cases (N=38) with the factors "modality", "presentation form" and "test pair" revealed the following results. First, there was a main effect of the modality variation. The error rate was significantly lower for the verbal condition than for the pictorial condition ($F(1,36)=4.46$ $p<.04$). Second, the factor "test pairs" produced a significant main effect ($F(5,180)=18.88$; $p<.001$;

$Diff_{(crit/p=.05)}=8$). The effect was mainly due to a high error rate for the inner pair (BC) and a rather low error rate for the pure end-term pair (AD). Finally, there was a highly significant interaction between the factors "test pairs" and "presentation form", which qualified the test pair main effect ($F(10,360)=5.85$; $p<.001$; $Diff_{(crit/p=.05)}=15$). An inspection of Table 1 shows that the interaction is primarily caused by very high error rates for the inner test pair (BC) under the presentation form "remote only". Subjects were at chance level for this test pair, when only the remote pairs were presented during encoding. The chance level performance for the inner pair suggests that subjects who were presented with the remote pairs only, could not set up a coherent representation of the whole four-term order. For this reason the presentation form "remote only" was excluded from the analyses of the decision latencies.

For each test pair under both remaining presentation forms independent t-tests were conducted to contrast the error rates of the pictorial group with those of the verbal group. This allowed an exploration of possible speed accuracy trade-offs of modality dependent reaction time differences. Table 2 shows the respective t-values and the corresponding probability scores.

Table 2. T-values and the respective alpha error probabilities of independent t-tests contrasting separately the pictorial and the verbal group for each test pair and for the presentation forms "adjacent only" and "adjacent+remote"

		AB	BC	CD	AC	BD	AD
adjacent only							
	t-value:	0.60	0.61	2.30	1.54	1.22	0.44
	p:	0.55	0.55	0.03*	0.12	0.23	0.66
adjacent+remote							
	t-value:	0.04	0.89	3.13	2.39	1.74	1.21
	p:	0.96	0.38	0.01**	0.02*	0.09	0.23

Note. df=36.

Although there is a general tendency, to produce more errors under the pictorial condition than under the verbal condition, this tendency is far from being reliable for the pairs AB, BC and AD. For these pairs the probability of an alpha error is above .20.

Decision latencies

The analyses of the latencies were carried out considering each item set as a case. Each subject provided 6 cases, 3 for the presentation form "adjacent pairs" and 3 for the presentation form "adjacent and remote pairs" resulting in a total of 228 cases. A 2*2*(6) ANOVA with the factors "modality", "presentation form" and "test pair" was conducted. Table 3 gives the mean reaction times.

Table 3. Mean reaction times dependent on "modality", "presentation form" and "test pair"

test pair:	AB	BC	CD	AC	BD	AD
adjacent only:						
pictorial	1246	1286	1053	1169	1015	998
verbal	1707	1819	1467	1612	1441	1311
adjacent + remote:						
pictorial	1293	1326	1059	1147	1023	1012
verbal	1653	1851	1468	1623	1399	1343

The analysis yielded a highly significant effect of modality ($F(1,224)=172.17$; $p<.001$). Learning and testing with pictorial input (1136 ms) led to considerably faster reaction times than learning and testing with verbal input (1558 ms). There was no effect of presentation form ($F(1,224)<1$; adjacent only: 1344 ms/ adjacent+remote: 1350 ms). The factor test pairs was highly significant ($F(5,1120)=81.29$; $p<.001$;

$Diff_{(crit/p=.05)}=83$). Table 4 shows the mean decision latencies for the six test pairs.

Table 4. Overall mean reaction times for the 6 test pairs

test pair:	AB	BC	CD	AC	BD	AD
	1475	1571	1262	1388	1220	1166

Inspections of the mean reaction time differences show that reactions to the inner pair (BC) are reliably slower than reactions to pairs where end-terms are included. The end-term D (1241 ms), which is the largest element, produces faster reaction times than the end-term A (1431 ms). Finally, there is a distance effect, if one concentrates on the pairs AB (1475 ms), AC (1388) and AD (1166). This distance effect is not so clear cut starting from the largest end (DC (1262 ms), DB (1220 ms) and DA (1166 ms)).

There is one significant interaction, the interaction between the factors "modality" and "test pairs" ($F(5,1120)=3.86$; $p<.01$). Table 5 gives the mean reaction times for the factor level combinations of this interaction.

Table 5. Mean reaction times as a function of "modality" and "test pair"

test pair:	AB	BC	CD	AC	BD	AD
pictorial	1269	1306	1056	1158	1019	1005
verbal	1680	1835	1468	1617	1420	1327

To explore the interaction further, a-posteriori tests were conducted. First, the pair which contained the two inner elements (BC) and the pair which contained the 2 end-terms (AD) were excluded. A 2-way ANOVA with the factors "modality" (2) and "test pair" (4) showed that without the two excluded pairs the interaction disappeared ($F(3,672)<1$). Second, considering the two excluded pairs alone in a 2*2 ANOVA (modality*test pairs (BC vs. AD)) corroborated the conjecture that the interaction was caused by the inner pair (BC) and the end-term pair (AD) ($F(1,224)=14.92$; $p<.001$). As can be seen in Table 5, the picture superiority effect is much more distinct for the inner pair (529 ms) than for the end-term pair (322 ms).

There are some further hints which suggest that having both end-terms is more important for the verbal condition than for the pictorial condition. Contrasting the end-term pair (AD) with the two pairs which contained only the D end-term (CD,BD), produced an effect under the verbal condition ($F(1,226)=15.50$; $p<.001$), but no effect under the pictorial condition ($F(1,226)=1.37$; $p>.20$). The conjecture that the A end-term is much more important for the verbal group is further confirmed by a contrast of the inner pair (BC) with the AB pair (which contains the smallest element). Under the pictorial condition there is no reliable difference between these two pairs ($F(1,226)<1$), whereas under the verbal condition this contrast is highly significant ($F(1,226)=14.77$; $p<.001$).

DISCUSSION

The picture superiority effect is taken as one piece of evidence which suggests that probably a modality-specific representation is consulted when subjects perform a size comparison task (Paivio, 1986). Richardson (1987) could not find a picture superiority effect and concluded from this and some other clues that no analogue representations are involved in the episodic memory variant of the task. We hypothesised that Richardson's finding could be due to the fact that his pictorial material did not contain the critical size information. We expected that a substantial picture superiority effect would come out, if pictorial material were used, which has the critical information within itself. This expectation was confirmed by our experimental data.

The overall higher error rate for the pictorial group might suggest that the picture superiority effect is simply due to the fact that the verbal group took more care to avoid errors than the pictorial group, resulting in longer reaction times. Such a speed-accuracy trade-off interpretation is, however, implausible if one compares the error rates and the latencies of both groups separately for different test pairs. First, one can concentrate on those test pairs for which the picture superiority effect was additive to the test pair effect (i.e. the pairs AB,AC,CD,BD). In this group there is at least one pair, the AB pair, for which it is very unlikely that the error rates for the two input modalities are different (see Table 1 and 2). Notwithstanding, the picture superiority effect for this pair was nearly identical to the effect for the CD pair, for which it is very likely that the two groups have different error rates.

Second, one can focus on the inner pair (BC) for which the error rates of both groups were nearly equal and high, and the reactions times were slow. For this pair the picture superiority effect was most distinct, a finding which is also not consistent with a speed-accuracy trade-off interpretation, because one should expect a positive correlation between the difference in error rates and the picture superiority effect.

Finally, the AD pair shows that there is a considerable picture superiority effect with fast decision latencies and overall low and equal error rates. This also weakens a trade-off hypothesis. Thus, we conclude that we observed a genuine effect of our experimental manipulation.

Although the finding of a picture superiority effect is not conclusive as regards the assumption of an analogue representational format, it is nevertheless an important phenomenon which is to be explained either in terms of different representational formats or in terms of different processes (Broadbent, 1984), e.g. having faster access to a uniform amodal representation (Banks & Flora, 1977). A first step towards more conclusive arguments in favour of representational or procedural differences would be to examine whether the picture superiority effect is an encoding phenomenon, a retrieval phenomenon, or both.

A further question related to the trade-off considerations above will be the question of the quality of the constructed representation. Designing the experiment, we expected that the error rates would be low and would not depend on the input modality. These implicit expectations proved to be wrong. The error rates were substantial and not independent of the stimulus modality. This result suggests that either the representations built from pictures are at least partly less adequate than the representations built from verbal material or that the constructions from pictures are more fragile than the constructions from words. Possibly representations from pictures are more susceptible to interference than representations from words.

In any case, there must be efforts to reduce the error rates on the one hand to get unequivocal reaction time results. On the other hand there must be systematic attempts to explain the obtained error rate differences. Using different secondary tasks within a dual task paradigm might be one way to shed some light on the question of adequacy and robustness of the representations set up from different input modalities.

Using "Stroop-like paradigms" (Fuchs, Goschke & Gude, 1989; Henik & Tzelgov, 1982) with pictures and words may be another one.

A third point which needs some further investigation is the interaction of "modality" and "test pairs". Our finding suggests that especially the role of the A end-term, i.e. the smallest element, is different under the two modalities. Reaction times under the pictorial condition seem to be a function of the D end-term (the largest element) and the distance only, whereas reaction times under the verbal condition seem to depend on both end-terms and on distance. To disentangle distance effects and end-term effects, longer series with more than one inner pair should be studied under both modalities (Potts, 1974).

Finally it should be noted that the implicit/explicit variation ("presentation form") did not interact with the modality factor. One could have expected that the pictorial condition, but not the verbal condition (Potts, 1972, 1974) profits from explicit presentations of all pairings, because a direct perceptual like storage of all six pictures takes place under the pictorial condition. Such an assumption seems implausible against the background of our findings. Moreover, the results suggest that under both conditions a coherent order is constructed from explicit as well as from implicit information. Whether the mental representation of this coherent order is of a certain modality under certain conditions remains a question for further research.

REFERENCES

Banks, W.P., & Flora, J. (1977). Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology: Human Perception and Performance* **3**, 278-290.

Broadbent, D.E. (1984). The maltese cross: A new simplistic model for memory. *The Behavioral and Brain Sciences* **7**, 55-94.

Cave, K.R., & Kosslyn, S.M. (1989). Varieties of size specific visual selection. *Journal of Experimental Psychology: General* **118**, 148-164.

Cech, C.G. (1989). Congruity and the expectancy hypothesis. *Journal of Experimental Psychology: Learning, Memory & Cognition* **15**, 1129-1133.

Fuchs, A., Goschke, T., Gude, D. (1988). On the role of imagery in linear syllogistic reasoning. *Psychological Research* **50**, 43-49.

Griggs, R.A., & Shea, S.L. (1977). Integrating verbal quantitative information in linear orderings. *Memory & Cognition* **5**, 287-291.

Henderson, J.M., & Well, A.D. (1985). Symbolic comparisons with and without perceptual referents: Is interval information used? *Memory & Cognition* **13**, 176-182.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition* **10**, 389-395.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance* **4**, 1-20.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition* **3**, 635-647.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Pohl, R.F. (1990). Position effects in chunked linear orders. *Psychological Research* **52**, 68-75.

Potts, G.R. (1972). Information processing strategies used in the encoding of linear orderings. *Journal of Verbal Learning and Verbal Behavior* **11**, 727-740.

Potts, G.R. (1974). Storing and retrieving information about ordered relationships. *Journal of Experimental Psychology* **103**, 431-439.

Richardson J.T.E. (1987). The role of mental imagery in models of transitive inference. *British Journal of Psychology* **78**, 189-203.

Shoben, E.D., Cech, C.G., Schwanenflugel, P.J., & Sailor, K.M. (1989). Serial position effects in comparative judgements. *Journal of Experimental Psychology: Human Perception and Performance* **15**, 273-286.

Trabasso, T., & Riley, C.A. (1975). On the construction and use of representations involving linear order. *In* R.L. Solso (ed.), *Information processing and cognition: The Loyola Symposium.* Hillsdale, N.J.: Erlbaum.

Woocher, F.D., Glass, A.L., & Holyoak, K.J. (1978). Positional discriminability in linear orderings. *Memory & Cognition* **6**, 165-173.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)
Elsevier Science Publishers B.V., 1991

Chapter 10

Description of perceived or imagined spatial networks*

Frédérique Robin and Michel Denis
Université de Paris-Sud, France

Over the last few years, several research programs have focused on the processes involved in the description of objects or scenes. They centre on the ways in which speakers order information in discourse when they describe places or routes so that another person can construct a representation of these objects (cf. Ehrich & Koster, 1983 Isaacs & Clark, 1987; Linde & Labov, 1975; Shanon, 1984; Ullmer-Ehrich, 1982).

One of the key features of language is sequentiality. There is no way for speakers to overcome the constraint of making statements one after the other. The linearization of discourse is not problematical when discourse is used to describe objects which are themselves linear. Trouble may arise when speakers have to describe entities which extend over more than one dimension. In particular, which element should the description start with? Which element should come next? People arrive at descriptions in a wide variety of ways even for very simple objects, but some descriptive strategies seem to be more typical, or better adapted, than others (cf. Denis, Carité, & Robin, in press; Denis & Denhière, 1990).

Suppose, for instance, that you have to describe the scene shown in Figure 1 to another person and that the only constraint is to start your description from the bus terminal. From the terminal, you can either describe the elements on the right branch first (Madison Street), then the elements on the left (Spring Street), or alternatively describe the elements on the left branch first, then those on the right. Both descriptions are

*The study reported in this chapter was supported by a research grant from the C.N.R.S., Programme de Recherche sur les Sciences de la Communication, A.S.P. "Communication multimédias : Interactions texte et image" (Contract Number 955869). Thanks to Milène France and Constance Greenbaum for their help in the preparation of the manuscript.

correct, but they illustrate directional choices which may create a differential cognitive load for the speaker (and probably also for the listener). Differential cognitive load is not related to the total number of elements to be reported on each side of the bus terminal, since the same number of elements have to be described on the left and the right branch. The problem is that when starting a description from the bus terminal in any direction, the speaker needs to keep track of the fact that after completing the description of the first branch, description of the second branch will have to start again from the bus terminal. If you describe the right side branch first, while describing it, you only have to maintain in memory the bus terminal as the element from which you will have to start the second part of the description. But if you describe the left branch first, you will have to maintain the same element for the same purpose, and in addition, at a given time, you will also have to keep in memory the park (when describing the school, for instance) as the point where to come back in order to report about the gas station.

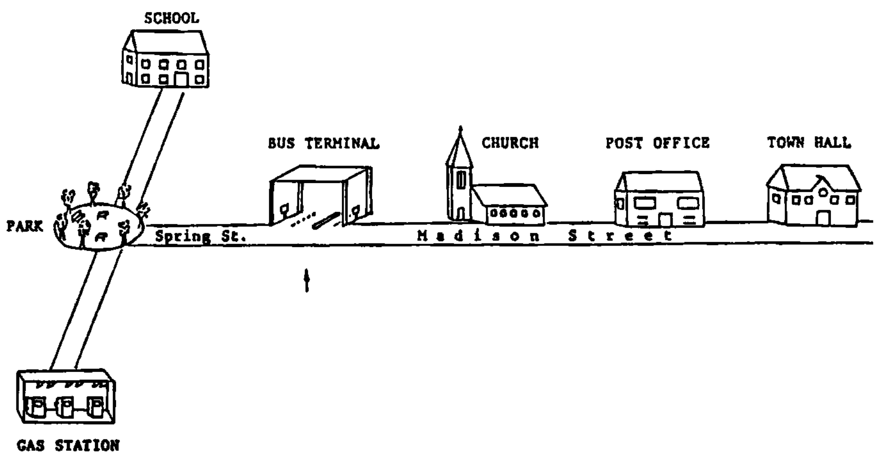

Figure 1. Example of an urban scene.

The available data indicate that such choices do not occur randomly, and that in fact, speakers generally prefer to make descriptions that minimize either the number of elements stored in working memory or the duration of such storage (Levelt, 1982, 1989; Robin, 1990). The present study was primarily designed to investigate some of the rules which govern the sequencing of discourse. In particular, we examined to what extent the structure of an object places constraints on the order in which its parts are entered into a description.

In addition to investigating descriptive strategies, there was a second objective to the present study. While one of the main functions of discourse is to describe objects and scenes in the environment, discourse also enables people to express their internal states to other persons. In particular, discourse can convey mental representations of objects which are temporarily (or permanently) absent from people's environments. Therefore, language can be seen as a means for externalizing internal representations, in particular mental images of objects. The second issue addressed here is thus whether descriptions of mental images exhibit similar regularities and sequencing as discourse describing objects.

There is ample evidence for structural and functional similarities between images and percepts (e.g., Denis & Cocude, 1989; Finke, 1989; Kosslyn, 1983; Sheehan, 1966; Shepard & Chipman, 1970). Images are assumed to be generated in a visual buffer exhibiting properties of coordinate space. If images are similar to percepts, constraints on the exploration of images should resemble constraints on perception. Consequently, when exploration of a pattern is the basis for verbalization, the processes which access visual images should exhibit the same regularities as those which access perceptual events.

Typical Strategies for Describing Spatial Networks

The experiment reported below was designed to compare descriptive strategies for either perceptual or mental representations. The materials were adapted from those used by Levelt (1982) to explore regularities in descriptive strategies. They consisted of networks in which coloured circles were connected to each other by horizontal and vertical lines. Subjects were visually presented with the networks, and asked to describe them so that another person could construct an accurate representation of them. Subjects only had one constraint: to start their description from a specific circle from which two branches went in opposite directions. There, subjects had to make a choice of which branch would be described first.

Levelt (1982) reports three main types of regularities for three classes of networks. An example of the *first class of networks* is shown in Figure 2a. Two branches, a short one and a long one, lead off from the green circle. The duration of maintenance of the starting point in working memory varies as a function of whether subjects have chosen to describe the short or the long branch first. Duration of storage before returning to describe the other branch is longer if the longer branch is described first. Thus, it should be cognitively more economical to describe the shorter branch first, since duration of storage will be shorter. The data collected by Levelt confirm that when subjects must start their description from the green circle, the probability is greater than 50% that they will describe the shorter branch before the longer one.

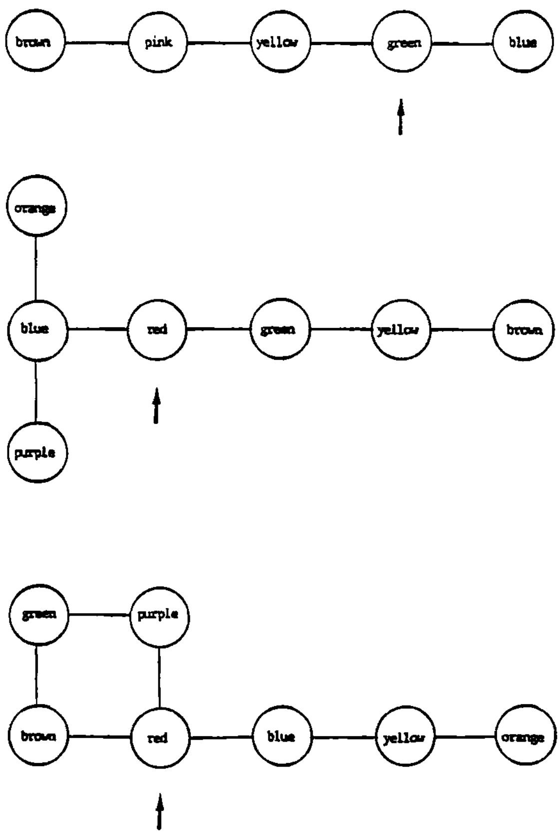

Figure 2. Three types of spatial networks. Adapted from Levelt (1982).

An example of the *second class of networks* can be found in Figure 2b. Subjects have to start their descriptions from the red circle. From there, two branches go in opposite directions. On one side, there is a linear branch and on the other side, a more complex branch (which itself leads to a further intersection). The linear branch and the complex branch have the same number of circles, and differ only for structural complexity. Consequently, less information needs to be stored transiently in memory during the description of the first branch for the linear branch since subjects only have to keep track of one circle, the red one, as the backtrack point before describing the other side of the network. If subjects describe the complex branch first, they need to store both the red and the blue circle in memory at some point in time. In fact, results show that when subjects start their descriptions from the red circle, the probability is greater than 50% that they will describe the branch without any embedded choice points before the branch with one or several embedded choice points.

The *third class of networks* is illustrated in Figure 2c. Three branches lead off from the red circle. One branch is linear, and the other two branches form a loop. Here, an adequate strategy may consist in describing the loop first, then the linear branch. Subjects who use this strategy simply avoid storing any intersection point. When following the loop, they go along a sort of a linear route and avoid the intersection which would force them to make a choice. However, when subjects start their descriptions from the red circle, they in fact have to choose among three directions. Consequently, the probability for them to describe the loop before the linear branch should be larger than 67%. This tendency, in fact, was only observed for a subset of subjects in the Levelt study.

Our experiment used the same kind of materials, but new variants on the network structures were introduced in order to explore several factors which were not dealt with in the Levelt study systematically. In some cases, the differential load between the two branches was increased by adding circles on one side. In other cases, we only changed the visual presentation of a branch. These modifications were expected to affect the probability of occurrence of typical descriptive strategies.

Fifty adult subjects participated in this experiment. Each was first involved in a perceptual condition, where they had to describe visually presented networks, and then in an imaginal condition (two weeks later). In the imaginal condition, a given trial consisted for the subjects firstly to memorize a network without time limitation. When subjects stated they had formed a clear visual image, the network was withdrawn. They were then asked to describe the network. In both perceptual and imaginal conditions, subjects had only one constraint: to start their description from a specific circle which was mentioned just before they started their description.

Description of Perceived Networks

Figure 3 shows the frequencies of typical strategies for the three classes of networks.

As expected, for the *first class of networks,* when both branches had the same length (network 1.1), no clear preference emerged for either direction. When circles were added to one side of the network, the cognitive load involved in the processing of the longer branch was expected to increase, and hence the probability for subjects to describe the shorter branch first. The data confirmed expectations, since 68% of the subjects preferred to describe the shorter branch before the longer one (network 1.2). This probability rose to 78% when the difference between the two branches was further increased (network 1.3). Therefore, variations designed to affect cognitive load during the description of the longer branch increased the probability of using the expected strategy.

For the *second class of networks,* the majority of the subjects (63%) preferred to describe the linear branch before the more complex branch (network 2.1), again confirming expectations. Our first variation consisted in adding one circle to each branch (network 2.2). The network thus still had same number of circles on both sides, but the relative complexity of the complex branch was increased, since the subjects had to return one additional time to the embedded circle during the description of the complex branch. This situation was expected to increase the probability for subjects to describe the linear branch before the complex branch. This was indeed the case, although the effect was weak (from 63% to 69%).

Greater impact was observed for the second variation on this class of networks. Presentation of the linear branch was modified to make it purely visually more complex. Instead of extending along a straight line, the linear branch was distorted. If one compares networks 2.1 and 2.3 for instance, both branches in each case have the same number of circles. Only the structure of the linear branch differs. This modification substantially lowered the probability of using the expected strategy. Results show that only 52% of the subjects described the linear branch first for network 2.3, and the probability further decreased to 43% for network 2.4.

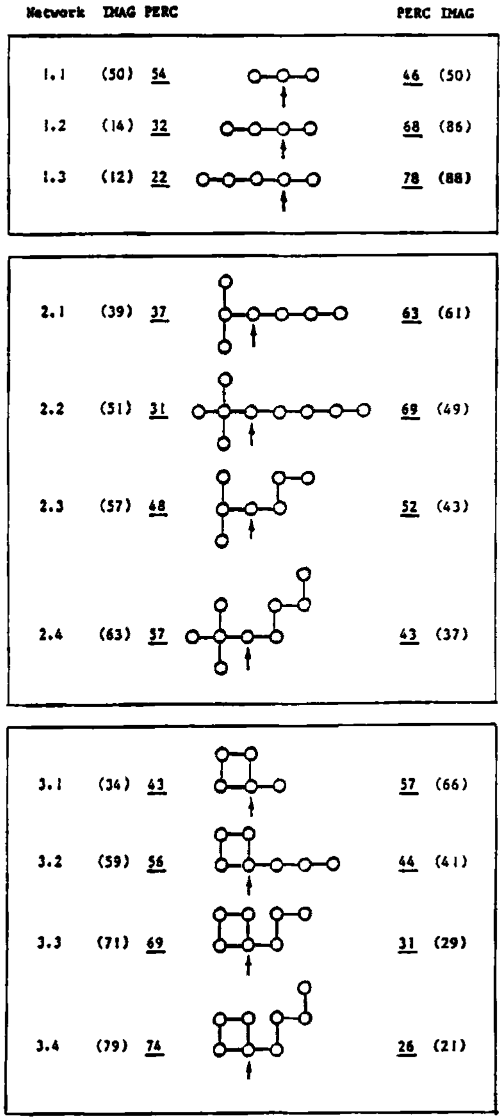

Figure 3. Materials used in the experiment. Arrows indicate starting points for descriptions. The proportions of subjects exhibiting each directional choice are given for each network (average proportion for each network and its mirror version) (underlined, perceptual condition ; in brackets, imaginal condition).

To sum up, when the linear branch run along a straight line, subjects apparently preferred to describe it first. But as soon as its presentation was made visually more complex, the probability of using this typical strategy decreased sharply. Subjects even tended to describe the complex branch first (in particular in the case of network 2.4) whose relative (perceived) complexity was lower. Subjects thus do not only judge branch complexity in terms of the number of embedded points, but also in terms of the structural complexity (or rate of distortion) of the linear branch.

The strategy expected by Levelt to be dominant for the *third class of networks* consists of describing the loop before the linear branch. The occurrence of this strategy, however, is apparently dependent to a large extent on the number of circles composing the linear branch. When this branch was very short (only one circle) (network 3.1), only 43% of the subjects described the loop first (which is far below Levelt's expectations). When the load of the linear branch was increased by adding further circles (network 3.2), the probability of describing the loop before the linear branch increased. Nevertheless, the proportion of subjects who used the loop-first strategy still remained moderate (56%).

In network 3.3, we modified the structure of the linear branch, making it visually more complex while having the same number of circles as in network 3.2. We expected that this modification would increase the probability that subjects describe the loop before the linear branch. It was found indeed that 69% of the subjects used the expected strategy in this case. The probability of using this strategy still increased when both visual complexity and load of the linear branch (in terms of number of circles) were increased (network 3.4). Here, 74% of the subjects used the expected strategy.

Our data thus clearly suggest that the choice of descriptive strategies is governed by cognitive factors. Modifications in the structure of the branches increased cognitive load associated with the processing of these branches. As a consequence, these modifications systematically increased the likelihood of some specific strategies.

Description of Imagined Networks

The aim of the second part of the experiment was to evidence similarities, if any, between the strategies used to describe visually presented networks and those used to describe visual images of networks.

Not surprisingly, similar proportions of subjects preferred to go left or right first on the *first class of networks* (network 1.1). As in the perceptual condition, subjects did not show any preference for either side. When the load of one branch was increased, the proportion of subjects exhibiting the typical strategy (i.e., describe the shorter branch first) dramatically increased to 86% (network 1.2) and 88% (network 1.3). This effect is even more marked than the one which occurred in the perceptual condition. Changes in the networks thus similarly increased the probability for subjects to describe the shorter branch before the longer one.

For the *second class of networks*, 61% of the subjects preferred to describe the linear branch before the complex branch (network 2.1). This strategy is similar to the one used in the perceptual conditions; that is, subjects preferred to describe the branch which involved the temporary storage of only one circle first. When the network was modified by adding one circle to each branch (network 2.2), the complex branch became relatively more complex. Whereas this modification slightly increased the probability for subjects to describe the linear branch first in the perceptual condition, the probability here decreased to 49%. In fact, this outcome can be accounted for by assuming that subjects experience more difficulty in processing the periphery of a visual image (cf. Kosslyn, 1980). Network 2.2 is the one whose linear branch extends the farthest at the periphery, and this may have entered into subjects' reluctance to process the linear branch first.

The probability of describing the linear branch first decreased still further when the presentation of this branch was distorted. As in the perceptual condition, this modification lowered the probability of implementing the typical strategy: 43% of the subjects described the linear branch first (network 2.3). When both presentation and load of the linear branch were again modified, the probability of using this strategy decreased even more : only 37% of the subjects preferred to describe the linear branch first (network 2.4).

When the linear branch forms a straight line, subjects obviously prefer to describe it first. But when the linear branch is made visually more complex, the probability of using the typical strategy (i.e., describe the linear branch first) decreases sharply. This effect occurs in quite similar ways in the perceptual and imaginal conditions.

For the *third class of networks*, in the perceptual condition, subjects preferred to describe the linear branch before the loop when the linear branch was short. A similar effect was obtained when subjects described this sort of network from a visual image : 34% of the subjects preferred to describe the loop before the linear branch (network 3.1). We expected that if subjects processed information in visual images and percepts in similar ways, increasing the load of the linear branch should increase the probability that they would describe the loop before the linear branch (network 3.2). The results confirmed expectations : in imagery, 59% of the subjects preferred to describe the loop first.

We also expected the probability of describing the loop before the linear branch to further increase when the presentation of the linear branch was more complex. As a matter of fact, when the linear branch was distorted (network 3.3), 71% of the subjects used the typical strategy. Again, similar to the perceptual condition, the probability increased when both load and visual presentation of the linear branch were modified (network 3.4): 79% of subjects preferred to describe the loop first.

Thus, for this class of networks, the factors which increase the probability of occurrence of the typical (loop-first) strategy affect perceptual and imaginal conditions in closely parallel fashions.

Additional analyses provided further information regarding within-subject consistency in strategies used for describing perceived or imagined networks. High consistency was found for those networks where a clear dominant strategy was evidenced in the perceptual condition. For instance, there was a significant tendency for subjects who used the typical strategy for networks 1.2 and 1.3 in the perceptual condition (i.e., describe the shorter branch first) to use this strategy again in the imaginal condition. Similar high levels of within-subject consistency were found for networks 2.1, 2.4, 3.3, and 3.4.

Further analyses also revealed that whereas the structures of descriptions were similar for perceptual and imagined networks, a greater amount of cognitive effort was required when subjects had to describe visual images than when they described visually presented networks. Here, the most informative index of cognitive effort was the time elapsing before subjects started their descriptions, once they had been informed of which circle the description should start from. Overall, these latencies were longer in the imaginal than in the perceptual condition. In addition, in both conditions, increasing the complexity of networks resulted in increase of latencies, but this effect was much more marked in the imaginal condition. Thus, if similar regularities can be evidenced in accessing the entity to be described in both perception and imagination, a larger amount of cognitive effort seems to be associated with the maintenance of a mental image during the process of description.

Concluding Comments

When people produce descriptions of spatial configurations, their dominant strategies clearly reflect the incidence of cognitive factors. In addition, more than one variable determines strategy choices. Furthermore, analysis of descriptions reveals close similarities between strategies used to describe objects and images of objects. There is a single case where strategies in the imaginal condition do not parallel those observed in the perceptual condition (network 2.2), but current assumptions regarding image processing easily account for this case.

Our findings are thus compatible with the hypothesis that similar processes are used to access perceptual events and visual images of corresponding objects. Factors weighing on descriptive strategies affect perceptual and imaginal conditions in a similar way. These similarities can be seen as reflecting similar constraints bearing on the processing of perceived objects and visual images.

REFERENCES

Denis, M., Carité, L., & Robin, F. (in press). Cognitive approaches to the processing of descriptions. *Zeitschrift für Psychologie.*

Denis, M., & Cocude, M. (1989). Scanning visual images generated from verbal descriptions. *European Journal of Cognitive Psychology* **1**, 293-307.

Denis, M., & Denhière, G. (1990). Comprehension and recall of spatial descriptions. *Cahiers de Psychologie Cognitive / European Bulletin of Cognitive Psychology* **10**, 115-143.

Ehrich, V., & Koster, C. (1983). Discourse organization and sentence form: The structure of room descriptions in Dutch. *Discourse Processes* **6**, 169-185.

Finke, R.A. (1989). *Principles of mental imagery.* Cambridge, MA : The MIT Press.

Isaacs, E.A., & Clark, H.H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology : General* **116**, 26-37.

Kosslyn, S.M. (1980). *Image and mind.* Cambridge, MA : Harvard University Press.

Kosslyn, S.M. (1983). *Ghosts in the mind's machine : Creating and using images in the brain.* New York : W. W. Norton.

Levelt, W. J.M. (1982). Linearization in describing spatial networks. In S. Peters & E. Saarinen (Eds.), *Processes, beliefs, and questions* (pp. 199-220). Dordrecht, The Netherlands : Reidel.

Levelt, W. J.M. (1989). *Speaking : From intention to articulation.* Cambridge, MA : The MIT Press.

Linde, C., & Labov, W. (1975). Spatial networks as a site for study of language and thought. *Language* **51**, 924-939.

Robin, F. (1990). La description de configurations spatiales : Approche développementale. *Cahiers de Psychologie Cognitive / European Bulletin of Cognitive Psychology* **10**, 433-447.

Shanon, B. (1984). Room descriptions. *Discourse Processes*, **7**, 225-255.

Sheehan, P.W. (1966). Functional similarity of imaging to perceiving : Individual differences in vividness of imagery. *Perceptual and Motor Skills* **23**, 1011-1033 (Monogr. Suppl. 6-V23).

Shepard, R.N., & Chipman, S. (1970). Second-order isomorphism of internal representations : Shapes of states. *Cognitive Psychology* **1**, 1-17.

Ullmer-Ehrich, V. (1982). The structure of living space descriptions. In R.J. Jarvella & W. Klein (Eds.), *Speech, place, and action : Studies in deixis and related topics* (pp. 219-249). Chichester : Wiley.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 11

Building referents of indeterminate sentences in the context of short narratives*

Manuel de Vega and José M. Díaz
University of La Laguna, Spain

INTRODUCTION

The purpose of this paper is to study the comprehension of indeterminate or "abstract" sentences included in short texts. We will explore how subjects build a representation of the reference for these indeterminate sentences, with an emphasis on the time course of the process. The processing of abstract versus concrete words has been a common topic of research within the framework of mental imagery. For instance, Paivio's dual-code hypothesis proposes that concrete words are usually coded in memory both verbally and imaginally, whereas abstract words are only coded in the verbal system (Paivio, 1971). However, most of the studies have analyzed the role of abstractness in tasks involving isolated words (e.g., James, 1975; Richardson, 1976; Kroll & Merves, 1986) or sentences (e.g., Begg & Paivio, 1969). Instead, the present study will explore how sentences involving an indeterminate word are processed in the context of short narratives.

Four-sentence texts were manipulated in order to produce several scenario-inducing conditions. A fifth verification sentence was added for each text as we will explain later. The texts involved some alternative words that were either indeterminate or scenario-inducing words. An example of the material is illustrated in Table 1.

There are two experimental factors: (a) Initial target word (in the first sentence): either an indeterminate or a resolution word, and (b) Final target word (in the fourth sentence): either indeterminate or

*This research was supported by the Spanish Ministerio de Educación y Ciencia, CICYT Grant PB86-0480 to the first author.

resolution. Therefore, four versions of each text were obtained by crossing both factors: Indeterminate & Indeterminate, Indeterminate & Resolution, Resolution & Resolution, Resolution & Indeterminate. The reading times for each sentence were recorded. In addition we recorded the time subjects spent in judging whether or not the fifth sentence was an appropriate continuation of the text.

Table 1. Example of the experimental material

	Sentence	Type of sentence
(1a)	He sat in the main seat comfortably.	Indeterminate
(1b)	He sat in the pilot seat comfortably.	Resolution
(2)	He looked around him.	Filler
(3)	Everything was in order.	Filler
(4a)	He started the operation skillfully.	Indeterminate
(4b)	He started the takeoff skillfully.	Resolution
(5)	He communicated with the control tower	Verification

By analyzing reading times under different task conditions we explored two related questions, first, whether indeterminate sentences are instantiated as specific referents or, conversely, they are merely coded as abstract (e.g., propositional) representations. For instance, does the reader exclusively encode an abstract proposition for (1a) or is he able to elaborate a specific mental model of its referent. A longer reading time for (1b) than for (1a) should indicate that the reader does initially encode the indeterminate sentence in a surface code (propositional) rather than trying to elaborate a specific representation of the referent. Otherwise, if the reading times for (1a) and (1b) do not differ, or if (1a) takes longer than (1b), then this might suggest that readers elaborate the referent of indeterminate sentences, as soon as possible.

Secondly, assuming that readers are able to instantiate a specific referent for indeterminate sentences, we will explore the temporal course of this instantiation. An interesting possibility is that the instantiation of (1a) does not occur on-line, but it is delayed till a later segment of the text is read. If this were the case, two predictions might be made. First, reading times for (1a) would be shorter than for (1b), namely the indeterminate sentence would require less cognitive resources because there is no initial attempt to elaborate its referent. Secondly, the reading time for sentence (4b) would be shorter after reading (1b) than after reading (1a). That is to say, the lack of a previous instantiation in the indeterminate sentence (1a) involves a later referent construction when the resolution is provided in (4b).

Let us focus on the preliminary question of whether indeterminate sentences are instantiated in terms of concrete referents or, conversely, they are encoded in more surface codes. According to the mental model approach to comprehension, the listener or reader of a text builds up a non-linguistic representation of discourse's reference. This representation has been called a mental model, situation model, possible world, etc. The main insight of the mental models framework is that the comprehender builds up a representation of the referent of the text (the objects, events, characters, and processes described in the text) rather than, or in addition to, a representation of the text itself (Johnson-Laird, 1983; Sanford & Garrod, 1981; van Dijk & Kintsch, 1983; Garnham, 1987; Glenberg et al., 1987). Several experiments provide evidence that comprehension involves the building in the reader's mind of a dynamic mental model. The mental model is established by the reader at the beginning of the text, and it is gradually updated as new segments of the text are processed. Eventually, the incoming text does not cohere with the previous model, and a new model is formed.

The representation format of mental models has not been studied in much detail so far. However, it is likely that at least some kind of situation models involve mental imagery (e.g., Mani & Johnson-Laird, 1982; Johnson-Laird, 1983; Perrig & Kintsch, 1985). On the other hand, most of the studies within the mental model approach involve texts describing specific environments with perceptual objects and events which are easily modeled as images.

The idea of an image-based mental model for concrete referents (e.g., physical scenarios with objects and events) is easy to accept. However, texts involving indeterminate words and sentences seem to challenge the mental models approach. It is unlikely that indeterminate sentences activate any form of situation model, in spite of which they are somehow understood. A possibility for indeterminate sentences is that they are encoded as generic or abstract traces. However, some problems arise from the idea of abstract meaning. Thus, objects belonging to abstract or superordinate categories like "furniture", or "game" share few, if any, semantic features (Rosch et al., 1976), therefore the abstract meaning conveyed by these categories is negligible. The solution proposed by Rosch (1975) is that generic categories are understood in terms of their prototypes or best exemplars that operate as "reference points". Later research has shown that subjects who read abstract terms (superordinate) in the context of sentences are able to instantiate their contextually appropriate exemplars (e.g., Anderson & Shiffrin; 1980; Roth & Shoben, 1983; Dubois & Denis, 1988).

For instance, Roth & Shoben gave the subjects one of three alternative contexts involving a superordinate word:

(5a) Mary watched the bird all day (Neutral context)
(5b) Mary saw the bird swimming (Test biasing)
(5c) Mary looked at the bird on the telephone wire (Biasing against test)
(6) Mary was very fond of ducks (Test sentence)

The reading times for the test sentence (6) were faster after receiving the biasing context (5b) than after reading the neutral context (5a). In addition, the reading time for (6) increased after reading the context (5c) that biases against the test sentence meaning. These results suggest that the linguistic context determines a particular instantiation of the categorical term bird. The contextual effect of instantiation is so strong that the standard values of exemplars typicality momentarily change. Further research by Dubois and Denis (1988) has shown that instantiation of superordinate terms can be figurative or akin to perception.

In the present paper we hypothesize that indeterminate words and sentences in the context of short narratives are usually instantiated as specific referents. In addition, abstract terms or sentences do not activate referents by themselves but they benefit from contextual cues provided by other sentences.

The previous expectations are consistent with the results of the literature described above. Notice that in those experiments the resolution context that favours a given instantiation of the abstract word was provided previously or in the same sentence as the critical word (previous context instantiation). However, in the present research we provided the resolution several sentences after the critical indeterminate sentence (late context instantiation). The late context instantiation was chosen because we believe that it is quite frequent in natural narrative texts. Furthermore, it was interesting to analyze the time course of referent construction under such a task.

Consequently a main goal of this paper is to reveal the time course of instantiation in late context situations. We will discuss two possibilities: (a) The immediate instantiation hypothesis: The reader computes the referent on-line while reading the indeterminate word/s, and (b) The delayed instantiation hypothesis: The instantiation of the indeterminate sentence takes places several sentences after the word has been read. Actually, the instantiation takes place when a later segment of the text provides a new semantic clue that is sufficiently specific. When the reader encounters the indeterminate term he or she might encode it in a surface code (propositional). The surface code would be kept in a buffer until a new word allows a backward instantiation of the target word.

The immediate instantiation hypothesis posits that the construction of a referent for indeterminate word/s takes place on-line. This proposal is derived from the immediacy hypothesis assumed by some students of reading (Just & Carpenter, 1980; Marslen-Wilson, 1975). According to the immediacy hypothesis, the reader processes each word "as soon as possible", typically while he or she is looking at the word. Immediacy seems to have some advantages in terms of processing economy. For instance, the immediate interpretation of each word does not require the temporary storage of surface codes.

However, in late context conditions as in our experiment, the immediate instantiation of indeterminate words might have undesirable consequences. Let us suppose that the first sentence in a text is (1a): "He sat in the main seat comfortably". The sentence is indeterminate because the adjective *main* does not specify what kind of seat, and consequently what kind of scenario, the sentence is referring to. An immediate instantiation of a sentence referent might determine a "mental model" in the reader's mind that very likely will be wrong. For instance, a given reader would interpret that the "main seat" is the boss's seat in an office. However, the incoming sentence (4b) would force a new interpretation. This might produce a very expensive (in terms of cognitive resources) garden-path, as the reader has to reject a previous model and to elaborate a new one.

According to the hypothesis of delayed instantiation, readers are conservative when they process indeterminate terms. They do not risk an incorrect interpretation of the sentence. Instead, the elaboration of the indeterminate word's referent can be delayed for several sentences. Actually, the instantiation only occurs when a later segment of the text provides a new semantic clue that is sufficiently specific. When the reader encounters the indeterminate word/s he or she might initially encode it in a linguistic or surface code. This code would be stored in a buffer until a new word or sentence allows a backward instantiation of the indeterminate word. This "wait-and-see" procedure avoids resource-demanding garden-paths. However, the model must include a buffer in order to store the non-instantiated indeterminate words and sentences.

This buffer presumably consumes working memory space and, therefore, there is a limit on the delay of the instantiation procedure. The system might collapse if the amount of non-instantiated words increases beyond some limit. This is exactly what happened in the classical research of Bransford and Johnson (1973). When the subjects received an indeterminate version of a relatively large text, in which it was extremely difficult to elaborate its referent, both memory and comprehension decreased.

METHOD

Subjects

Twenty four students of the University of La Laguna (14 females and 10 males) participated as subjects as a means of fulfilling an introductory psychology course requirement. They all were native Spanish speakers.

Materials and design

Forty eight short texts were selected for the experiment. Each text involved 5 sentences (see Table 1). There were two different versions of the first sentence that differed in a single target word, that was either an indeterminate or a resolution word. There were also two different versions of sentence 4 involving an indeterminate or a resolution target word, respectively. Sentences 2 and 3 were filler sentences that remained the same in all versions of the task. Finally, sentence 5 involved a verification sentence that also was common to all versions of the task.

There were four experimental conditions resulting from manipulating the initial target word (first sentence) and the final target word (fourth sentence): Indeterminate & Indeterminate, Indeterminate & Resolution, Resolution & Resolution, Resolution & Indeterminate.

To ensure that the manipulation of the target words in the critical sentences was appropriate the following normative data were collected. A sample of 55 subjects, who did not participate in the experiment reported here, received an initial set of 72 four-sentence texts. Each subject read each text in one of the four possible versions. They were instructed for each text to suggest a title and then to write a sentence that was a plausible continuation of the text.

Six judges classified the subjects' answers as coincident or non-coincident with the authors interpretation. For instance let us suppose that a subject, who receive an "indeterminate & indeterminate" version of the text described in Table 1, gives it the title "In the office". Such a response might be classified as non-coincident. However, the title "Taking off" would be considered coincident. To be selected, a text should satisfy two requirements: (a) Most subjects' answers were coincident in their interpretations when they received a resolution version of the text; (b) Most subjects' answers were non-coincident in the version without any resolution ("indeterminate & indeterminate"). The 48 experimental texts were selected because they satisfied both criteria.

From these experimental texts, four sets of material were constructed. Within a material set, there was an equal number of experimental texts in the four experimental conditions. Across material sets, each experimental text occurred in all four of its experimental conditions. Six subjects were randomly assigned to each material set.

Procedure

Subjects sat in front of a computer screen on which they were shown the texts. Subjects read each text sentence by sentence in a self-paced mode. In the fifth sentence subjects were instructed to judge whether or not the sentence was "a reasonable continuation of the text". The YES and NO responses were previously assigned to two alternative keys of the keyboard. Immediately after the verification response, a seven point rating scale was presented on the screen, and subjects had to assess their confidence in the answer. The computer recorded the reading times corresponding to each sentence as well as the judgment times of the verification sentence and the confidence ratings.

Each subject received a given material set involving texts in all four experimental conditions. For each subjects, the order of presentation was automatically randomized by the computer.

After completing the whole set of texts, subjects received an unexpected memory task. They were given a booklet with the first three sentences of each text, and they were asked to write down the two final sentences.

RESULTS AND DISCUSSION

Analyses of variance (ANOVA) was performed for each dependent measure. The independent variables were the Initial target word and the Final target word. In the analyses of reading times for the first and fourth sentences two covariants were introduced. The covariants were the length (number of characters) and the frequency in print of the target words. These covariants were introduced as an a posteriori control in order to dissociate their effect from the experimental manipulation of the target words. None of the covariants were significant in any analysis, therefore the effects of the experimental variables cannot be attributed to any bias in these lexical parameters.

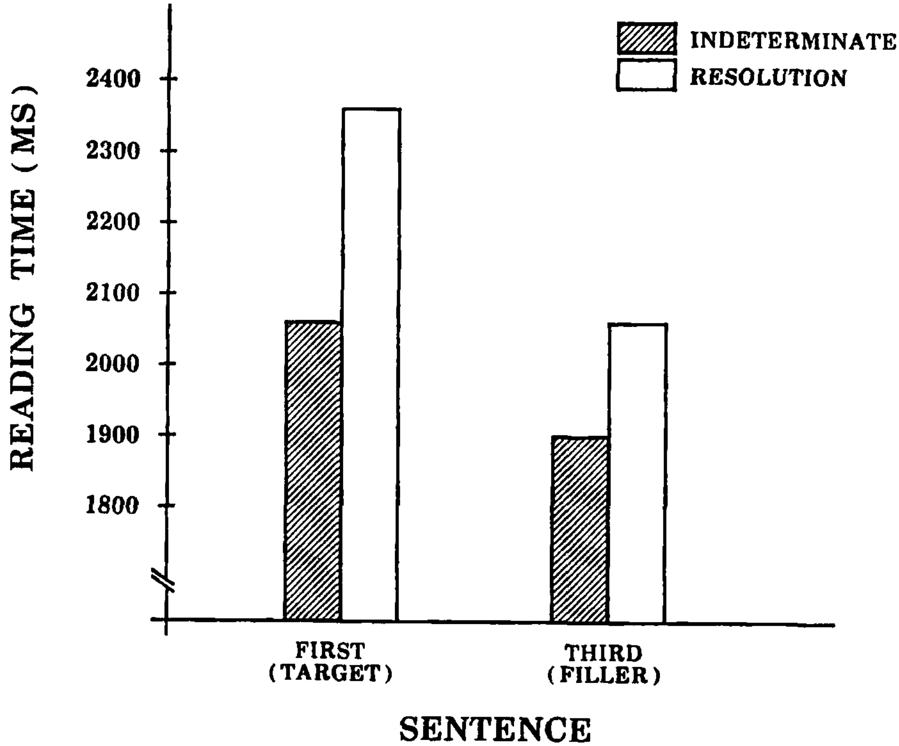

Figure 1. Reading times for the first sentence (target) and the third sentence (second filler) as a function of Initial Target Word.

Reading times for the three first sentences

The only independent variable on reading times for the first three sentences was the Initial target word. Subjects read the first sentence more slowly, when it included a resolution word rather than an indeterminate word ($F(1,22)=8.13$, $p<.01$, for subjects; $F(1,44)=8.90$, $p<.01$, for items). Furthermore, there was a main effect of the Initial target word on reading time for sentence three ($F(1,22)=6.14$, $p<.05$, for subjects; and $F(1,46)=5.79$, $p<.05$, for items). In particular this filler sentence was read more slowly when there was an initial resolution word. These results are illustrated in Figure 1.

The longer reading times with initial resolution suggest that subjects elaborate a representation of the referent in this condition. Conversely, under the indeterminate condition readers spend less time on each sentence probably because they build a less demanding linguistic code and they do not try to elaborate the referent. It is remarkable that the indeterminacy or resolution of a single word in the first sentence determines differences on reading time of a filler sentence. This long term effect indicates that the experimental manipulation has to do with processes at the level of text integration rather than at the lexical level. It is likely that in the initial resolution condition, once the reader has elaborated the referent (e.g., a mental model) in the first sentence, he or she has to spend some effort in integrating sentence three in order to update the initial mental model. This integrative effort is not required when the reader is relying on a propositional code (indeterminate condition).

Reading times for the fourth sentence

The Initial x Final target word interaction was significant ($F(1,22)=5.54$, $p<.05$, for subjects; and $F(1,44)=7.22$, $p<.01$, for items). Figure 2a illustrates this interaction.

The slowest reading time corresponded to the "Indeterminate & Resolution" condition. In this condition subjects spent significantly more time than in the "Resolution & Resolution" condition ($LSD=281$; $p<.05$). A likely explanation of this difference is that in the "Indeterminate & Resolution" condition, subjects have to elaborate for the first time a representation of the text referent, and this is a demanding operation because it involves a backward elaboration of the previous sentences, that have been stored in a propositional code. Instead, in the "Resolution & Resolution" condition, subjects elaborate at the beginning, a

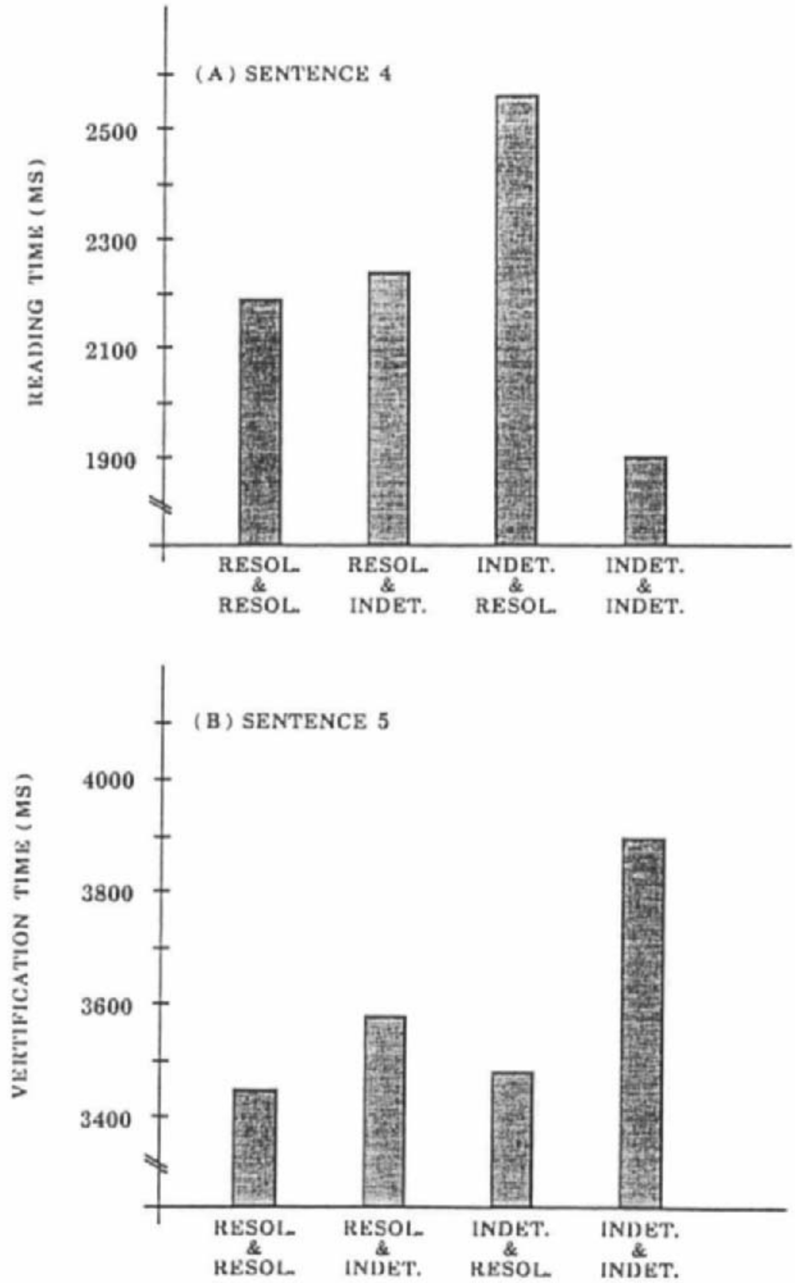

Figure 2. Reading times for the (a) fourth sentence and (b) the fifth sentence as a function of Initial Target Word and Final Target Word.

representation of the referent and when they read the fourth sentence they expend effort only in the normal integrative process. In other words, in the former condition subjects have to lay the foundation of a mental model from relatively surface codes, whereas in the latter condition subjects updated an ongoing mental model.

Another finding illustrated by Figure 2a is that reading time for the "Indeterminate & Indeterminate" condition was faster than for any other condition (LSD=281; $p<.05$). Apparently when readers received this highly indeterminate version of the task they did not make any special effort to elaborate a top-down representation of the referent either at the beginning or at the end of the text. However, this conclusion can be premature as the subject had a last chance to generate a model in the last verification sentence. So let us comment next on the verification data.

Verification time

It took longer to verify the fifth sentence when the previous sentence included an indeterminate word ($F(1,22)=4.86$, $p<.05$, for subjects; $F(1,46)=6.67$, $p<.05$ for items). There was a sort of trade-off between fourth-sentence reading times and verification times as can be seen by comparing Figure 2a and 2b. This trade-off is particularly clear in the "Indeterminate & Indeterminate" condition. Apparently, under this condition the reader does not elaborate any referent during the four-sentence text. However the price of this low demanding text reading is an increase in the difficulty of the verification task. The reader in the "Indeterminate & Indeterminate" condition has no referent representation and therefore the verification requires a late elaborative effort to make sense of the text. It is probably that the fifth sentence provides a cue to resolve the reference. Instead, the other task conditions result in a relatively speedy verification, because they all involve a previous construction of the referent.

Non-chronometric measures

Subjects felt more confident of their judgments when the first sentence provided a resolution word rather than an indeterminate word (the mean rates were 6 and 5.7, respectively). The difference, although small, was significant ($F(1,22)=5.97$, $p<.05$, for subjects; and $F(1,46)=6.58$, $p<.05$, for items). Therefore, readers experience uncertainty when there is a lack of resolution in the first sentence. This is consistent with the suggestion that in such a situation they read three sentences without any attempt to solve the referent.

Concerning recall, the index was obtained by counting the number of content words (or their paraphrases) correctly recalled. The ANOVA on recall showed that recall was better when the first sentence provided a resolution word ($F(1,22)=8.74$, $p<.01$, for subjects; and $F(1,46)=9.66$, $p<.01$ for items). Furthermore, recall also improved when the fourth sentence include a resolution word ($F(1,22)=6.87$, $p<.05$, for subjects; and $F(1,46)=8.34$, $p<.01$ for items). The interaction between Initial target word and Final target word was not significant ($F<1$).

GENERAL DISCUSSION

From the above results, a coherent picture emerges of the dynamic processes involved in building the reference of a text. Readers build up a mental model of the referent as soon as possible. However, when the first sentences do not provide specific information (indeterminate target word) subjects do not spend time trying to build on-line a top-down model, as their fast reading times suggest. Therefore, the hypothesis of immediate instantiation of indeterminate words can be rejected. Instead, subjects appear to be using a wait-and-see procedure. Namely, they initially build on-line a simplified representation (a linguistic-like or propositional code), while they expect that further text sentences will provide them with better clues to solve the reference. Under initial indeterminate conditions subjects not only code superficially the first sentence but also the incoming sentences. This claim is suggested by the fact that one of the filler sentence (the third sentence) was also read faster under the initial indeterminate condition.

It is clear from the analyses of covariance that these results cannot be attributed to purely "lexical effects" such as differences in word-frequency or word-length between the indeterminate and the scenario-inducing words. On the other hand, indeterminate and resolution conditions were associated with "abstract" and "concrete" target words, respectively. However, the effect of the target words cannot be explained in terms of a local concreteness effect at the lexical level. Rather it was a whole-text effect. For instance, the manipulation of the initial target word extended its effects to the reading of filler sentences as well as the fourth and the verification sentences.

The wait-and-see strategy we have described is similar to that described in the literature on elaborative inferences (McKoon & Ratcliff, 1986; O'Brien et al. 1988; Garrod et al. 1990). According to these authors, when subjects read texts that involve predictable events they only encode minimally the corresponding inference at the time of reading. Thus, if a text describes someone who falls off of a 14th story roof, the "death" of the character is highly predictable. However, subjects do not encode such predictable events, but they appear to represent only a few semantic features, such as "something bad". In the present experiment readers follow a similarly conservative strategy. They avoid engaging in elaborative processing of indeterminate information. Instead subjects code the indeterminate sentences in a content-free format and store it in a buffer until further resolution cues were provided.

The long term consequences of the wait-and-see strategy were analyzed in fourth and fifth sentences. When the fourth sentence provides the expected resolution cues, after a first indeterminate sentence, then readers elaborate a model of the referent, with a considerable demand on resources. This late referent construction is most likely used to make a backward integration of the previous texts contents. However, under highly indeterminate conditions (indeterminate & indeterminate) the reader does not try to elaborate a conceptually driven referent as is suggested by the fast reading times.

Elaborating the text referent seems to be a very demanding activity. Not only the resolution sentences are read more slowly but even filler sentences require more time after reading a resolution sentence. However, once the referent has been elaborated some functional advantages emerge. Thus, subjects who read texts involving resolution cues are faster in verifying sentences about them, and their recall and confidence improve. Referent construction provides an integrative mechanism that facilitates incorporation of new pieces of information, and the assessment of other contents. In addition, integrated memory traces are more easily retrieved in a later memory test.

The ideas developed above involve two levels of representation in text comprehension. Readers elaborate immediately a linguistic representation for each sentence, and eventually they build an integrated mental model of the referent. The above notion is similar to the double-stage model that has been defended by some authors in the comprehension literature (e.g. Cutler, 1976; Mani & Johnson-Laird, 1982). The linguistic code is a content-free representation of the text that demands few cognitive resources and it is built on-line. It operates as a back-up code that allows a superficial encoding of the text even when the referent construction is not immediately available. The elaboration of a representation of the referent (mental model) is more resource demanding and less text-bounded. It may include thematic inferences based on world knowledge. When the sentences or the texts are explicit enough the mental model is built and updated almost on-line. It is likely that the referent model updating occurs at the end of sentences or clauses (Haberlandt & Graesser, 1985; de Vega et al., 1990). However, when the text does not provide explicit information the reader can delay by several sentences the elaboration of the referent, as this experiment has demonstrated.

Concerning the representational format of the two levels of representation, the linguistic code is a discursive representation that can be described as "propositional". On its side, the referent representation or mental model may involve isomorphic features such as mental images and their transformations. Furthermore, it is also isomorphic with the serial structure of the representation's entities that mimics the temporal structure of events in the world.

REFERENCES

Anderson, R.C. & Shiffrin, Z. (1980). The meaning of words in context. *In* R.J. Spiro, B.C. Bruce, & W.F. Brewer (Eds.), *Theoretical issues in reading comprehension* . Hillsdale, NJ: Erlbaum Associates, pp 331-348.

Begg, I. & Paivio, A. (1969). Concretness and imagery in sentence meaning. *Journal of Verbal Learning and Verbal Behavior* **8**, 821-827.

van Dijk, T.A. & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Bransford, J.D. & Johnson, M.K. (1973). Considerations of some problems of comprehension. *In* W.G. Chase (Ed.) *Visual information processing.* New York: Academic Press, pp 383-438.

Cutler, A. (1976) Beyond parsing and lexical look-up. An enriched description of auditory sentence comprehension. *In* R. J. Wales and E. Walker (Eds.) *New approaches to language mechanism.* Amsterdam: North-Holland.

Dubois, D. & Denis, M. (1988). Knowledge organization and instantiation of general terms in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**, 604-611.

Garnham, A. (1987). *Mental Models as Representations of discourse and text.* Chichester: J. Wiley & Sons.

Garrod, S., O'Brien, E.J., Morris, R.K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory and Cognition* **16**, 250-257.

Glenberg, A.M., Meyer, M. & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language* **26**, 69-83.

Haberlandt, K. & Graesser, A.C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology: General* **114**, 357-374.

James, C.T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance* **1**, 130-136.

Johnson-Laird, P.N. (1983). *Mental models.* Cambridge, England: Cambridge University Press.

Just, M.A. & Carpenter, P.A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review* **87**, 329-354.

Kroll, J.F., & Merves, J.S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 92-107.

Mani, K., & Johnson-Laird, P.N. (1982) The mental representation of spatial descriptions. *Memory and Cognition* **10**, 181-187.

Marslen-Wilson, W.D. (1975). Sentence perception as an interactive parallel process. *Science* **189**, 226-227.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 82-91.

O'Brien, E.J., Shank, D.M., Myers, J.L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur online? *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**, 410-420.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

Richardson, J.T.E. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology* **67**, 315-325.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology* **7**, 532-547.

Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M. & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology* **8**, 382-439.

Roth, E. M. & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology* **15**, 346-378.

Sanford, A. & Garrod, S. (1981). *Understanding written language: Exploration of comprehension beyond the sentence.* Chichester: Wiley.

de Vega, M., Carreiras, M., Gutiérrez-Calvo, M. & Alonso-Quecuty, M.L.(1990). *Lectura y comprension. Una perspectiva cognitiva.* Madrid: Alianza Editorial.

This Page Intentionally Left Blank

Part 4

Imagery and Memory

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 12

Imaging objects, routines, and locations

Martin A. Conway, Hank Kahney[*]*, Kathryn Bruce, and Helen Duce*
University of Lancaster & [*]The Open University, United Kingdom

INTRODUCTION

A common view of the nature of conceptual knowledge is that such knowledge is *decontextualised* and *abstract*. Recent research, however, suggests that conceptual knowledge may be closely associated with contextually specific knowledge such as generic knowledge, autobiographical memories, and person-specific factual knowledge (Conway, 1987, 1990a, 1990b). For instance, Conway (1990b) found that goal-derived concepts such as *birthday present* speeded autobiographical memory retrieval to related cues (e.g. *jumper*). In contrast, other types of concepts such as taxonomic categories (e.g. *clothing*), which were also directly related to the memory cues, did not speed retrieval. Conway (1990b) argued that goal-derived concepts are closely associated in memory with limited numbers of discrete event knowledge structures and this association mediates fast access to memories of specific events, resulting in the priming effect described above. Taxonomic categories on the other hand are associated with a multiplicity of event knowledge structures and these diverse associations do not facilitate direct access to specific memories of experienced events.

Thus, the study of priming has proved one useful way in which to explore associations between concepts and different types of knowledge (see also, Conway & Bekerian, 1987). Another approach which we have been developing involves the use of an image generation task. Subjects are presented with words or phrases naming different types of concepts and required to bring to mind an image of whatever they take that word or phrase to refer to. Image generation times (IGTs) are recorded and subjects make a number of judgments of their images. In particular, subjects judged whether their image constituted a specific, datable, *autobiographical memory*, a *generic autobiographical memory* (that is an image directly based on personal experience or experiences but one which is not localisable to a specific time or event), or a *semantic image*. Semantic images comprise knowledge that is not *in principle* localisable to any specific experienced event or time. For example, an image of George Washington confessing to having chopped down a fruit tree might be brought to mind in response to the abstract concept *honesty* (see Conway, 1990a) and this would be classed as a semantic image. Using this technique it is possible to examine the types of knowledge which spontaneously come to mind when people process different types of concepts. Thus, it is possible to map out the dominant type of knowledge drawn upon when people instantiate a concept in the form of an image.

Conway (1990a) used this image generation task to explore the knowledge-base associated with a range of different concepts. The main findings were that concepts naming emotions, e.g. *love, anger, sadness*, and goal-derived concepts, e.g. *things to do at the seaside, places to go on holiday, birthday presents*, were dominated by the spontaneous retrieval of autobiographical memories. Very few images to these concepts were described as generic or semantic images. Other types of concepts such as taxonomic categories and abstract concepts, e.g. *honesty, democracy, hour*, were described with equal frequency as autobiographical memories, generic images, and semantic images. Finally, highly self-relevant concepts denoting personality traits (i.e. *charming, artistic, aggressive*) were dominated by generic images. These findings suggest that emotions and goal-derived concepts are closely associated with autobiographical knowledge, personality traits with generic knowledge, and taxonomic and abstract concepts with a heterogeneous knowledge-base.

The purpose of the present study was to extend the image generation method to a new range of concepts and, in particular, to select concepts which other research has suggested might be selectively associated with autobiographical knowledge. Experiment 1, then, examines the knowledge drawn upon when people image common routines such as *making a cup of tea*, and Experiment 2 examines images of common locations such as *an airport*. In both these experiments taxonomic concepts (e.g. *furniture, clothing, trees*) are employed as baseline concepts and are expected to show a heterogeneous knowledge base in terms of image content. Routines and locations are, however, expected to be dominated by autobiographical knowledge, (Reiser, Black, & Abelson, 1985; Barsalou 1988).

One further manipulation employed in the present experiments is that the level of abstraction of the to-be-imaged concepts is systematically varied. In Experiment 1 three levels of abstraction are employed and concepts comprise equally superordinates (e.g. *Furniture, Eating at a restaurant*), basic level concepts (e.g. *Chair, Ordering a meal,* and subordinates (e.g. *Kitchen Chair, Eating a starter*). In Experiment 2 the concepts comprise equally basic level concepts, subordinates, and attributes of locations and objects. The reasoning behind this level of abstraction manipulation derives from a finding by Rosch (1975) who established that basic level concepts are most likely to be represented analogically and, therefore, these concepts should lead to fastest image generation times. Of course, the question of interest given present concerns was whether these concepts would result in selective generation of autobiographical memories, generic images, or semantic images.

EXPERIMENT 1

Design

A two factor within subjects design was employed. The first factor was Type of Concept with two levels, Object Concepts and Routine Concepts. The second factor was Level of Abstraction with three levels, Superordinate, Basic, and Subordinate. The dependent measure was Type of Image (TOI) with three levels, autobiographical memory, generic image, and semantic image. In addition to this two further measures were taken, Image Generation Time (IGT) and ratings of image vividness on a 5-point rating scale where '1' denoted a very vivid image and '5' a vague and unclear image. Subjects also self assessed their everyday use of imagery on a 5-point scale where '1' denoted very frequent use of imagery and a '5' denoted no use of imagery.

Subjects

Twenty-one student volunteers were recruited from the undergraduate population of the University of Lancaster. Average age of subjects was 20.6 years and there were eleven females and ten males.

Materials

Stimuli for Routines were taken from norms collected at Lancaster University which were highly similar to those reported by Galambos (1983), e.g. *Making a cup of tea, Doing the laundry, Taking a book out of the library.* The stimuli for object concepts were taken from Conway (1990b), e.g. *Furniture, Fruits, Vehicles.* At each of the three levels of abstraction there were four concepts. Thus each Type of Concept was represented by 12 items and there were 24 items in all. In addition to this eight items representing the conditions in the experiment were employed on practice trials.

Procedure

The experiment was run on an Apple Macintosh microcomputer and all stimuli were presented in upper case in white against a black background. Subjects were tested individually and were instructed to 'bring to mind an image of whatever you take a concept to refer to'. Subjects were told to bring to mind their images as quickly as possible and pressed a response key to indicate that they had an image in mind. Subjects were led to believe that the IGT was the critical variable in the experiment and additional instructions emphasising speed of responding and ensuring that an image was in fact in mind, were provided. When a response was made the concept was wiped from the screen and replaced by a message instructing the subject to complete the various ratings of their image. At this point subjects turned to a response booklet and indicated whether the image had been a semantic image, generic image, or autobiographical memory. Subjects also rated memory vividness. For these judgments subjects followed the procedures described in Conway (1990a). After completing the ratings the subject then turned back to the screen and pressed the space for the next trial. When no image came to mind within ten seconds the stimulus for that trial was wiped from the screen and replaced by a message informing subjects that they had been timed-out. In this case subjects marked the corresponding section of the response booklet with Xs.

Results and Discussion

In the first phase of the analysis the data for frequency of TOI, IGT, and vividness were analysed separately. For each variable a 2 X 3 X 3 within subjects analysis of variance was conducted. The first variable was type of concept with two levels, routines and locations. The second variable was level of abstraction with three levels, superordinate, basic, and subordinate. Finally, the third variable was type of image also with three levels, autobiographical memory, generic image, and semantic image. Note that, subjects failed to generate images on less than 1.6% of trials. In addition to this the design was inherently unbalanced because a subject could only generate one type of image to each item. For the purposes of analysis the missing cases where subjects had failed to generate any image, and missing cases arising due to the nature of the design (these latter missing cases only apply to the variables IGT and vividness) were replaced as in Conway (1990a).

A main effect of frequency of TOI was observed, $F(2,40)=5.144$, $p<0.01$. Forty-one percent of images were generic images, 34% were autobiographical memories, and 24% were semantic images. Post-hoc t-tests found that the frequencies of autobiographical memories and generic images did not differ reliably whereas frequency of semantic images was significantly less than either of the other two. A significant interaction of TOI with Type of Concept was also observed, $F(2,40)=5.568$, $p<0.01$, and the means for this interaction are shown in Figure 1.

It can be seen from Figure 1 that the interaction arises because object concepts are dominated by semantic and generic images whereas concepts naming routines are dominated by autobiographical memories and generic images. This suggests that concepts denoting common routine action sequences are closely associated in memory with records of experienced events. In contrast, concepts referring to common objects are associated with generic experiential knowledge, rather than memories of specific experiences, and with decontextualised semantic knowledge.

No significant effects of IGT were observed and it can be concluded that different types of knowledge are equally available for different classes of concepts at different levels of abstraction. Moreover, the basic level was not found to lead to fastest IGTs suggesting that analogical knowledge is equally available at different levels of concept specificity.

In the analysis of vividness ratings a reliable effect of TOI was observed, ($F(2,38(2))=19.9$, $p<0.01$) and mean vividness ratings were, 2.06 for autobiographical memories, 2.42 for generic images, and 3.26 for semantic images. (Numbers within parentheses for the degrees of freedom indicate the adjusted DF for missing cases). Post-hoc tests found that autobiographical memories and generic images did not differ in vividness and images based on either type of knowledge were highly vivid. Semantic images were, however, reliably less vivid than the other two types of images and generally were rated as being only moderately vivid.

Finally, a series of multiple regressions were conducted. For these analyses a regression model was developed in which subjects were the first factor to be entered, followed by type of concept, level of abstraction, and TOI. After entering these main effects the various interactions of the factors and their levels were then entered into the model. With each change in the model the percentage of additional variance accounted for was computed. The variates IGT and memory vividness were analysed separately with this model. The total IGT variance accounted for was 21.5%, and subjects accounted for 20.8% of IGT variance, ($F(20,231)=4.34$, $p<0.01$). No other effects were significant.

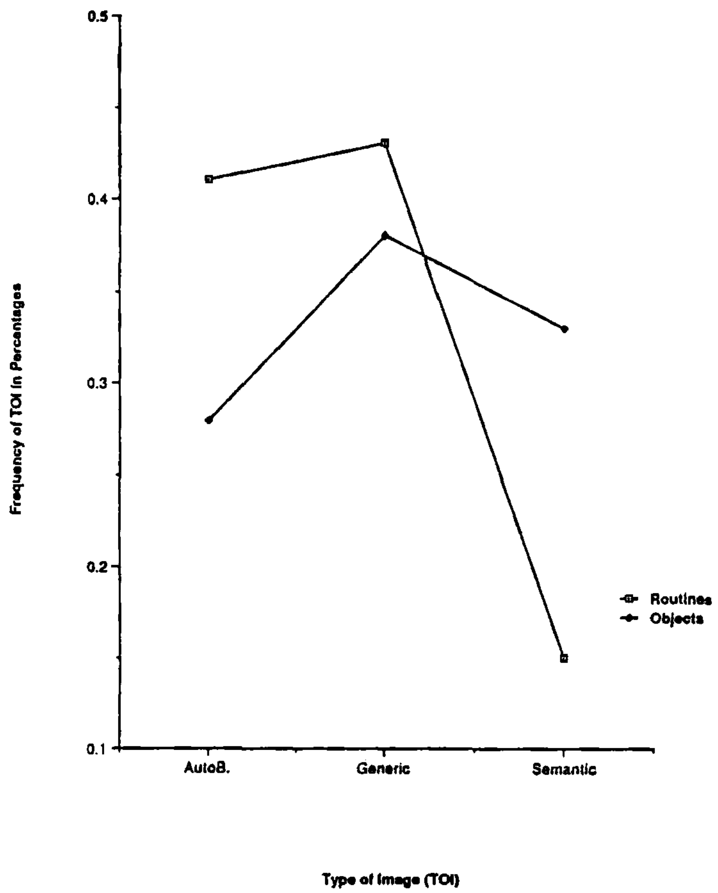

Figure 1. Frequency of type of image (TOI) for routines and objects.

For image vividness, however, subjects only accounted for 4.6% for the variance, $F(1,251)=12.02$, $p<0.01$, whereas TOI accounted for 23.5% of the variance, $F(1,251)=78.54$, $p<0.01$. The complete regression model accounted for 31.5% of the variance in image vividness ratings, $F(4,248)=25.5$, $p<0.01$. Thus, some of the interactions accounted for small and nonsignificant proportions of the vividness variance. Further analyses found the correlations between vividness and IGT to be small and unreliable.

The findings from this experiment indicate that concepts denoting routines may be closely associated with autobiographical knowledge whereas concepts denoting taxonomic categories appear to be equally associated with generic autobiographical knowledge and semantic knowledge. And this is the case regardless of the level of abstraction of concepts. All three types of knowledge are equally available to the image generation process and no differences in IGTs were noted. Finally, images based on autobiographical knowledge were rated as being reliably more vivid than images based on semantic knowledge.

EXPERIMENT 2

Experiment 2 was the same as Experiment 1 with the following modifications. The two concept types were Locations and Objects. The three levels of abstraction were basic level, subordinate, and attributes of the objects and locations. The stimuli for locations were taken from Tversky and Hemenway (1983). This slight alteration in the levels of abstraction was necessitated by the fact that there are apparently only two superordinates for locations: *outside and inside*. Indeed this was one of the main findings of Tversky & Hemenway who examined attribute overlap between location concepts at different levels of abstraction. In their study outside and inside were found to be the main superordinates for locations, whereas other locations such as *lake, mountain,* and *park* were found to be at the basic level. Thus, basic level location concepts were found to share many attributes in common whereas items at the superordinate level were disjunctive in their attributes. In the present case two superordinates did not provide sufficient stimuli for the experiment and, in order to obtain a larger pool of stimuli the levels of abstraction were shifted to basic, subordinate, and attribute as opposed to the superordinate, basic, and subordinate levels used in Experiment 1.

Eighteen subjects took part, all unpaid volunteers from the undergraduate population of Lancaster University. There were 12 women and six men with a mean age of 20.9 years.

Results & Discussion

The analyses were the same as those performed for Experiment 1. For frequency of TOI the only significant effect was that of type of image, ($F(2,34)=6.03$, $p<0.01$), and more generic images (43%) and autobiographical memories (35%) were generated than semantic images (20%). Analysis by simple main effects found that this effect was present for both Locations, ($F(2,34)=3.84$, $p<0.035$), and Objects, ($F(2,34)=5.55$, $p<0.01$), although Objects were associated with slightly more generic images than Locations, 45% vs. 41%, and fewer autobiographical memories, 33% vs. 37%. These findings suggest that locations, like routines, are primarily associated with autobiographical knowledge.

For IGT a reliable effect of TOI was observed, ($F(2,32(2))=3.87$, $p<0.03$), and autobiographical memories and generic images were generated more quickly that semantic images, mean IGTs were 1321ms, 1382ms, and 1773ms, respectively. Also present was a main effect of level of abstraction, ($F(2,34)=4.38$, $p<0.02$), and an interaction of level of abstraction with type of concept, ($F(2,33(1))=3.65$, $p<0.04$), and Figure 2 shows the mean IGTs for this interaction. It can be seen from Figure 2 that for objects IGTs gradually lengthened as the level of abstraction becomes more specific, so fastest IGTs are present for basic level object concepts (e.g. *Apple, Car*) and slowest for attributes of concepts (e.g. *Core with pips, Steering wheel*). In contrast, basic level locations (e.g. *a park, an airport*) and attributes of locations (e.g. *a bench, a duty-free shop*) give rise to equally fast IGTs whereas subordinate locations (e.g. *a wild-life park, Heathrow airport*) lead to reliably slower retrieval times. In general, these findings support Rosch's (1975) view that for objects the basic level is the level at which concepts can be most easily imaged. Locations, unlike objects, do not show this basic level bias.

In the analysis of memory vividness the only reliable effect was that of TOI, ($F(2,32(2))=33.23$, $p<0.01$,) and autobiographical memories and generic images did not differ reliably in rated vividness and both types of image were judged to be highly vivid. The means were 1.9 and 2.4 respectively. Semantic images were, however, reliably less vivid with a mean vividness rating of 3.33, indicating that these images were judged to be only moderately vivid.

In the regression analyses Subjects accounted for 2.3% of IGT variance, ($F(1,227)=5.26$, $p<0.01$), concepts 1.4%, levels 1.7%, ($F(1,227)=3.964$, $p<0.05$), and TOI 1.2%. Overall, then less than 7% of the variance in IGT was accounted for by the factors in the design, although this was significant, ($F(4,244)=3.62$, $p<0.01$). For ratings of vividness the total variance accounted for was 32.2%, ($F(4,224)=26.6$, $p<0.01$). However, TOI alone accounted for 31.6%, ($F(1,227)=104.9$, $p<0.01$), of the total vividness variance and this means that all the other factors in the design accounted for less than 1% of the vividness variance.

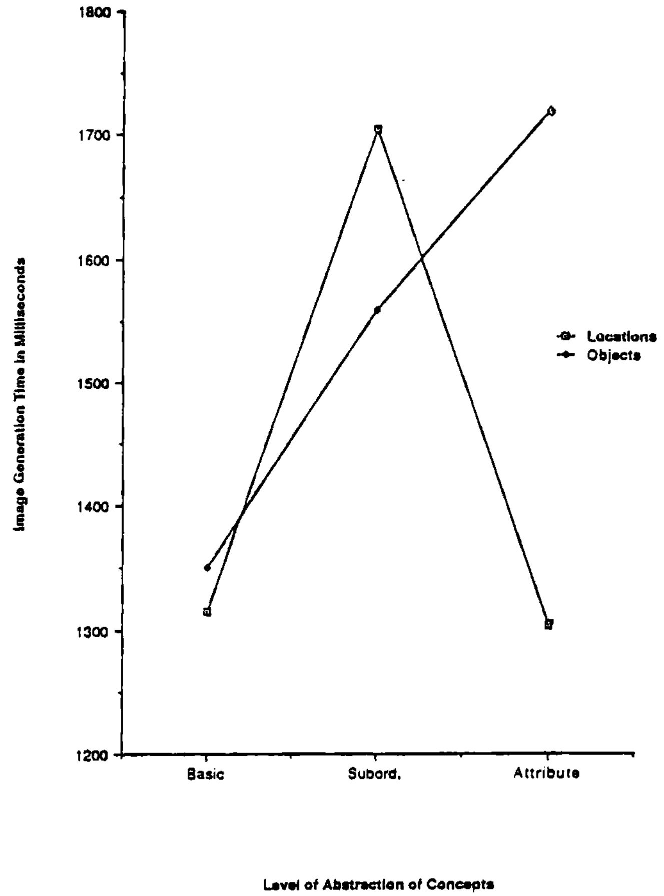

Figure 2. Image generation times (IGTs) for locations and objects at different levels of abstraction.

DISCUSSION

Concepts denoting common routines (e.g. *making a cup of tea*) and locations (e.g. *an airport*) were found to be dominated by knowledge of specific experiences. These results lend support to the findings of previous autobiographical memory research which have established that action sequences (Reiser, Black, & Abelson, 1985) and locations (Barsalou, 1988) can serve as effective cues for autobiographical memory retrieval (see Conway, 1990c for a review).

For taxonomic concepts, however, the findings were somewhat mixed. In Experiment 1 taxonomic concepts were reliably associated with generic and semantic images. In Experiment 2 taxonomic concepts were reliably associated with autobiographical and generic knowledge. Also in Experiment 1 no reliable differences were found between the different levels of concept abstraction and IGT, but in Experiment 2 basic level taxonomic concepts were imaged more quickly than more specific taxonomic concepts. The main difference between the two experiments was that Experiment 1 contained superordinate taxonomic categories whereas in Experiment 2 these were replaced by attributes of objects. Thus, the findings indicate that attributes are selectively associated with autobiographical knowledge rather than a range of knowledge types and hence the TOI differences between the experiments. Similarly, the highly specific nature of attributes of objects may require longer searches of memory prior to locating knowledge which can be used for image generation and hence the IGT advantage to basic level concepts observed in Experiment 2 only. Clearly, the IGT findings of Experiment 1 indicate that superordinate taxonomic categories are as imageable as basic level taxonomic categories and this suggests that high imagability is not a unique feature of basic level concepts as Rosch (1975) originally suggested. Despite these differences between the two experiments the findings are highly compatible with previous research (Conway, 1990a) that used imagery to explore conceptual knowledge and, overall, it can be concluded that taxonomic categories are associated with a heterogeneous knowledge-base. In contrast, other types of concepts, routines, locations, goal-derived categories, and emotions, have been found to be selectively associated with autobiographical knowledge.

One particularly striking aspect of the present findings was that in both experiments autobiographical memories and generic images were judged to be reliably higher in vividness than semantic images. Moreover, TOI accounted for a large and significant proportion of the vividness rating variance. Why should images based directly on specific experienced events be more vivid than images which do not have such a direct association with memories of experienced events? One answer is provided by the work of Johnson, Foley, Suengas, and Raye (1988). These authors found that memories of events which had been experienced were strongly associated with perceptual knowledge. In contrast, memories for events which had been imagined were only weakly, if at all, associated with perceptual knowledge. To the extent that semantic knowledge structures represent decontextualised knowledge which has been abstracted from experienced events (see also Kosslyn, 1980) then semantic knowledge may be similar to memories for imagined events and specific perceptual details arising from experience may not be preserved in such representations. In this way then, images generated from semantic knowledge may be less vivid than images generated from autobiographical knowledge.

Finally, the limitations of using imagery to probe conceptual knowledge will be briefly considered. The present findings and those of Conway (1990a) suggest that imagery may represent a useful technique with which to explore conceptual knowledge. The knowledge-base accessed in generating an image gives important clues about the nature of the underlying knowledge with which a concept is associated. However, there are two major problems which limit the inferences which can be drawn of the basis of such findings. First, it may be that specific types of knowledge are only accessed when a person generates an image and when the same concept is processed in some other, non-imaged, way then different types of knowledge are accessed. One obvious solution to this problem is to examine the types of knowledge which are accessed in different processing contexts, which was one of the aims of Conway's (1990b) autobiographical memory priming studies. Thus, imaging concepts might be thought of as one method to adopt when a convergent operations approach is taken to investigate conceptual knowledge. Second, it seems highly unlikely that all, or even most, types of conceptual knowledge can be investigated by the use of imagery. For example, conceptual relations such as a *robin is a bird,* mathematical operations such as addition, and processing of procedural knowledge, seem unlikely to be amenable to investigation by use of imagery. Nevertheless, the present study suggests that when the scope of the investigation is limited to declarative knowledge concerning common event, action, and object concepts then imagery may provide a useful way in which to map relations between concepts and different types of knowledge.

REFERENCES

Barsalou, L.W. (1988). The content and organization of autobiographical memories. *In* U. Neisser & E. Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory.* New York: Cambridge University Press, pp 193-243.

Conway, M.A. (1987). Verifying autobiographical facts. *Cognition* **25**, 39-58.

Conway, M.A. (1990a). Conceptual representation of emotions: The role of autobiographical memories. *In* K.J. Gilhooly, M.T.G. Keane, R.H. Logie & G. Erdos (Eds.), *Lines of Thinking* **Vol. 2**, 133-143.

Conway, M.A. (1990b). Associations between autobiographical memories and conceptual knowledge. *Journal of Experimental Psychology: Learning, Memory and Cognition* **16**, 799-812.

Conway, M.A. (1990c). *Autobiographical memory: An introduction.* Open University Press: Buckingham, England.

Conway, M.A., & Bekerian, D.A. (1987). Organization in autobiographical memory. *Memory & Cognition* **15**, 119-132.

Galambos, J.A. (1983). Normative studies of six characteristics of our knowledge of common activities. *Behaviour Research Methods & Instrumentation* **15**, 327-340.

Johnson, M.K., Foley, M.A., Suengas, A.G., & Raye, C.L.(1988). Phenomenal characteristics of memories for perceived and imagined autobiographical events. *Journal of Experimental Psychology: General* **117**, 371-376.

Kosslyn, S.M. (1980). *Image and mind.* Cambridge, MA: Harvard University Press.

Reiser, B.J., Black, J.B., & Abelson, R.P. (1985). Knowledge structures in the organization and retrieval of autobiographical memories. *Cognitive Psychology* **17**, 89-137.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General* **104**, 192-233.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology* **15**, 121-149.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 13

Bizarre imagery: Mnemonic benefits and theoretical implications

Mark A. McDaniel
Purdue University
and
Gilles O. Einstein
Furman University

One issue in studying imagery and memory concerns how the quality of the image may influence its mnemonic properties. One qualitative factor that has received attention since the ancient Greeks is image bizarreness, the standard suggestion being that bizarre images improve memory (Lorayne & Lucas, 1974). The early research on the mnemonic effects of bizarre imagery, however, showed that unusual or bizarre images were no better recalled than common or normal images (e.g., see Cox & Wollen, 1981). Several later studies, however, showed positive benefits of bizarre imagery on memory. In reviewing the literature, Einstein and McDaniel (1987) noted that the type of design seemed to influence whether or not bizarre imagery effects were found. Specifically, between-subjects (unmixed list) manipulations of imagery type did not produce bizarre imagery effects (e.g., Collyer, Jonides, & Bevan, 1972; Cox & Wollen, 1981), whereas within-subjects (mixed list) manipulations of imagery type did lead to enhanced free recall for bizarre images (Pra Baldi, de Beni, Cornoldi, & Cavedon, 1985; Wollen & Cox, 1981). McDaniel and Einstein (1986) provided support for this analysis by directly comparing mixed and unmixed list manipulations of the bizarreness variable. Bizarre imagery produced higher recall than common imagery with the mixed list but not the unmixed lists.

Two other factors that are apparently important for obtaining bizarre imagery effects are the processing task and the test task. In terms of the processing task, subjects usually are presented with verbal material (e.g., sentences) that either specify a bizarre event or a common event, and they are asked to image the referents of the sentence. Implicit in this procedure is the assumption that imagery is involved in the bizarreness effect (when found). Recent studies have tested this assumption by having subjects perform either semantic or imaginal processing on a mixed list of bizarre and common sentences (McDaniel, Anderson, Einstein, & O'Halloran, 1989, Exps. 1-3; McDaniel & Einstein, 1986, Exp. 2). The results consistently showed that the bizarreness effect is obtained with imagery instructions but not with semantic instructions (for similar results also see Cornoldi, Cavedon, de Beni, & Pra Baldi, 1988). It seems, then, that imagery instructions are essential for obtaining the bizarreness effect. With regard to the test task, the bizarreness effect is obtained with free recall but not with cued recall or recognition (See Einstein & McDaniel, 1987, for a detailed review).

We (Einstein & McDaniel, 1987; McDaniel & Einstein, 1986) used a distinctiveness hypothesis to explain why the type of design is instrumental for observing bizarre imagery effects. (In a subsequent section we address how the distinctiveness view relates to the effect being limited to free recall.) According to most accounts of distinctiveness, the distinctiveness of an event is determined by two factors. First, a distinctive item is one that shares few features in common with other information in memory. Because bizarre images contain distorted relationships among objects, they satisfy this criterion of distinctiveness. Second, distinctiveness is typically defined in a relative sense--that is, relative to other information in the learning context. According to this criterion, items become functionally distinct only when they are presented in the context of common items. Assuming that distinctiveness is positively related to recall, then this view nicely describes our results. Bizarre imagery effects are found only when bizarre items appear in the context of common material (i.e., mixed lists).

DISTINCTIVENESS OR AFFECTIVE ELABORATION?

An alternative to the distinctiveness hypothesis is that bizarre items cause subjects to have some sort of startle response or emotional reaction, and these emotional reactions serve as important cues for retrieval (Hirshman, Whelley & Palij, 1989). Moreover, Hirshman et al. invoke cue overload theory (e.g., Watkins & Watkins, 1975) to argue that the emotional cues become less effective with increasing numbers of bizarre items. There are two implications of this perspective. First, mixed lists may not be necessary for obtaining bizarre imagery effects. Second, the determining factor for obtaining bizarre imagery effects should be the number of bizarre items appearing in the list. Bizarre imagery effects should be obtained as long as there are not too many bizarre items in a list. With too many bizarre images, the emotional response cue is rendered ineffective.

By contrast, according to the distinctiveness framework outlined earlier, the determining factor for obtaining bizarre imagery effects is having both bizarre and common events in the learning context. Previous experiments fail to resolve this issue, because all of the ones that have compared mixed with unmixed lists have confounded mixing with the number of bizarre items in the list. For example, in our previous research, our mixed lists always contained half the number of bizarre items as our unmixed lists.

To test these competing predictions, we (with Natalie Coté) factorially manipulated the type of list (mixed vs. unmixed) and the number of bizarre items in the study list. The experiment consisted of 6 groups, with 12 subjects in each group. Two groups of subjects received mixed lists; one group was presented with 6 bizarre and 6 common sentences, and the other was presented with 12 bizarre and 12 common sentences. There were also 4 groups receiving unmixed lists. One received 6 bizarre sentences, one received 6 common sentences, one received 12 bizarre sentences, and one received 12 common sentences.

For all subjects, the target list consisted of noun triplets embedded within sentences. The noun triplets were presented in capital letters in each sentence, and type of imagery was manipulated by varying the relationship among the nouns. A bizarre and common sentence was created for each of the 24 noun triplets. For example, the bizarre context for the items DOG, BICYCLE, and STREET was the "The DOG rode the BICYCLE down the STREET" and the common context was "The DOG chased the BICYCLE down the STREET." Noun triplets were counterbalanced across conditions, such that each triplet was presented equally often in each condition.

Each sentence was presented with a slide projector for 10 seconds. For each sentence, subjects were told to form an interactive image and to rate the vividness of their image on a 5-point scale. After imaging and rating all of the sentences in their list, subjects were given word problems to solve for 12 minutes. Finally, subjects were given as much time as they needed to recall the capitalized nouns from the sentences.

For all analyses, the alpha level was set at .05. Also, we used Erlebacher's (1978) method for contrasting the mixed- and unmixed-list effects of bizarre imagery. Generally, subjects recalled a greater proportion of target nouns from short lists than long lists. More importantly, this analysis produced a reliable interaction between the type of list and type of image. As can be seen in Table 1, items from bizarre sentences were better recalled than items from common sentences in mixed lists ($F(1,64) = 7.95$) but not in unmixed lists ($F < 1$). As shown in the table, list length did not affect the relative recall levels of bizarre and common images. Regardless of list length, the bizarre imagery effect occurred with a mixed list but not with an unmixed list.

Table 1. Results from experiment contrasting mixed versus unmixed list manipulations of imagery type

		LIST TYPE			
	Number of each sentence type	Mixed		Unmixed	
Measure		Bizarre	Common	Bizarre	Common
Proportion of nouns recalled	6	.48	.37	.62	.62
	12	.38	.22	.40	.39
Sentence Access[a]	6	.60	.43	.72	.68
	12	.48	.26	.56	.50
Items per Sentence[a]	6	2.36	2.47	2.51	2.62
	12	2.33	2.40	2.13	2.34

[a]See text for full explanation

Next, we analyzed the proportion of sentences recalled. A sentence was scored as recalled if at least one word from the sentence was recalled. This measure, then, reflects subjects' ability to access the imaged sentences. As shown in Table 1, and consistent with our previous research, the results of this measure parallel those of the overall recall measure. Most importantly, bizarre sentences were better accessed than common sentences in the mixed lists and not the unmixed lists. This pattern was evident regardless of the number of bizarre items in the list. To get a measure of subjects' ability to recover the constituents of the sentences, we also analyzed recall for the number of words recalled per sentence, given that a sentence was recalled. As can be seen in Table 1, recall levels for bizarre and common sentences were approximately the same, regardless of list length and whether a mixed or unmixed list was used.

In sum, the results do not support the emotional cue overload explanation of bizarre-imagery effects, which implies that the number of bizarre images in the list should be the critical factor affecting the relative recall levels of bizarre and common images. The results seem clear in supporting the claim that a mixed list is necessary for obtaining bizarre imagery effects. As such, the results are consistent with a distinctiveness framework outlined earlier, wherein bizarre items are not functionally distinctive unless presented in the context of common items.

The distinctiveness framework presented thus far is purely descriptive in the sense that it describes the necessary conditions for obtaining the bizarreness effect. It does not explicate the processes underlying bizarre-imagery effects. A potentially fruitful approach for understanding bizarre-imagery effects is to assume that bizarre images are encoded differently in mixed and unmixed lists. This was the approach taken by Hunt and Elliott (1980) in explaining the positive effects of distinctive orthography in mixed but not unmixed lists. Briefly, Hunt and Elliott argued that the distinctive features of orthographically distinctive words are encoded only when orthographically common items are presented. Their view is that subjects do not perceive unique letter patterns when all the items in the list are orthographically distinct. Similarly, it may be that the bizarre or distinctive qualities of bizarre images are better encoded and hence more useful in mixed lists.

Support for this assumption was provided by Cindi May, a student of G. Einstein, in a replication of the present experiment that required subjects to make bizarreness ratings. Bizarre sentences in mixed lists were perceived to be more bizarre than bizarre sentences in unmixed lists, and this interaction held regardless of list length. These results indicate that bizarre sentences are encoded differently in mixed lists relative to unmixed lists, and are consistent with Hunt and Elliott's view that mixing affects the degree to which distinctive qualities of events are encoded.

CHALLENGING THE DISTINCTIVENESS FRAMEWORK

Encoding of distinctive qualities of the bizarre events when they are mixed with common events may not be the whole story, however. The literature reports consistent failures to obtain a mnemonic advantage of bizarreness when the target sentences in a mixed-list design are made more complex by adding several adjectives before each of the target nouns (Kroll, Schepeler, & Angin, 1986; McDaniel & Einstein, 1989; Richman, Dunn, Kahl, Sadler, & Simmons, 1990). For instance, the bizarreness effect is obtained when simple sentences like "The BANKER dropped the NEWSPAPER into the PUDDLE" are imaged (in a mixed-list design), but not when complex sentences like "The tired, crabby BANKER dropped the torn, crumpled NEWSPAPER into the muddy PUDDLE" are imaged.

Richman et al (1990) accounted for this pattern by maintaining a focus on the influence of the encoding context on the degree to which bizarre events are distinctively encoded. "We assume that the distinctive encoding of a sentence requires a limited number of words per sentence. Increasing the number of words per sentence beyond this point encourages subjects to break the sentence down into phrases. If these phrases are logical (the noun modifiers make sense), they are considered by the subject to be common and nondistinct." (p187). Within this view, then, the extra modifiers lead subjects to encode both bizarre and common sentences as sets of common phrases, thereby functionally equating the bizarreness levels of the two types of sentences that nominally differ.

Instead of assuming that adding adjectives disrupts distinctive encoding, an alternative idea is that the additional cues (elaborative adjectives) present in the complex sentences override the retrieval advantage conferred by the distinctiveness of the bizarre sentences (cf. McDaniel & Einstein, 1989). There is evidence that increasing the complexity of a sentence frame in which a target word is embedded produces increased recall levels for the target (Craik & Tulving, 1975). This result, in concert with the bizarre imagery results, suggests that either elaboration or distinctiveness can aid retrieval, with bizarreness adding distinctiveness and sentence complexity entailing elaboration. On this view it should be possible to obtain bizarre imagery effects with complex sentences by reducing the potency of the adjectives as retrieval aids. One way to reduce the usefulness of a cue for retrieval is to associate that cue with many target items (Watkins & Watkins, 1975). Bridget Robinson, a student of M. McDaniel, implemented just this situation in a recently completed study. She selected a small pool of adjectives, and used these adjectives repeatedly throughout the list so that each adjective was paired with several target nouns (adjectives were paired with nouns in a way that made sense). We reasoned that repeated use of the same adjectives across sentences should reduce the effectiveness of these modifiers as memory cues (due to cue overload). Subjects may then be forced to exploit event distinctiveness as a retrieval aid, and consequently bizarre imagery effects should emerge with these complex sentences.

Note that in modifying the sentences, Robinson did not reduce the number of words per sentence, and she tried to maintain the degree to which the modifiers associated with each noun made sense. Thus, by the first view sketched above, the bizarre sentences should be processed in a phrase by phrase fashion, with the result that they would be encoded in a common, nondistinct fashion. Accordingly, on this view the bizarreness effect should not be reinstated.

In this study 16 subjects processed a mixed list of complex sentences (6 bizarre and 6 common) in which the adjectives (48 in all) appeared once across the list (adopted from Kroll et al, 1986), and 16 other subjects processed a mixed list in which the same 12 adjectives were used repeatedly throughout the list (these two groups were tested at different times, accordingly their data are not combined into one analysis). Subjects were instructed to form a mental image of the event suggested by each sentence. After forming their image, subjects were asked to rate the vividness of the image. Sentence presentation was subject paced, and no mention was made of a subsequent memory test. After completing the imagery task for the sentence list, subjects worked on a mental rotation task instrument, and then they attempted to free recall the nouns in the imaged sentences.

The results are presented in Table 2. Bizarre images took longer than common images to form ($F(1,15)=4.48$ and $F(1,15)=9.91$ for the standard and repeated-adjective lists, respectively), and bizarre images were rated as less vivid than common images ($F(1,15)=11.67$ and $F(1,15)=96.84$ for the standard and repeated-adjective lists). For the proportion of nouns recalled, items from bizarre sentences were better recalled than items from common sentences from the repeated-adjective list ($F(1,15) = 5.09$) but not for the standard list ($F < 1$). Moreover, the bizarre imagery effect for the repeated-adjective list reflected the same components observed in all of our previous work. The bizarre sentences were better accessed than common sentences (.58 vs. .39, respectively; $F(1,15)=14.10$), whereas there was no significant difference across bizarre and common sentences in the number of words recalled per sentence (1.41 vs. 1.53, respectively). In short the results supported the view that bizarreness may not aid recall for complex sentences because the additional elaboration provided by the added adjectives may override the use of distinctive information at retrieval.

Table 2. Processing times, vividness ratings, and proportion of target nouns recalled in the Robinson Study (Complex Sentence Materials)

		List Type[a]	
Measure	Sentence Type	Standard	Repeated-Adjectives
Recall	Bizarre	.44	.42
	Common	.38	.31
Vividness Ratings[b]	Bizarre	2.94	3.32
	Common	2.05	1.63
Processing Time[c]	Bizarre	9.18	8.17
	Common	8.32	7.02

[a]See text for full explanation
[b]Lower ratings indicate increased vividness
[c]Average per sentence in seconds

CONCLUSIONS

The bizarreness of an image does appear to have mnemonic value, but in a restricted sense. First, the encoding context has to be such that the bizarre event is distinctively encoded. This occurs optimally when the event is presented in the context of more common events and when the event is encoded through visual imagery. Secondly, the distinctiveness of the encoding needs to be useful at retrieval. Explicit retrieval cues as provided in cued recall and recognition seem to better serve retrieval than do distinctive qualities of bizarre images. Even in free recall, retrieval routes afforded by elaborative encoding of the target (cf. Anderson & Reder, 1979) can override the use of distinctive qualities of a bizarre image, as evidenced by the lack of bizarre imagery effects with complex sentences. That these distinctive qualities are present and potentially useful, however, is evident by the manifestation of the bizarreness effect for complex sentences in which the elaboration has been rendered functionally less effective due to cue overload.

REFERENCES

Anderson, J.R., & Reder, L.M. (1979). An elaborative processing explanation of depth of processing. *In* L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 385-403). Hillsdale, NJ: Erlbaum.

Collyer, S.C., Jonides, J., & Bevan, W. (1972). Images as memory aids: Is bizarreness helpful? *American Journal of Psychology* **85**, 31-38.

Cornoldi, C., Cavedon, A., de Beni, R., & Pra Baldi, A. (1988). The influence of the nature of material and of mental operations on the occurrence of the bizarreness effect. *The Quarterly Journal of Experimental Psychology* **40**, 73-85.

Cox, S.D., & Wollen, K.A. (1981). Bizarreness and recall. *Bulletin of the Psychonomic Society* **18**, 244-245.

Craik, F.I.M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General* **104**, 268-294.

Einstein, G.O., & McDaniel, M.A. (1987). Distinctiveness and the mnemonic benefits of bizarre imagery. *In* M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp. 78-102). New York: Springer-Verlag.

Erlebacher, A. (1978). The analysis of multifactor experiments designed to contrast the within- and between-subjects manipulation of the independent variables. *Behavior Research Methods and Instrumentation* **10**, 833-840.

Hirshman, E., Whelley, M.M., & Palij, M. (1989). An investigation of paradoxical memory effects. *Journal of Memory and Language* **28**, 594-609.

Hunt, R.R., & Elliott, J.M. (1980). The role of nonsemantic information in memory: Orthographic distinctiveness effects on retention. *Journal of Experimental Psychology: General* **109**, 49-74.

Kroll, N., Schepeler, E.M., & Angin, K.T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 42-54.

Lorayne, H., & Lucas, J. (1974). *The memory book*. New York: Stein & Day.

McDaniel, M.A., Anderson, D.C., Einstein, G.O., & O'Halloran, C.M. (1989). Modulation of environmental reinstatement effects through encoding strategies. *American Journal of Psychology*, 102, 523-548.

McDaniel, M.A., & Einstein, G.O. (1986). Bizarre imagery as an effective mnemonic aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 54-65.

McDaniel, M.A., & Einstein, G.O. (1989). Sentence complexity eliminates the mnemonic advantage of bizarre imagery. *Bulletin of the Psychonomic Society* **27**, 117-120.

Pra Baldi, A., de Beni, R., Cornoldi, C., & Cavedon, A. (1985). Some conditions for the occurrence of the bizarreness effect in free recall. *British Journal of Psychology* **76**, 427-436.

Richman, C.L., Dunn, J., Kahl, G., Sadler, L., & Simmons, K. (1990). The bizarre sentence effect as a function of list length and complexity. *Bulletin of the Psychonomic Society* **28**, 185-187.

Watkins, O.C., & Watkins, M.J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory* **104**, 442-452.

Wollen, K.A., & Cox, S. (1981). Sentence cuing and the effectiveness of bizarre imagery. *Journal of Experimental Psychology: Human Memory and Learning* **104**, 442-452.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 14

Concreteness, imagery, and memory for prose*

Marc Marschark and John Warner
University of North Carolina at Greensboro, USA

Roxann Thompson
Massachussetts Institute of Technology, USA
and
Charles Huffman
University of North Carolina at Greensboro, USA

INTRODUCTION

Among the more consistently reported findings in cognitive psychology are those indicating (a) better memory for high-imagery (concrete) materials than low-imagery (abstract) materials and (b) improvement of memory through the use of visual imagery (see Paivio, 1971, 1986). These effects are not nearly as ubiquitous as we assumed just 5 or 10 years ago, however, and there have now been several demonstrations that it is possible to attenuate or eliminate concreteness effects in virtually all memory paradigms (Marschark & Hunt, 1989; Marschark, Richman, Yuille, & Hunt, 1987; Marschark & Surian, 1989; Wattenmaker & Shoben, 1987). From our perspective, such findings indicate that we do not fully understand the role of concreteness and imagery in memory.

*This research was supported in part by a grant from the United States Department of Education to the UNC-Greensboro Human Development Research Institute and by grants KO4 DC00028 and RO1 NS 27243 from the National Institutes of Health to Marc Marschark. Experiment 2 is part of a larger study by Marschark, Huffman, and Warner. We wish to thank Arnie Glass for his helpful comments with regard to Experiment 1, Tom Trabasso for his structural and conceptual help with regard to Experiment 2, and Johannes Engelkamp and Bob Logie for their extensive comments on an earlier version of this chapter. All correspondence concerning this article should be addressed to Marc Marschark, Department of Psychology, University of North Carolina at Greensboro, Greensboro, North Carolina 27412, U.S.A. E-mail: MARCO@UNCG.BITNET or MARCO@STEFFI.ACC.UNCG.EDU

In our continuing attempts to decipher the locus of concreteness effects in recall of verbal material, memory for prose is one task which presents an especially important and yet confusing picture. Imagery research in this area began after several studies in the late 1960's demonstrated that the gist of coherent paragraphs forms the stuff of memory, whereas exact wording and grammatical structure are quickly lost. Those studies uniformly used relatively concrete paragraphs, however, leading Yuille and Paivio (1969) to suggest that the observed effects might be the result of visual images' preserving meaning in some non-verbatim form.

Yuille and Paivio presented subjects with the words of concrete or abstract paragraphs, one word at a time, in either random or syntactic order. Not only were the concrete paragraphs remembered better than the abstract ones, but the randomization manipulation affected memory for the abstract words more than the concrete words. These findings provided strong support for the emerging dual coding theory wherein concrete materials are thought to be retained in both imaginal and verbal forms whereas abstract materials are thought to be retained only in verbal form akin to the verbatim input strings (Paivio, 1971). Overlooked at the time was Yuille and Paivio's additional finding that there were more intrusions of thematically consistent material in memory for their abstract than concrete paragraphs (e.g., "war" replaced "active struggle towards conquest for power"). This result contradicts Yuille and Paivio's assumption that abstract text is retained in some verbatim form, but it is consistent with other findings indicating some common integrative mechanism for concrete and abstract materials (Marschark & Paivio, 1977).

Yuille and Paivio's (1969) reported results fit in so well with a variety of other imagery-related findings emerging during the early 1970's that it was not until a study by Marschark (1978) that the role of concreteness in memory for prose was examined in detail. Marschark took concreteness effects in text memory as a "given" and focused on possible differences in the reading strategies employed by subjects as a possible contributor to those memory effects. Marked and consistent differences in the processing strategies used for concrete and abstract texts were in fact observed in that study (see Marschark, 1979). Surprisingly, however, no concreteness effects were obtained in recall. Marschark (1985) and Wattenmaker and Shoben (1987) later replicated that null finding and argued that the results could not be explained by any obvious dual memory model. Instead, those investigators opted for explanations emphasizing the contextual or relational aspects of prose processing and began to search for alternative accounts of concreteness effects in other paradigms. Left unresolved, however, was why the later prose studies obtained such different results from Yuille and Paivio (1969) and several subsequent studies employing the same materials (e.g., Ernest, in press).

One possible explanation of the discrepant findings is suggested by a study by Bunn (1986). Using a subset of Marschark's (1978) paragraphs, she found a reliable interaction of concreteness and presentation order, such that the proportion of concrete nouns recalled was significantly smaller than the proportion of abstract nouns recalled (.39 versus .45, respectively) when concrete paragraphs were presented and recalled first (C-A order), whereas a significantly greater proportion of concrete than abstract nouns were recalled (.53 versus .47) when abstract paragraphs were presented and recalled first (A-C order); overall, the concreteness effect was not reliable. A similar interaction also was obtained when gist recall was considered, but tests of the simple effects were not reported.

Considering Bunn's complete design, a reliable concreteness effect was obtained in gist recall when concreteness was evaluated within-subjects. However, when she compared only the first test blocks of subjects in the C-A and A-C presentation conditions (i.e., a between-subjects concrete-abstract comparison), a concreteness effect was not obtained, consistent with Marschark's (1978, 1985) findings with between-subjects designs. Taken together, these studies suggest the possibility that the presence or absence of concreteness effects in prose memory, in general, might depend on whether within- or between-subjects designs are employed, a conclusion that would be consistent with Marschark et al.'s (1987) account of concreteness effects in terms of the distinctiveness of concrete relative to abstract materials in memory (see McDaniel & Einstein, 1986).

"Distinctiveness," as used by Marschark and his colleagues, refers to the relative contrast of any particular thing to other things from which it must be discriminated in a particular context (Jacoby & Craik, 1979). Distinctiveness, and hence memory, thus can be enhanced by more complete descriptions, more meaningful processing of relevant information, or the emphasis of distinguishing features. In a within-subjects manipulation of concreteness, we would expect that the relative distinctiveness of concrete materials would be emphasized through a complex of variables such as imageability, specificity, and familiarity (see Marschark & Cornoldi, 1990, for discussion); and relatively greater memory for concrete materials would be observed. That enhancement typically would not obtain in a between-subjects design, there being no manipulation of a distinguishing variable, and reliable concreteness should be eliminated, all other factors being equal (Marschark & Cornoldi, 1990).

Ransdell and Fischler (1989) recently provided further support for the suggestion that concreteness effects in memory for prose might depend on the design in which the materials are presented. In a study employing a subset of Marschark's (1978) materials, they obtained concreteness effects when within-subjects designs were used (Experiment 1 and 3), but not when a between-subjects design was used (Experiment 2). However, Ransdell and Fischler (1989) did not use the same orienting task in the two experiments intended to contrast the two designs. In their Experiment 2, subjects read the paragraphs for comprehension (as in Marschark's studies) whereas in Experiment 3, they rated the materials for concreteness. To the extent that drawing subjects' attention to the image-evoking qualities of the concrete materials might increase their distinctiveness, Ransdell and Fischler's (within-subjects design) Experiment 3 might have been biased in favor of finding concreteness effects.

In Experiment 1, we examined this subjective contrast effect and explicitly examined the effects of within- and between-subjects comparisons of concrete and abstract paragraphs under identical conditions.

EXPERIMENT 1

Each subject in this experiment read a concrete and an abstract paragraph, in counterbalanced order, and attempted recall after each. If concreteness effects in memory for prose depend on the relative contrast of concrete and abstract paragraphs in within-subjects designs, concreteness effects should not be obtained in (C-A and A-C) subjects' first paragraph recall, but should be obtained in second paragraph recall.

Method

Subjects. Forty undergraduates participated and were tested in small groups. Approximately half received credit toward course research requirements; the remainder were volunteers. Subjects were randomly assigned to the four conditions of the design on the basis of the testing session in which they participated.

Materials. Two concrete and two abstract paragraphs were drawn from those developed by Marschark (1978); paragraphs used by Marschark

(1985), Bunn (1986), and Ransdell and Fischler (1989) were excluded from selection. Marschark's paragraphs comprised pairs of concrete and abstract passages equated for length, comprehensibility, syntactic structure, and conceptual structure; the members of each pair share the same theme and many of the same words but differ in the concreteness of their component content words and, to some extent, their subject matter. Because of the overlap in content and structure, subjects received concrete and abstract paragraphs from different pairs.

Design and Procedure. The complete design was a 2 (imagery or comprehensibility rating) x 2 (C-A or A-C presentation order) x 2 (concrete or abstract paragraphs); only concreteness was within-subjects. Subjects were shown the paragraphs via an overhead transparency for 30 seconds each. They were asked to read the paragraph to themselves and informed that this would be followed by a brief interpolated task, recall, and then a normative rating task. After the first paragraph was shown, they solved multiplication problems for 30 seconds. This was followed by a recall test in which they were requested to reproduce the paragraph as accurately as possible but encouraged to write words, phrases, or ideas where exact wording was not available. Following three minutes of recall, they rated the sentences of the paragraphs using standard imagery or comprehensibility instructions. The procedure was then repeated for the second paragraph.

Results and Discussion

Before subjects were tested, the paragraphs had been divided into idea units; each paragraph contained 11 such units. Gist recall of these idea units was scored independently by two of the authors; disagreements were resolved in conference. The recall data then were analyzed using a 2 x 2 x 2 ANOVA. An alpha level of .05 was adopted for this and the subsequent experiment; all and only those effects reported were reliable at that level. The means for all cells of Experiment 1 are listed in Table 1. To be consistent with the data reported for Experiment 2, they are reported as proportions; raw scores can be obtained by multiplying by 11.

Overall, concrete paragraphs were recalled better than abstract paragraphs, $F(1,36)=4.10$, $MSe=2.39$. However, this was qualified by a reliable task by concreteness interaction, $F(1,36)=4.10$: There was no concreteness effect at all when subjects rated comprehensibility (means=.52), whereas concrete paragraphs were marginally better recalled than abstract paragraphs when subjects rated imageability (means=.52 and .39, respectively, $p=.05$, by a Newman-Keuls test).

Consistent with the findings of Marschark (1985), the first row of Table 1 reveals that the between-subjects comparisons (i.e., first-test data) did not show concreteness effects in either condition. The second-test data (in the second row of Table 1) yielded a reliable effect only in the imagery rating condition. These results subsequently were replicated in an experiment run with the identical procedure but more naturalistic materials (selected from Aldous Huxley's *Heaven and Hell*). Once again, a reliable, overall, within-subjects concreteness effect was obtained, but the first-test, between-subjects comparisons were not reliable.

Table 1. Mean Recall for Concrete (C) and Abstract (A) with C-A and A-C presentation orders: Experiment 1

	Orienting Task			
	Comprehensibility		Imagery	
Presentation Order	C-A	A-C	C-A	A-C
First	.45(C)	.56(A)	.44(C)	.39(A)
Second	.47(A)	.58(C)	.39(A)	.60(C)
Concreteness Effect	-.02	.02	.05	.21

Examination of Table 1 reveals that in both rating conditions, concrete paragraphs were better remembered when they were presented second than when they were presented first. Abstract paragraphs, in contrast, appeared equally well remembered regardless of presentation position. This result, together with the finding of concreteness effects only when assessed within-subjects, suggests that there is a difference between concrete and abstract paragraphs, in favour of concrete ones, that is enhanced by the prior presentation of abstract material. This finding is similar to that obtained by McDaniel and Einstein (1986) in their study of bizarreness effects in paired-associate learning (see also, Marschark & Hunt, 1989), and both findings appear to be interpretable in terms of a common distinctiveness mechanism.

Before we can accept the distinctiveness account, however, there remains at least one empirical puzzle to be solved. Although all of the above studies employing texts drawn from Marschark's (1978) pool have consistently demonstrated within- but not between-subjects concreteness effects, both the original Yuille and Paivio (1969) study and a more recent one by Ernest (in press) obtained concreteness effects using between-subjects designs. In fact, both of those studies used the same materials, and therein may lie the answer to the puzzle.

As noted in Experiment 1, the materials developed by Marschark (1978) and used by Bunn (1986), Marschark (1985), and Ransdell and Fischler (1989), comprised pairs of concrete and abstract passages that were equated for comprehensibility, syntactic structure, and conceptual structure. In addition, the members of each pair shared many of the same words. The paragraphs used by Yuille and Paivio (1969) and Ernest (in press) comprised concrete and abstract pairs, each having similar themes, emotionality, and distributions of word frequencies. However, the members of each pair were not equated for comprehensibility, conceptual structure, or syntactic structure, and shared common themes only generally, as for example, in two paragraphs that concerned military conflict (see Appendix).

Possible structural differences between the concrete and abstract paragraphs in the Yuille and Paivio materials are relevant here, because several recent studies by Trabasso and his colleagues (e.g., Trabasso & Sperry, 1985; Trabasso, van den Broek, & Suh, 1989) have suggested that such differences should have marked effects on memory. The primary focus of the Trabasso work concerns the nature of connections among the events of a text and, in particular, differences between causal, enabling, and temporally related events. Although other structural analyses are certainly possible with the same texts, Trabasso's model is particularly interesting in the present context because of the obvious differences between the Marschark (1978) and Yuille and Paivio (1969) materials in this regard.

In Marschark's materials, both paragraphs of a pair are about the same person and comparable episodes are described, albeit differing in their concreteness and detail. Thus, their causal structures are essentially identical. In Yuille and Paivio's materials, in contrast, only the concrete passages are about specific episodes, whereas the abstract passages describe more general states of affairs. Thus, the concrete passages are more causally interconnected than the abstract ones, a difference that is readily apparent in reading the paragraphs (examples of which are presented in the Appendix). This difference, a priori, would be expected to lead to better memory for the concrete than abstract passages independent of their rated concreteness. Experiment 2 thus was designed to determine both the relative structural coherence of the Marschark (1978) and Yuille and Paivio (1969) paragraphs and the effects of that coherence on memory. In order to maximize the possibility of concreteness effects, a within-subjects design was employed.

EXPERIMENT 2

This experiment provided structural information concerning the concrete and abstract paragraphs used in the several studies described above. Subjects were given the sentences of four paragraphs, each in a random order on a separate page, and were asked to re-order them in their "original form." Marschark (1985) demonstrated that when the sentences of his paragraphs were presented as randomly-ordered lists, concreteness effects were obtained. When the same lists were presented in paragraph order, concreteness effects were not obtained. That study, however, did not involve an orienting task that would encourage subjects to notice the relational information among the sentences, an apparently necessary condition for the demonstration of concreteness effects with word lists (see Marschark & Surian, 1989, for a review).

If explicit relational processing of concrete and abstract paragraphs is sufficient to produce apparent "concreteness effects" in memory for prose, the effect of paragraph type should be reliable with both sets of materials in this experiment. Because a within-subjects design was used, effects of paragraph type should be optimized if the distinctiveness account is sufficient to explain "concreteness effects" in memory for prose. Alternatively, if the structural differences between Yuille and Paivio's paragraphs are the source of the "concreteness effects" reported in their 1969 study and that of Ernest (in press), those passages would produce effects of paragraph type while those of Marschark (1978) would not.

Method

Subjects, Design, and Procedure.

Thirty-three introductory students participated in this experiment and received credit toward a course research requirement. Each received four paragraphs: two of the paragraphs used by Yuille and Paivio (1969) and two of those used by Marschark (1985). Concrete and abstract versions of

Marschark's paragraphs were counterbalanced across subjects, because they could not receive the two members of a pair without noticing their obvious similarity. All of the paragraphs used in Experiment 2 are presented in the Appendix.

Each paragraph appeared on a separate page of a 4-page test booklet. On each page, sentences were listed in random order with a blank beside each. Subjects were told that the experimenters were interested in the coherence of texts that had been used in previous experiments, and they were asked to try to reconstruct the original passages from the randomly-ordered sentences, numbering them to denote order. The orienting task was subject-paced; when finished, subjects turned their test booklets face down. When everyone had finished (approximately 6-7 minutes), an unexpected free recall test was given. The instructions indicated that original wording was desired, but that words, phrases, or ideas could be written if necessary.

Table 2. Correlations of original paragraph structures with subjects' reordering of sentences in Experiment 2 - and their subsequent recall.

	Paragraphs			
	Yuille & Paivio (1969)		Marschark (1985)	
	Concrete	Abstract	Concrete	Abstract
Mean Coherence Correlation	.64	.19	.78	.90
Mean Proportion Recalled	.23	.02	.23	.19

Results and Discussion

Orienting task

Recall that according to Trabasso's causal analyses of story structure, paragraphs with more coherent structures would be expected to be remembered better than those with less coherent structures. In order to assess relative differences in perceived structural coherence between the concrete and abstract versions of the Yuille and Paivio (1969) and Marschark (1985) paragraphs, subjects' reconstructions of each paragraph were correlated with the original ("correct") orders of the sentences. As can be seen in Table 2, those correlations proved highly reliable for Marschark's concrete and abstract paragraphs. The correlation for Yuille and Paivio's (1969) concrete paragraph also was highly reliable, but that for their abstract paragraph, in which individual subjects' correlations ranged from -.94 to +.94, was not.

Consistent with structural analyses of these paragraphs provided by Tom Trabasso, the above results indicate that the abstract passage from the Yuille and Paivio study was significantly less coherent (or relationally interconnected) than the other materials employed here. On this basis alone, one would expect a larger difference in recall between the concrete and abstract paragraphs used by Yuille and Paivio than those used by Marschark. As observed, that difference would be expected to derive primarily from low recall of the Yuille and Paivio abstract paragraph. It is now clear, however, that any memory difference obtained should not really be called a "concreteness effect."

Recall

Gist scoring of recall was based on predetermined propositions and was analyzed using a 2 (paragraph type - concrete or abstract) x 2 (material set) within-subjects design. The proportions of propositions recalled from both the Marschark materials (out of 18 per paragraph) and the Yuille and Paivio materials (out of 23 per paragraph) are shown in Table 2. The obvious patterns among those means were verified by the statistical analyses. Overall, there were reliable effects of material set, $F(1,32)=8.36$, $MSe=.02$, and paragraph type, $F(1,32)=25.03$, and these effects were qualified by their highly reliable interaction, $F(1, 32)=13.08$. As can be seen in Table 2, there was a robust effect of paragraph type with the Yuille and Paivio materials but not the Marschark materials, according to Newman-Keuls tests.

Taken together with the data from the orienting task, the results for recall from this experiment indicate that the previous reports of reliable concreteness effects using the paragraph materials of Yuille and Paivio (1969) are largely due to artifactual differences in the materials and not their rated concreteness. Several studies, using various subsets of paragraphs drawn from the Marschark (1978) materials, have now failed to obtain "concreteness effects," and we are led to conclude that those effects generally are not obtained in memory for prose when the to-be-remembered materials are equated for comprehensibility and structure

(Schwanenflugel & Shoben, 1983). This finding clearly holds in all of the between-subjects comparisons in which it has been made (Experiment 1; Bunn, 1986; Marschark, 1985; Wattenmaker & Shoben, 1986).

When within-subjects designs are employed, however, the case is not entirely clear. Bunn (1986) obtained a reliable within-subjects concreteness effect in recall of gist but not in recall of nouns; and Ransdell and Fischler (1989) also obtained reliable effects in gist recall. In this experiment, a concreteness effect was not obtained with comparable materials. Further, the lack of a concreteness effect was replicated in an additional experiment involving the same subset of Marschark's paragraphs. In that study, 96 subjects rated either the imageability or the comprehensibility of the correctly ordered sentences. No concreteness effects were obtained either overall or in any of the separate cells of the design.

Although acceptance of null results is usually a tenuous affair, the consistency of the present findings across independent replications gives us confidence in their reliability. The results contrast with those of Ransdell and Fischler (1989), however, with no apparent reason for the difference. In the condition closest to their (Experiment 3) demonstration of concreteness effects, the imagery rating condition in the replication experiment yielded a concrete-abstract difference of less than 1%.

SUMMARY AND CONCLUSIONS

The results of these and previous experiments indicate that concreteness effects might be robust in some memory tasks, but that prose memory is not one of them. When concreteness is manipulated between-subjects using well-controlled materials, concreteness effects have not been demonstrated. When manipulated within-subjects, concreteness effects are variable, even when the same materials are used. The findings of reliable concreteness effects using a within- but not a between-subjects design in Experiment 1 and previous studies is consistent with a view of imagery and concreteness that places the locus of memory effects on their enhancement of relative item distinctiveness (Marschark & Surian, 1989; Marschark et al., 1987) rather than on the production of an extra memory code.

Imagery seems to us one particularly powerful source of distinctiveness in verbal as well as nonverbal memory tasks, insofar as images constructed to capture the meanings of particular items are likely to be relatively unique and specific. Imagery is not the only variable that can serve this function (see Marschark & Cornoldi, 1990), and it does not always have beneficial effects (see Marschark & Hunt, 1989). However, to the extent that such processing leads to memories that are easily discriminated in memory relative to those produced by other encoding strategies (e.g., superficial reading or categorization), such items should be relatively better remembered.

This framework has provided explanations of concreteness effects in free recall of word lists (Marschark & Surian, 1990) and paired-associate learning (Marschark & Hunt, 1989), as well as in memory for sentences and paragraphs. Some of those results also can be accounted for in terms of either the dual coding model of Paivio (1971, 1986) or the context availability model of Schwanenflugel and Shoben (1983), in which concrete materials are seen to evoke more information from memory than abstract materials. Results such as those obtained in Experiment 1, however, are not easily reconciled with either of those models (Ransdell & Fischler, 1989), and the context availability model has further difficulty in explaining the demonstration of concreteness effects in free recall tasks in which context availability has been controlled (Marschark & Surian, 1990).

An essential aspect of the present framework is that distinctive and relational information go "hand-in-hand" in affecting memory (Marschark et al., 1987). The lack of concreteness effects in memory for well-controlled prose materials and the results of Experiment 2, in particular, reflect the essential role of relational (i.e., conceptual, structural) information in text processing. Unlike list learning tasks, in which relational information is at a minimum and individual item distinctiveness is most important (Marschark & Hunt, 1989), memory for prose should be less dependent on images relating to any particular proposition or idea. This is not to say that imagery might not play an important role in the comprehension of concrete prose, especially when spatial descriptions are involved (e.g., Denis & Denhière, 1990; see Robin & Denis, this volume). However, traditional explanations of concreteness effects in prose memory in terms of the number of alternative memory codes that are retained clearly cannot account for the results described above.

We believe that the inconsistent occurrence of concreteness effects in experiments involving within-subjects designs, even when the same materials are used, can best be interpreted as indicating the relatively peripheral role of imagery in memory for prose and the influence of individual differences in memory strategies. In fact, we have observed similar inconsistencies in attempts to explore the role of imagery in various other memory tasks, as concreteness effects appear and disappear from one experiment to another using the same materials and orienting tasks. This phenomenon may be widespread with other researchers, but it is not the stuff that journal editors are likely to publish. For the most part, published research tells us of those situations in which imagery and concreteness serve positive roles in memory, however narrowly circumscribed the tasks. We believe that there is far more to the story, and that we have just begun to understand how imagery affects memory.

REFERENCES

Bunn, C. (1986). Individual differences in spatial-imagery and memory for concrete and abstract prose. Unpublished Honors Thesis, Trent University, Peterborough, Ontario.

Denis, M. & Denhière (1990). Comprehension and recall of spatial descriptions. *European Bulletin of Cognitive Psychology* **10**, 115-143.

Ernest, C. (in press). Ability differences and prose learning. *Intelligence.*

Jacoby, L.L. & Craik, F.I.M. (1979). Effects of elaboration of processing at encoding and retrieval: Trace distinctiveness and recovery of initial context. *In* L.S. Cermack & F.I.M. Craik (Eds.), *Levels of processing in human memory* (pp. 1-21). Hillsdale, NJ: Lawrence Erlbaum.

Marschark, M. (1978). *Prose processing: A chronometric study of the effects of imagability.* Unpublished doctoral dissertation, University of Western Ontario, London, Ontario.

Marschark, M. (1979). The syntax and semantics of comprehension. *In* G. Prideaux (Ed.), *Perspectives in Experimental Linguistics.* Amsterdam: John Benjamins.

Marschark, M. (1985). Imagery and organization in the recall of prose. *Journal of Memory and Language* **24**, 734-745.

Marschark, M. & Cornoldi, C. (1990). Imagery and verbal processing. *In* C. Cornoldi & M. McDaniel (Eds.), *Imagery and cognition* (pp. 133-182). New York: Springer-Verlag.

Marschark, M. & Hunt, R.R. (1989). A reexamination of the role of imagery in learning and memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition 15,* 710-720.

Marschark, M., & Paivio, A. (1977). Integrative processing of concrete and abstract sentences. *Journal of Verbal Learning and Verbal Behavior 16,* 217-231.

Marschark, M., Richman, C.L., Yuille, J.C., & Hunt, R.R. (1987). The role of imagery in memory: On shared and distinctive information. *Psychological Bulletin* **102**, 28-41.

Marschark, M. & Surian, L. (1989). Why does imagery improve memory? *European Journal of Cognitive Psychology* **1**, 251-263.

Marschark, M. & Surian, L. (1990). Where are the concreteness effects in free recall? *Manuscript submitted for publication.*

McDaniel, M.A., & Einstein, G.O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **12**, 54-65.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart, and Winston. Reprinted by Lawrence Erlbaum.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Ransdell, S.E. & Fischler, I. (1989). Effects of concreteness and task context on recall of prose among bilingual and monolingual speakers. *Journal of Memory and Language 28*, 278-291.

Schwanenflugel, P. & Shoben, E.J. (1983). Differential context effects in the comprehension of abstract and concrete prose. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **9**, 82-102.

Trabasso, T. & Sperry, L.L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language* **24**, 595-611.

Trabasso, T., van den Broek, P., & Suh, S.Y. (1989). Logical necessity and transitivity of causal relations in stories. *Discourse Processes* **12**, 1-25.

Wattenmaker, W.D. & Shoben, E.J. (1987). Context and the recallability of concrete and abstract sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **13**, 140-150.

Yuille, J.C. & Paivio, A. (1969). Abstractness and the recall of connected discourse. *Journal of Experimental Psychology* **82**, 467- 471.

APPENDIX

Prose material used for Experiments 1 and 2

From Yuille and Paivio (1969)

Major Smith watched horrified as withering fire crumpled an immediate superior. Now he must guide the small group on their important journey. Urging his men to keep silent they moved through dense snake infested forests.

Suddenly clattering caterpillar treads sounded nearby. After rapidly cranking in one round he estimated distance and elevation. His finger carefully tightened and the steel monster was blown apart. Fast thinking saved them. This had been only a minor incident among endless desperate encounters.

Despite honest intentions of many countries to avoid conflicts disagreements occur. When different national interests clash contrary effects arise. Resulting events involve confrontations where senseless happenings take place. While men sacrifice their lives for certain ideals they must temporarily suspend other basic principles regarding humanity and justice as a consequence of this intergroup strife. Normal life aims become subordinated to active struggle towards conquest for power. Individuals in serving military goals meet painful encounters effecting small gains and successes.

From Marschark (1985)

While he was a military judge in Washington Allen Thomas was involved in a trial involving three wounded prisoners of war. These ragged prisoners included a blind British admiral. Usually high ranking prisoners were questioned in a small dark cell and then released. However Thomas had received reports that the admiral had beaten his own sailors and had tortured captive officers. He quickly had the prisoners put in chains and denied them all food and water. After two days of heated arguments his aging commander ordered him to release them.

While he was a public magistrate in Washington Allen Thomas was involved in an incident involving several captured enemy agents. These particular agents included an important enemy official. Usually even important agents were questioned in a purely routine manner and then paroled. However Thomas had received reports that this official was unfair to his own personnel and had mistreated captured agents. He quickly had the agents put in custody and denied them all privilege of communication. After some time of careful consideration his troubled conscience convinced him to release them.

Early in his college career John Williams was chosen by his football coach to receive a trophy from a professional team. The ceremony proved to be crowded and was prolonged over a two hour period. John was excited however to receive awards for scoring and sportsmanship from their star player. Further his large size and fast running helped to improve opinions of his college's players. Williams went on to become a star halfback and gained an international reputation.

Early in his college career John Williams was chosen by his department to receive a scholarship to a foreign university. The program proved to be difficult and was prolonged over a several year period. John was successful however in obtaining degrees in philosophy and humanities from the institution. Further his keen interest and tactful personality helped to improve opinions of American students. Williams went on to become a brilliant philosopher and gained an international reputation.

This Page Intentionally Left Blank

Part 5

Imagery, Reasoning, and Problem Solving

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 15

Why does mental visualization facilitate problem-solving?

Alessandro Antonietti
Università Cattolica del Sacro Cuore, Milan, Italy

INTRODUCTION

It has been argued that visual imagery can play a symbolizing and a simulation function in thinking (Kosslyn, 1983). Mental images may provide a simplified representation of the content of the cognitive task. For instance, schematic images highlight the salient elements of a given situation allowing such elements to be processed easily and rapidly. Moreover, images permit representation of objects and events in an analogue format, so that it is possible mentally to anticipate transformations of the external world.

The anticipatory role of reasoning is fundamental in the solution of problems, where the requirement is to reach a solution from a given starting state. Thus, the involvement of visual images in problem-solving may result in a notable degree of success. In fact, imagery can be regarded as a mental code with characteristic properties which are relevant to the generation of valid conclusions beyond the information initially given (Denis, 1990).

These remarks are consistent with the implicit and intuitive grounds of many instructional procedures which are aimed at improving problem-solving by means of the productive use of figural representations (e.g. Schoenfeld, 1979; Wicker et al., 1978). Furthermore, relationships between mental imagery and creativity suggest that visualization can facilitate the activation of original reasoning strategies (Forisha, 1978; Suler & Rizziello, 1987). Nevertheless, little experimental effort has been directed towards studying the specific role of imagery in the solution of insight problems (see Richardson, 1983; Kaufmann, 1988).

As far as this topic is concerned, the review of the literature suggests that visualization plays heuristic functions in cognition because:
1) mental images are a set of processes that allow subjects to avoid the mechanical use of algorithms elicited by the verbal formulation of the problem (Shepard, 1978);
2) mental images are flexible representations which permit the manipulation of problem elements in unusual ways (Kabanova-Meller, 1971);
3) mental images are holistic representations specialized for parallel processing of information (Kaufmann, 1985), so that they facilitate simultaneous consideration of the whole problem field. Thus they mitigate the tendency to examine the single elements sequentially;
4) mental images may help subjects to restructure the problem because they hint at dynamic transformations of the material given (Bejat, 1972).

In some problems the solution is difficult to achieve because subjects are induced to make use of complex, unnecessary procedures. More precisely, data quoted in the problem may produce the tendency towards an automatic application of familiar reasoning strategies (e.g. calculations, theorems, rules etc.) that are misleading. If these biases, often referred to as "mechanization" or "set" effects (Luchins, 1942) are working in problem-solving, the representation of the problem elements in a figural mode may result in the suppression of the deceptive tendencies elicitated by the verbal formulation of the task.

Another main obstacle to problem-solving closely related to mechanization is "functional fixedness" (Duncker, 1963). This refers to the tendency to consider some critical elements of the problem in their familiar role, while the solution requires that they are employed in more unusual ways. In these situations, the visual representation of such elements may produce a flexible mental elaboration of them and suggest some alternative uses.

These hypotheses have been supported indirectly by positive correlations between imagery abilities and factors of divergent thinking (Durndell & Wetherick, 1976; Kaufmann, 1981). The aim of the first two experiments described in the present paper was to provide more direct evidence about the effects of mental visualization on mechanization (Experiment 1) and on functional fixedness (Experiment 2) in reasoning.

According to Gestalt psychologists (e.g. Wertheimer, 1959), subjects solve insight problems by means of a simultaneous reorganization of the cognitive field. This operation allows subjects to avoid the inclination to consider separately the elements of the problem. While the sequential, verbal formulation of the problem enhances this tendency, pictures and mental images may counteract it, because information is available simultaneously.

Some empirical data corroborate this interpretation. For instance, Khatena (1977, 1983) found that creative thinking is correlated with the mental elaboration of whole complex images. Furthermore, a series of investigations (Antonietti, Barolo, & Masini, 1987; Barolo, Masini, & Antonietti, in press) showed that the solution of insight problems is associated with the ability to manipulate holistic mental images in two or three dimensions. On the other hand, sequential imagery elaboration is associated with unproductive strategies in problem-solving.

The restructuring of the problem field depends not only on the simultaneous analysis of its elements, but also on changes in its internal organization. Visual imagery may facilitate the latter process because it suggests free transformations of the problem structure. Positive correlations between transformation of mental images and problem-solving support this hypothesis (Frandsen & Holder, 1969; Kaufmann, 1979).

The goal of the last two experiments of this study was to obtain confirmation that mental visualization facilitates the holistic processing of the problem field (Experiment 3) as well as its internal reorganization (Experiment 4).

EXPERIMENT 1

The aim of this experiment was to assess whether mental imagery helps subjects to get over mechanization biases in reasoning. Two problem-solving tasks were selected: the alarm-clock problem (Raudsepp, 1980) and the high tide problem (Mosconi & D'Urso, 1974).

a) Alarm-clock problem: "John went to sleep at 8 p.m., having previously wound up his old alarm clock and set the hands to wake him up at 9 a.m. He slept soundly until the alarm rang. How many hours did John sleep?". Subjects usually answer "13 hours" because they count the number of hours from 8 p.m. to 9 a.m. and do not realize that at 9 p.m., that is after 1 hour, the hand of the alarm makes it ring.

b) High tide problem: "A rope ladder was hanging from a boat so that the ladder had six rungs above the sea. The distance between any two rungs was 30 cm. At high tide, the sea level rose 70 cm. How many rungs were above the sea at high tide?". Subjects seldom realize that the tide lifted the whole boat and, consequently, six rungs were always above the water.

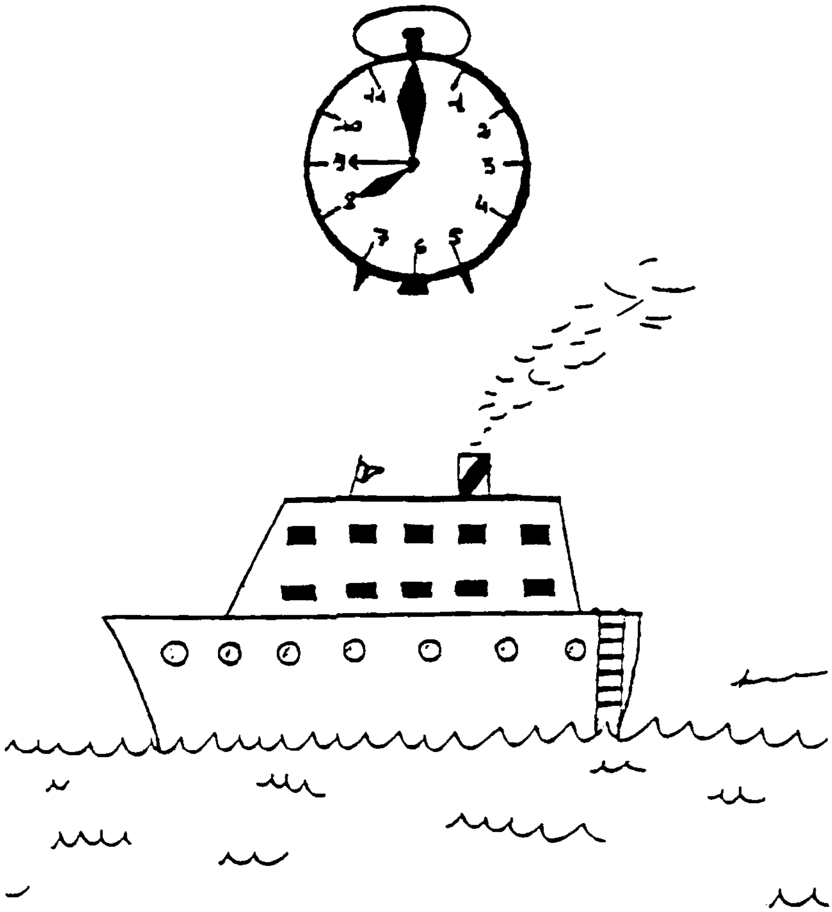

Figure 1. Pictures used for the alarm clock and high tide problems in Experiment 1.

The statements of both problems prompt the automatic application of mathematical operations which are misleading. The imaged representation of the situations should suppress such a tendency by means of the mental simulation, in an analogue code, of the physical events mentioned in the problems.

Three different versions of each problem were randomly distributed among three independent groups of 13-14 year old children attending secondary schools. Each group comprised 35 subjects. The "verbal" group received only the statement of the problem. The "picture" group received the problem with an illustration of the scene described. More precisely, the picture of the first problem depicted an old alarm-clock with the hour-hand set at 8 and the alarm-hand set at 9 (see Figure 1a); the picture of the second problem depicted a boat with a hanging rope ladder, six rungs of which were above the sea (see Figure 1b). The subjects in the "mental visualization" group received the problem without a picture but with the instruction to represent visually in their minds the situation described in the problem. Problems were individually presented in written form. Five minutes were allotted for the solution of each problem.

The distributions of solvers and nonsolvers in the alarm-clock problem and in the high tide problem are reported, respectively, in Tables 1 and 2.

Table 1. Distribution of solvers and nonsolvers in the alarm clock problem in Experiment 1 (Proportions of solvers and nonsolvers within each group of subjects are reported in parenthesis)

Groups	Verbal	Picture	Mental Visualization
Solvers	0 (0.00)	2 (0.06)	3 (0.09)
Nonsolvers	35 (1.00)	33 (0.94)	32 (0.91)

Comparison between the proportions of solvers in the Verbal and Picture group: $z=-1.43$, n.s.. Comparison between the proportions of solvers in the Verbal and Mental visualization group: $z=-1.77$, $p<0.05$

In both problems visualization yielded the highest frequency of correct solutions. So we are induced to maintain that mental simulation, in a visual image, of the problem situations may help subjects in avoiding the mechanization biases elicitated by the verbal formulation of the problems. In the alarm-clock problem this effect was partially produced also by pictures.

On the contrary, in the high tide problem, pictures had negative effects. The graphic illustration of the scene decreased the number of correct responses and, as seen from the analysis of the protocols, increased the number of the response "three rungs". This response is mathematically correct if it is erroneously assumed that the boat was stationary while the sea level rose along its sides.

Table 2. Distribution of solvers and nonsolvers in the high tide problem in Experiment 1 (Proportions of solvers and nonsolvers within each group of subjects are reported in parenthesis)

Groups	Verbal	Picture	Mental visualization
Solvers	7 (0.20)	5 (0.14)	10 (0.29)
Nonsolvers	28 (0.80)	30 (0.86)	25 (0.71)

Comparison between the proportions of solvers in the Verbal and Picture group: z=0.63, n.s.. Comparison between the proportions of solvers in the Verbal and Mental visualization group: z=-0.84, n.s.

We can conclude that the static visual representation (the pictures condition) of the problem situation has at best a weak effect in counteracting the mechanization of thinking. On the other hand, the kinetic representation, yielded by mental visualization instructions, may allow subjects to overcome this deceiving bias in reasoning. In both problems studied in the present experiment, mechanization effects depend on the tendency to consider the initial positions of the critical objects (the hour-hand in the alarm-clock problem and the boat in the high tide problem) as fixed. Presumably this bias towards a static consideration of the critical elements of the task may be enhanced by pictures. Pictures, in fact, provide a "frozen" representation of the starting situation described in the problem. In contrast, visualization may "break" this tendency and facilitate the mental simulation of the continuous movements of the critical objects.

EXPERIMENT 2

The goal of the second experiment was to study the effects of mental visualization on functional fixedness. In this regard two problems were employed: the pea problem (Raaheim, 1974) and a modified version of the cord problem (Maier, 1930).

a) Pea problem: "Two bowls were set on a table, one within the subject's reach and one further away. One bowl contained a number of peas and the other was empty. Also on the table were a newspaper, scissors, paper clips and rubber bands. The subject's task was to devise as many ways as possible, using the material provided, of transferring the peas from the filled to the empty bowl without leaving the seat". The most efficient solution consists of rolling the newspaper to form a tube (or in folding the newspaper to form a gutter), positioning the newspaper between the two bowls and then rolling the peas through the tube (or the gutter).

b) Cord problem: "Suppose you are in a room where two cords are hung from the ceiling. The two cords are of such a length that when you hold one cord in either hand, you cannot reach the other. Your task is to tie the ends of these cords together. The room is empty. You have only a bunch of keys". The "pendulum solution" of this problem consists of tying a weight (for instance, the bunch of keys) to one cord and to swing it so it becomes a pendulum. The other cord can be brought toward the centre, so that the swinging cord can be caught as it approaches the midpoint, and then the two cords can be tied together.

The solutions, described above, of both problems require the subjects to overcome functional fixedness because the critical objects (respectively, the newspaper and the bunch of keys) have to be employed in uncommon ways.

The problems were presented to two independent groups of 25 students, ranging in age from 16 to 20 and attending secondary schools. The "control" group received only the verbal formulation of the problem. The "visualization" group, before being given the problem, was instructed to represent mentally and to transform freely the scene and/or the objects described in the problem. More precisely, the subjects were told that they had to visualize the situation and to imagine any change in it if they wished. The duration of each visualization task was five minutes. Finally subjects verbally reported the contents of their visualizations. Problems were individually presented in a written form and ten minutes were allowed for the solution of each problem.

The distributions of productive and unproductive solvers in each group for each problem are reported in Table 3 (pea problem) and Table

4 (cord problem). Subjects were classified as productive solvers if they employed the critical objects in unusual ways (that is, the paper as a tube or as a gutter and the bunch of keys as a pendulum), thus overcoming functional fixedness. On the other hand, subjects were classified as unproductive solvers if they gave other responses involving common uses of the critical objects.

Table 3. Distribution of productive and unproductive solvers in the pea problem in Experiment 2 (Proportions of productive solvers and unproductive solvers in each group of subjects are reported in parenthesis)

Groups	Control	Mental visualization
Productive solvers	16 (0.64)	22 (0.88)
Unproductive solvers	9 (0.36)	3 (0.12)

Comparison between the proportions of productive solvers in the Control and Mental visualization group: $z=-1.97$, $p<0.05$

In both problems the visualization tasks produced a more flexible use of the critical objects than in the control groups. In the pea problem, the newspaper was employed in functions (as a tube or as a gutter) different from the usual ones. In the cord problem, a large number of subjects in the visualization group considered the ropes in dynamic ways (that is, swinging it) and proposed the use of the bunch of keys as pendulum. It is worth noting that, for both problems, in the experimental groups, subjects who then produced the solutions involving the suppression of functional fixedness, had not reported the corresponding transformations of the critical elements of the problem during the previous visualization tasks.

In conclusion, the results of the experiment suggest that in problem-solving the free mental visualization of the available objects leads subjects to employ them successively in a very flexible manner.

Therefore, mental visualization can be regarded as a useful strategy to avoid functional fixedness in thinking.

Table 4. Distribution of productive and unproductive solvers in the cord problem in Experiment 2 (Proportions of productive solvers and unproductive solvers in each group of subjects are reported in parenthesis)

Groups:	Control	Mental visualization
Productive solvers	12 (0.48)	21 (0.84)
Unproductive solvers	13 (0.52)	4 (0.16)

Comparison between the proportions of productive solvers in the Control and Mental visualization group: $z=-2.69$, $p<0.01$

EXPERIMENT 3

In order to verify whether visual imagery facilitates the simultaneous processing of the problem field, the sock problem (Mosconi & D'Urso, 1974) was chosen: "In a drawer you have five white socks and five black socks. In the dark, what is the smallest number of single socks which you would have to take out to be sure of getting a matching pair?".

The correct answer is "three socks", but subjects often respond that a greater number of socks are necessary or that it is impossible to know how many socks are needed. In fact, subjects tend to consider separately each extraction, whose result is unpredictable. However, if we consider together various extractions, we can make some predictions about their results. Thus, it is possible to hypothesize that the simultaneous representation of the elements of the problem facilitates its solution.

The problem was presented to a total of 105 children aged 13-14 years. Subjects were randomly assigned to three independent groups as in Experiment 1. The procedures were as for Experiment 1. Subjects in the picture group were presented with an illustration, depicting five white and five black socks in two rows (see Figure 2). Subjects of the visualization group were instructed to imagine five white and five black socks set tidily in a drawer.

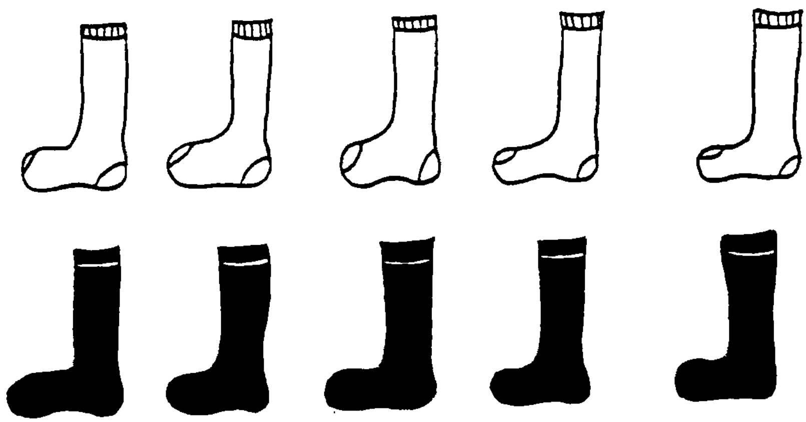

Figure 2. Illustration of the sock problem used in Experiment 3.

The distribution of solvers and nonsolvers in the verbal, picture and visualization groups is reported in Table 5.

Pictures facilitated the solution of the sock problem in comparison with the control group. In contrast, visualization caused inhibitory effects. The data showed that pictures, but not mental visualization, helped subjects to simultaneously represent the elements of the problem.

Negative effects of visualization can be explained as follows. Firstly, we can argue that the visualization of the problem situation and of its successive transformations produced an high mental load; subjects were not able to keep in their mind a distinct and lasting image of the situation, so that interference effects occurred in problem-solving. In contrast, pictures provided a clear and permanent representation of the initial state of the problem which facilitated the simultaneous consideration of the results of the extraction of socks. Secondly, the formulation of the problem might have implicitly encouraged subjects to visualize mentally the socks in the dark, even if the instructions of the visualization task excluded it. If subjects imagined a whole dark scene, obviously they could hardly "see" in their minds the results of the extractions of socks from the drawer, or could not even "see" them at all.

Table 5. Distribution of solvers and nonsolvers in the sock problem in Experiment 3 (Proportions of solvers and nonsolvers in each group of subjects are reported in parenthesis)

Groups	Verbal	Picture	Mental visualization
Solvers	13 (0.37)	24 (0.69)	7 (0.20)
Nonsolvers	22 (0.63)	11 (0.31)	28 (0.80)

Comparison between the proportions of solvers in the Verbal and Picture group: $z=-2.63$, $p<0.01$. Comparison between the proportions of solvers in the Verbal and Mental visualization group: $z=1.59$, n.s.

In conclusion, in problems whose solution requires considering various steps, subjects may have difficulty in mentally forming and updating a whole visual image representing a large number of discrete elements. In that event, pictures may be more useful, because they, contrary to what happens with mental images, are not affected by limitations in size, vividness, stability, duration, etc..

EXPERIMENT 4

The experiment was designed to study the restructuring properties of mental visualization. Two problems were selected whose productive solutions require the reorganization of the material given: the square and rhombus problem (Kosslyn, 1983) and the Gauss problem (Wertheimer, 1959).

a) Square and rhombus problem: Subjects were asked to determine the ratio between the area of a square and the area of the rhombus inscribed in the square. Instead of the application of geometric formulae,

an intuitive solution (described by Plato in the Meno) can easily be achieved by means of a spatial transformation of the figures. If we consider the four triangles formed by the sides of the rhombus and the half-sides of the square, and we imagine folding them inward, we obtain a second rhombus as wide as the one inscribed in the square: thus the area of the square is the double the area of the rhombus.
b) Gauss problem: subjects were requested to find a quick and simple method to sum the first 10 integer numbers. The solution (found by Gauss when he was attending the primary school) consists of decomposing the numerical series into an ascending progression (1+2+3+...10) and a descending one (10+9+8+...1) and realizing that the sum of the symmetric ends of the numerical series is always the same (1+10=11, 2+9=11, etc.), so that the solution can be reached rapidly as follows: 11 multiplied by 5 is equal to 55.

The problems were presented to two independent groups of 25 secondary school students (age: 16-20 years) as in Experiment 2. In the experimental group subjects were instructed to visualize and freely transform the square and rhombus figure (in the first problem) and a ten-step stair (in the second problem).

Table 6. Distribution of solvers and nonsolvers in the square and rhombus problem in Experiment 4 (Proportions of intuitive solvers, demonstrative solvers and nonsolvers in each group of subjects are reported in parenthesis)

Groups	Control	Mental visualization
Intuitive solvers	5 (0.20)	20 (0.80)
Demonstrative solvers	5 (0.20)	1 (0.03)
Nonsolvers	15 (0.60)	4 (0.17)

Comparison between the proportions of intuitive solvers in the Control and Mental visualization group: $z=-4.24$, $p<0.001$

It is hypothesized that in the square and rhombus problem the previous visualization of the geometric figure may suggest the folding operations involved in the intuitive solution mentioned above. Similarly, in the Gauss problem the visualization of the stairs may suggest ascending and descending movements, isomorphic with the ascending and descending sequences underlying the series of integer numbers.

The distribution of solvers and nonsolvers in the square and rhombus problem is reported in Table 6. Solvers were divided into intuitive solvers (subjects who solved the problem only by means of the spatial transformations described above) and demonstrative solvers (subjects who solved the problem by applying geometrical formulae).

Table 7 shows the distribution of productive and unproductive solvers in the Gauss problem. Subjects were classified as productive solvers if they followed the solution procedure found by Gauss and mentioned above. On the other hand, subjects were classified as unproductive solvers if they employed other common or non restructuring procedures.

In the square and rhombus problem intuitive solutions were markedly more frequent in the visualization group than in the control group. A similar significant difference was found in the productive solvers of the Gauss problem.

The analysis of the protocols concerning the visualization tasks revealed that, as observed in Experiment 2, cognitive operations involved in solving the problems were not anticipated by subjects during visualization. Furthermore, it is worth noticing that the facilitatory effects in the Gauss problem depended on free flow of visual images produced not, as in the previous problems, starting from the direct representation of a problem element (the newspaper in the pea problem, the ropes hanging from the ceiling and the bunch of keys in the cord problem, the rhombus and square figure in the first problem of the present experiment), but from an analogue cue (the stairs as analogues of the numerical series).

In general, data indicated that visualization facilitates the reorganization of the material given both in geometrical and in mathematical problems. In the first case visualization suggested folding operations which allowed solution of the square and rhombus problem only by means of an internal transformation of the problem field, without the use of geometric formulae. In the second case, visualization raised the probability that subjects would consider the series of integers from a different, uncommon point of view.

Table 7. Distribution of productive and unproductive solvers in the Gauss problem in Experiment 4 (Proportions of productive and unproductive solvers in each group of subjects are reported in parenthesis).

Groups	Control	Mental visualization
Productive solvers	5 (0.20)	12 (0.48)
Unproductive solvers	20 (0.80)	13 (0.52)

Comparison between the proportions of productive solvers in the Control and Mental visualization group: $z=-2.09$, $p<0.05$

CONCLUSIONS

Problem-solving performance in the four experiments described above support the notion that visual images allow subjects to avoid some of the main obstacles to productive thinking and to transform problem situations in unusual but effective ways.

In the present study three different experimental procedure were employed in order to produce a visual representation of the problem elements:

i) presentation of the problem together with illustrative pictures

ii) presentation of the problem with instructions to mentally visualize the scene described

iii) previous instructions to mentally visualize and freely elaborate the problem situation and subsequent presentation of the problem.

Pictures resulted in a high degree of success in tasks whose solution required simultaneous consideration of various elements (Experiment 3) but had no facilitatory effects on mechanization biases in thinking (Experiment 1). This is evidence that pictures aid subjects to form the visual representation of the whole problem field, but enhance the examination of the problem elements in a static form.

The tendency towards mechanization can be partially overcome by visualizing the problem scene (Experiment 1). The mental simulation, in a format which is analogous to the physical world, of the problem situation counteracts the misleading forms of thinking elicitated by the problem as presented. Furthermore imagery, with its isomorphic link to the external objects and events that it represents, may well reveal significant details and relationships that are not adequately represented in a purely verbal formulation of the problem.

Finally, the free flow of visual images contributes to the generation of subsequent flexible representations so that the problem elements easily can be reorganized in productive ways by means of unusual functional and structural transformations (Experiments 2 and 4).

In conclusion, visualization can operate in problem-solving both after and before the problem is given. Once the problem is posed, figural representations are useful mentally to simulate the situation and the transformations described in the problem. In this way, familiar, but misleading, strategies of reasoning can be substituted with new and productive directions of thinking which avoid the "traps" created by the verbal formulation. In fact, the verbal communication system is the primary vehicle for maintaining and perpetrating established ideas and entrenched traditions. Rather, the challenges to such a system are likely to come from outside the system itself", for instance from a visual-spatial recoding of the verbal formulation of the cognitive tasks (Shepard, 1978, p.156). Moreover, previous visual elaborations break the initial rigid configuration of the problem and introduce a high dynamism in the cognitive field, so that the subsequent restructuring of the problem is facilitated. These heuristic functions of mental imagery arise out of the fact that it is a flexible code through which it is possible to prefigure easily and quickly various original transformations of the situations that are represented.

REFERENCES

Antonietti, A., Barolo, E., & Masini, R., (1987). *L'immagine mentale nella scoperta cognitiva. Contributi sperimentali.* Unicopli, Milano.

Barolo, E., Masini, R., & Antonietti, A., (In press). Mental rotation of solid objects and problem-solving in sighted and blind subjects. *Journal of Mental Imagery* **14**.

Bejat, M., 1972. The role of image in problem-solving. *Revue Roumaine des Sciences Sociales* **16**, 169-180

Denis, M., (1990). Imagery and thinking. *In* C. Cornoldi & M.A. McDaniel (Eds.), *Imagery and cognition.* New York: Springer, pp 103-131

Duncker, K., (1963). *Zur Psychologie des produktiven Denkens.* Springer, Berlin

Durndell, A.J. & Wetherick, N.E., (1976). The relation of reported imagery to cognitive performance. *British Journal of Psychology* **67**, 501-506

Forisha, B.L., (1978). Mental imagery and creativity: review and speculations. *Journal of Mental Imagery* **2**, 209-237

Frandsen, A.N. & Holder, J.R., (1969). Spatial visualization in solving complex verbal problems. *Journal of Psychology* **73**, 229-233.

Kabanova-Meller, E.N., (1971). The role of the image in problem solving. *Soviet Psychology* **9**, 346-360.

Kathena, J., (1977). Analogy strategies and production of original verbal images. *Journal of Creative Behavior* **11**, 213.

Kathena, J., (1983). Analogy imagery and the creative imagination. *Journal of Mental Imagery* **7**, 127-134.

Kaufmann, G., (1979). *Visual imagery and its relation to problem solving.* Universitetsforlaget, Bergen.

Kaufmann, G., (1981). The functional significance of visual imagery in ideational fluency performance. *Journal of Mental Imagery* **5**, 115-120.

Kaufmann, G., (1985). A theory of symbolic representation in problem-solving. *Journal of Mental Imagery* **9**, 51-70.

Kaufmann, G., (1988). Mental imagery and problem-solving. *In* M. Denis, J. Engelkamp, & J.T.E. Richardson. (Eds.), *Cognitive and neuropsychological approaches to mental imagery,* Dordrecht-Boston-Lancaster: Nijkoff, pp 231-240.

Kosslyn, S.E., (1983). *Ghost's in the mind machine.* New York: Norton.

Luchins, A.S., (1942). Mechanization in problem-solving: the effect of Einstellung. *Psychological Monographs* **54**, n.248

Maier, N.R.F., (1930). Reasoning in humans. I: On direction. *Journal of Comparative Psychology* **10**, 115-143.

Mosconi, G. & D'Urso, V., (1974). *Il farsi e il disfarsi del problema.* Firenze: Giunti.

Raaheim, K., (1974). *Problem-solving and intelligence,* Oslo: Universitetsforlaget.

Raudsepp, E., (1980). *More creative growth games.* New York: Putnam.

Richardson, J.T.E., (1983). Mental imagery in thinking and problem solving. *In* J.S.B.T Evans (Ed.), *Thinking and reasoning: Psychological approaches.* London: Routledge and Kegan Paul, pp 197-226.

Schoenfeld, A.H., (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education* **10**, 173-187

Shepard, R.N., (1978). Externalization of mental images and the act of creation. *In* B.S. Randhawa & W.E. Coffman (Eds.) *Visual learning, thinking and communication.* San Francisco: Academic Press, pp 133-189.

Suler, J.R. & Rizziello, J., (1987). Imagery and verbal processes in creativity. *Journal of Creative Behavior* **21**, 1-6.

Wertheimer, M.,(1959). *Productive thinking,* New York: Harper.

Wicker, F.W., Weinstein C.E., Yelich C.A., & Brooks J.D., (1978). Problem-reformulation training and visualization training with insight problems. *Journal of Educational Psychology* **70**, 372-377

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 16

Imagery in mental construction and decomposition tasks

Tore Helstrup
University of Bergen, Norway
and
Rita E. Anderson
Memorial University of Newfoundland, Canada

Although exploration of issues in mental construction is fairly recent, debates are already developing and alternative task situations are being designed (cf. Finke, 1989). In the research project to be discussed, we made use of a mental construction task developed by Finke and Slayton (1988). The questions we examined pertain to the effects of internal and external representation and to the effects of visual and verbal strategies on mental construction and decomposition tasks. The experiments are part of a research project still in progress. This chapter presents some of the main data and some of our attempts at theoretical analysis.

In the mental construction task developed by Finke and Slayton, a well-defined and restricted element population is first identified from which new composite figures are to be constructed. The element population might, for instance, comprise the capital letters of the alphabet, the digits 1-9, plus the geometrical forms square, circle, and equilateral triangle. The subjects are presented with two or three elements, most often verbally, and asked to combine them mentally into a meaningful configuration. To give an example: The letters D and J can easily be combined to look like an umbrella (D rotated with its bow pointing upwards, and the J placed in its normal position below the rotated D). The construction rules are specified in advance (e.g. that the elements may be increased or decreased in size, rotated in space, etc.).

The experiments by Finke and Slayton demonstrated that subjects are generally able to perform this task, many coming up with unexpected, creative solutions. Such experiments suggest that mental images, constructed from scratch, may be interpreted in ways that are not apparent from the constituent parts. The "meaning" of the image is therefore not necessarily given before its construction. Interpretation of the constructed image seems to be an operation that can be performed in mental imagery. On this point, however, the opinion is still divided (cf. Fodor, 1975; Chambers & Reisberg, 1985).

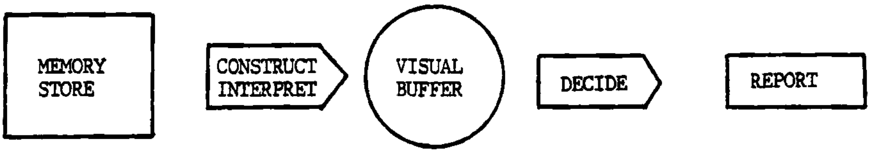

Figure 1. A task-analysis model of mental figure construction.

Figure 1 presents our task-analysis model of mental construction, based on Farah's analyses of various imagery tasks (Farah, 1984). Each trial starts with specification of the elements to be used. The specified elements are assumed to be fed into a memory store, either in visual form or as verbal descriptions, possibly both. Then a configuration must be constructed which is consistent with the rules of combination and which allows a meaningful interpretation. The construction work presumably involves the use of a visual buffer. Operationally, an adequate interpretation requires that a suitable verbal label is attached to the assembled configuration.

The processes called "construction" and "interpretation" may in principle be carried out separately and sequentially. But it is more likely that they interact. For example, a construction attempt may suggest an interpretation, which then influences further construction, and so on. In any case, construction is thought to be a controlled process. To assemble a configuration, one may zoom in, rotate, enlarge, diminish, insert, pile, place the given elements side by side, and so forth. However, while this goal oriented activity is taking place, more automatic, gestalt-like figure formation processes probably also contribute to the building of a configurational unit. What we have termed "construction" might therefore be conceived of as referring to controlled, high-level visual-spatial processes, whereas automatic figure formation processes might constitute low-level processes (cf. Pinker, 1984).

The interpretation processes may be considered as partly controlled, partly automatic. On the basis of a deliberate feature analysis of the emerging pattern, one may search for meaning fragments, or one might "free associate" interpretations to the figure or figure elements, as one does when associating to clouds or ink blots.

Once an interpreted figure has been produced, a decision has to be made as to whether this outcome will be accepted or rejected. This decision has to match the assembled figure against the set of elements in memory to check whether all elements, and only those elements, have been correctly employed. A check must further be made as to whether the derived meaning relation between the configuration and the interpretation is tolerably adequate. If the decision is positive, the result can be reported. If the decision is negative, the construction process has to produce new configurations until an acceptable outcome is reached or until the subject gives up.

In terms of Figure 1, the crucial point in mental construction is to come up with an adequate configuration in the visual buffer. If a main problem concerns memory capacity, e.g. the probability of losing some of the information about elements and construction rules kept in the memory store, then working with paper-and-pencil support should be of an advantage compared to a condition where all of the work has to be done with the information in mind (cf. Reisberg, 1987). On the other hand, if automatic, low-level figure formation processes are important, they may work more efficiently with "internal" perceptual-like representations than with external representations produced by paper-and-pencil support. Lack of a paper and pencil advantage might indicate that automatic figure formation is as important as controlled construction and interpretation, the latter presumably being supported by external representation.

To explore these speculative inferences, Rita E. Anderson and I carried out some experiments comparing mental construction with and without external paper-and-pencil support (Anderson & Helstrup, 1989). In the first experiment subjects were asked to produce one configuration per trial. The procedures were generally similar to those used by Finke and Slayton (1988). Three figure elements were orally presented, and the subjects given 2 minutes in which to assemble a figure. After the assembly period, the subjects were first to write a brief description of any figure they had created and then to draw it to the best of their ability. On half of the trials the subjects had to assemble the patterns without external support, whereas on the other half of the trials, the subjects could use paper and pencil. Independent judges, who were blind with respect to the condition under which each construction had been produced, later rated the correspondence between descriptions and drawings on a scale from 5 (easy to identify from its description) to 1 (impossible to identify). All patterns receiving an average rating of at least 4.0 were classified as good patterns (cf. Finke & Slayton, 1988). No significant differences were found between the external and internal representation conditions, either in terms of the number of good constructions or the mean rating of the constructions.

In a second experiment we compared the number of constructions produced under the same two representation conditions. But now the subjects were requested to produce as many alternative constructions per trial as possible. The results showed that while the production rate was significantly higher with paper and pencil support than without, the proportion of good constructions was about the same for the external and internal representation conditions.

In terms of our task-analysis model (Figure 1) these findings suggest that controlled construction processes work best with external support. Subjects can derive constructions more readily when operations such as rotation, insertion, piling, etc. can be wholly or partially externalized. Yet the emergence of good patterns may depend on "internal dynamics" in the developing pattern. The results thus are consistent with the hypothesis that lower-level configuration processes work better with internally represented information, in this way compensating for the lack of external support.

Would similar results be found if the task had been to decompose the figures, i.e. to break up a given configuration into a specified number of meaningful subparts? The experimental procedures for decomposition would be a reversal of the construction task. The subjects would be presented with a ready-made configuration, with or without an accompanying interpretation, and asked to find its constituent elements, again defined for instance as capital letters, digits, and geometrical forms.

Decomposition control processes will of course be different from those used in construction, e.g. split the given configuration into its top and bottom parts, into its left and right parts, select the most easily identified parts and test for remaining parts, and so on. These decomposition operations probably would interact with interpretation operations. Aside from the different control operations, construction and decomposition seem to be symmetrical on all but one point. To us it seemed questionable that the low-level figure synthesizing processes of the construction model would have parallel low-level fragmentation processes. Natural perception is directed towards pattern formation, not towards visual fragmentation. On this assumption, one might expect that paper-and-pencil support would be relatively more important for decomposition than for construction.

We also believe that people in general have more experience with drawing outline figures than with splitting up figures into a stated number of subparts. Intentional or incidental doodling, interpreting coffee grounds or drifting clouds, represent everyday experiences. There certainly also are situations when one has to find hidden parts of configurations (e.g. in games). Nonetheless there seems to be a clear frequency difference in favour of construction tasks. On the assumption that the importance of external representations increases with decreasing task familiarity, we expected paper-and-pencil support to result in stronger effects in a decomposition task.

As a first step in testing the construction/decomposition analysis, we constructed 56 meaningful configurations. Half of the figures were constructed from two figure elements, the other half from three elements. All configurations were then rated for ease of interpretation, and for ease of decomposition. For the interpretation rating, the task was to rate the ease of finding a meaning for the whole configuration. For the decomposition rating, the task was to rate the ease of breaking the total figure down into the required number of meaningful subpatterns. In the rating the subjects were not asked to give the solutions, only to rate the likelihood of finding the two types of solution.

The constructed figures were on the average rated as easy both to interpret and to decompose. No difference in decomposition ratings was found between 2-element and 3-element configurations, but the 3-element figures were rated as somewhat harder to interpret than the 2-element figures. More interesting is the modest correlation observed between interpretation and decomposition ratings ($r = 0.39$). Although highly significant, the size of the correlation indicates that interpretation and decomposition reflect partially independent operations. The rating data thus made it possible to select configurations with different combinations of high/low values in the two dimensions. From the 56 rated figures, 16 were selected such that there were 4 figures of each of the 4 interpretation/decomposition rating values. These 16 figures were then divided into two equivalent sets of 8 figures, with two for each rating value combination.

In a third experiment we examined the external/internal representation question with regard to task materials. It seemed reasonable to expect that external support should be most important with the figures rated as being most difficult. In one experimental condition subjects were given the figure elements (visually presented) and asked to produce one identifiable pattern per trial in accord with the usual construction rules, being allowed and encouraged to use paper and pencil during the assembly period. At the end of the assembly period, all subjects had to describe the mentally constructed pattern and then draw it.

The other groups were given ready-made figures, and asked to suggest a figure interpretation and to decompose the figure into 2 or 3 elements, depending on figure type. Half of the decomposition subjects worked with paper and pencil, the other half without. For all four experimental conditions the visually presented information, elements or configurations, could be consulted again during the work time period. All subjects were tested for 16 trials.

These procedures meant that all subjects were given identical treatment, except for the external/internal and construction/decomposition variations. Consequently, the decomposition subjects only had to interpret the configurations, while the construction subjects also had to construct the figures. The prediction was that subjects in the decomposition conditions would propose more interpretations, since they did not have to construct the figures. A difference in favour of external representation was expected to be more pronounced for decomposition than for construction. As in previous experiments independent judges rated the correspondence between the description and the constructed pattern.

The results (see Table 1) revealed that more good interpretations were reported when subjects did not have to construct the figures. However, as in the first experiment, no significant difference was found between internal and external representation conditions. As can also be seen in Table 1, the number of failures (misconstructions or no interpretations/constructions) varied significantly between construction and representation conditions. More failures occurred when the figures had to be produced and when the work had to be done with internal representations.

As regards decomposition task performance, the difference was in the predicted direction, but statistically insignificant. Contrary to expectations, decomposition seems to work about as well without as with paper-and-pencil support. Performance did not differ between the first and second set of 8 trials in any analysis. Nor did the results reveal any interactions between figure types (as defined by the ratings) and representation conditions or between figure types and construction conditions. In general, more good figures were reported among the tasks independently rated as easy to interpret. But the decomposition ratings did not relate to the observed performance, not even to the number of decompositions successfully achieved.

Table 1. Mean number of good figures (failures in parentheses) over representation and construction conditions.

Representation conditions	Construction	Decomposition
Internal	9.8 (2.1)	12.6 (1.2)
External	10.3 (1.0)	12.8 (0.4)

The results of our third experiment clearly disconfirmed the prediction that decomposition is different from construction as regards the importance of external representation support. A possible implication is that there exist parallel figure fragmentation processes to the assumed low-level figure formation processes. If so, automatic fragmentation processes may work as well, or even better, with images than with drawings. The difference in number of interpretations between the construction and decomposition conditions lends support to the hypothesis that construction and interpretation are separate operations which individually make important contributions in mental construction.

Our belief that an asymmetry should obtain between construction and decomposition rested on two assumptions: the lack of automatic figure fragmentation processes, and a general lack of experience with figure decomposition. The two assumptions are probably connected, in that sufficient practice with figure decomposition probably would have resulted in the establishment of automatic or semiautomatic fragmentation operations, especially in the event that such fragmentation processes do not belong to the basic visual processes.

To gain some information about people's experience with construction and decomposition of simple figures, a group of subjects were shown how to construct and decompose figures of the kind used in our experiments, and asked to rate how familiar they were with performing the two types of operations. For this purpose a previously developed task familiarity rating scale was used (Helstrup, 1989). The results showed no familiarity difference between construction and decomposition. Indirectly this observation may be taken to support the assumed existence of low-level figure decomposition operations.

The results of the third experiment suggested that interpretation and construction/decomposition are partially independent operations.

Providing subjects with ready-made figures resulted in more good interpreted constructions than when figures also had to be constructed from the given elements. We therefore wanted to see whether simultaneous provision of elements and interpretation would lead to a corresponding improvement. New subjects were given the same 16 figures as verbally presented elements and were asked to construct configurations having specified interpretations. Because these subjects were permitted to use paper and pencil, the external representation groups of our third experiment could be used for comparison purposes. Presentation of elements only (with paper and pencil) resulted in 65% good solutions. Providing in addition the configuration, increased the percentage to 80. Giving both elements and interpretations, lead to 78% good solutions in terms of the correspondence between drawing and interpretation. These comparisons indicate that having the configuration, or the interpretation plus elements, are about equally helpful. Further studies are needed to explore how interpretation and construction/decomposition processes interact.

If solution of mental construction tasks depends on both controlled and automatic processes, one might expect strategies to influence task performance. However, so far the results suggest the contrary. In the first two experiments, no systematic performance differences were observed between the first and second half of trials. Similar results were found for construction/decomposition. This stability seems to speak against development or change of strategies with increased task experience. As a further check on the strategy question, we had two students carry out an experiment where the subjects were given a strategy inducing questionnaire in the interval between the first and second set of four figure construction trials (Tvedt & Wilhelmsen, 1989). The strategy questionnaire forced the subjects to reflect on questions pertaining to change of element size, rotation, insertion, etc. The idea was that experience with the first four figures, together with the focus on element manipulation created by the questionnaire, should induce the use of improved strategies on the second set of trials. But again the results failed to show significant performance differences between the first and second set of trials.

The consistent lack of experience-based performance improvement could mean that the postulated control processes are not easily manipulated, or that the repertoire of construction operations is too restricted. Previous research on metaphor processing has nevertheless demonstrated that induced strategies may produce significant performance differences (Helstrup, 1988a,b). In these experiments visual strategies resulted in better performance than verbal strategies. In several respects mental construction tasks resemble figurative language tasks. Both, for instance, draw on imagery and linguistic functions. In our last experiment we therefore decided to compare the effects of visual and verbal strategies on mental construction. The difference between external and internal representation was kept, so that we could compare the results with the previous experiments. Verbal strategies were expected to be more successful with external paper-and-pencil support, whereas visual strategies were expected to profit more from internal imagery representations. Sixteen 3-element figures were selected from the pool of 56 figures previously rated for ease of interpretation and decomposition. The 16 figures were selected so as to fit the task design used in the third experiment, which meant that four figures were rated as easy both to interpret and decompose, four figures easy to interpret but hard to decompose, and so on. The visual strategy instructions encouraged the subjects to make use of as clear and vivid visual images as possible. To emphasize the importance of imagery, the experimental session began with administration of Gordon's test of visual imagery control. The verbal strategy instructions correspondingly requested the subjects to approach the tasks with a focus on language, pointing out the importance of clear concepts and an awareness of the meaning of the words. This strategy condition began with administration of the vocabulary test taken from the WAIS battery. With both strategies the subjects' task was to construct as many meaningful figures as possible.

Table 2. Mean number of good figures (failures in parentheses) over strategy and representation conditions.

Representation conditions	Strategies Visual	Verbal
Internal	10.6 (1.1)	6.6 (4.3)
External	11.3 (0.8)	6.6 (1.9)

The results as shown in Table 2, revealed a marked difference in favour of the visual strategy condition, but no difference between internal and external representation. As regards the number of failures (misconstructions or no answers), strategy type as well as form of representation were significant. Most failures occurred with verbal processing and with internal representation. But there were no interactions between strategies and representation types, neither for number of good figures, nor for number of failures.

Table 3. Mean number of good figures produced in the first and second block of eight trials.

Processing condition	First block	Second block
Visual/internal	4.8	5.8
Visual/external	5.0	6.3
Verbal/internal	2.3	4.3
Verbal/external	3.1	3.5

The strategy effect would be further supported given enhanced performance level on the second block of 8 trials. Whereas the subjects in our other experiments most likely relied on previously established routines, the subjects in this last experiment probably had to break away from such routines in order to focus on specific verbal or visual measures. Improvement in both strategy conditions was therefore expected, and in fact observed. Table 3 shows that a significant increase was found in all processing conditions.

Finally the number of good figures produced were analysed in terms of figure types. The results of this analysis showed that neither representation type, nor strategy interacted with figure type. As before, although more good figures were reported among the figures previously rated as easy to interpret, the decomposition ratings were not found to affect the performance.

The final experiment thus showed that mental construction performance *can* vary as a function of strategy employed, that the number of good figures produced is equally high with or without external paper-and-pencil support, and that failures increase when external support is removed and with a change from visual to verbal processing.

In sum the experimental evidence has led us to modify our task-analysis model to include low-level construction *and* fragmentation processes. A symmetrical model for construction and decomposition of mental constructions thus seems appropriate. In terms of this conception, the advantage of visual over verbal strategies may be due either to facilitation of low-level figure formation processes, or to an improved interaction between construction and interpretation processes, possibly both. The present strategy instructions were intended to create a broad processing mode, rather than to install specific operations (cf. Helstrup, 1989). A processing mode may be deliberately controlled, but still have as one of its main effects, the preparation of the grounds for automatic organization processes. Compared to the verbal strategy, visual processing resulted in fewer failures and more good figures. This combination would be expected if construction and interpretation processes interact in a positive way. The extent to which a visual strategy effect in mental construction is due to automatic and to controlled processes, future experiments must decide.

All four of our experiments comparing mental construction or decomposition with and without external paper-and-pencil support, consistently showed that people work surprisingly well with mental representation only. However, as indexed by the performance level, and familiarity ratings, our tasks must generally be characterized as rather easy. For harder tasks the results suggest that external representation support will be of greater importance. There is thus no question about the importance of external support. Our conclusion is that a change from mainly internal processing to greater dependency on external support represents a gradual increase in the use of controlled compared to automatic processing. Because increased reliance on controlled processes is gradual, there should be no dramatic difference between working with and without external representations.

REFERENCES

Anderson, R.E., & Helstrup, T. (1989). Visual discovery in mind and on paper. Paper presented at the 1989 meeting of the Psychonomic Society in Atlanta, Georgia.

Chambers, D., & Reisberg, D. (1985). Can mental images be ambiguous? *Journal of Experimental Psychology: Human Perception and Performance* **11**, 317-328.

Farah, M.J. (1984). The neurological basis of mental imagery: A componential analysis. *Cognition* **18**, 245-272.

Finke, R.A. (1989). *Principles of mental imagery*. Cambridge, Mass.: MIT Press.

Finke, R.A., & Slayton, K. (1988). Explorations of creative visual synthesis in mental imagery. *Memory & Cognition* **16**, 252-257.

Fodor, J.A. (1975). *The language of thought*. New York: Crowell.

Helstrup, T. (1988a). Imagery as a cognitive strategy. *In* M. Denis, J. Engelkamp & J.T.E. Richardson (Eds.), *Cognitive and neuropsychological approaches to mental imagery*. Dordrecht: Nijhoff.

Helstrup, T. (1988b). The influence of verbal and imagery strategies on processing of figurative language. *Scandinavian Journal of Psychology* **29**, 65-84.

Helstrup, T. (1989). Active and passive memory: States, attitudes, and strategies. *Scandinavian Journal of Psychology* **30**, 113-133.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition* **18**, 1-63.

Reisberg, D. (1987). External representations and the advantages of externalizing one's thoughts. *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Seattle, Washington. Hillsdale,NJ: Erlbaum.

Tvedt, K.S., & Wilhelmsen, S. (1989). Mental forestilling. En undersøkelse om forestillingsaktivitet under ulike forsøksbetingelser. *Term paper. Department of Cognitive Psychology*, University of Bergen, Norway.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 17

Understanding spatial descriptions: Test of a mental-capacity model*

Sergio Morra
Università di Padova, Italy
and
Juan Pascual-Leone, Janice Johnson, Raymond Baillargeon
York University, Toronto, Canada

This paper reports an experiment on a spatial descriptions task, originally designed by Ehrlich and Johnson-Laird (1982) to study the integration of sentences in a comprehensive representation. Items in the original studies all comprised three sentences, describing the positions of four objects. The authors concluded that subjects integrated the spatial descriptions in the form of a "spatial mental model", rather than a network of propositions. According to Johnson-Laird (1983), a spatial mental model is an analogue representation, consisting of a finite set of tokens and a finite set of spatial relations between them; the relations are represented by mentally locating the tokens within a dimensional space.

Morra (1989a) modified the task for a study involving both children, aged 6 to 9, and adults. To avoid floor levels of performance, descriptions comprising two sentences (referring to three objects) were used with the children. The result that adults follow a "spatial mental model" strategy was replicated. Children, however, were found generally to follow a different strategy: They tried to store the sentences verbatim, and then to recall and overtly represent them one at a time. Children's performance was predicted well by a probabilistic model, which

*This research was supported by a grant from the Consiglio Nazionale delle Ricerche (Comitato Scienze Biologiche e Mediche) to S.Morra, by a Social Sciences and Humanities Research Council of Canada research grant to J.Pascual-Leone and by a York University Faculty of Arts research grant awarded to J.Johnson. Thanks are also due to Josie Lumia for help in collecting and scoring part of the data.

expressed the probability of success in the task as a function of subjects' ability to remember the sentences and their ability to represent a single sentence. This model largely underestimated the performance of adults, however.

Children and adults clearly were following different strategies, yet the source of this difference remained unclear. One might suggest that young children are unable to construct spatial mental models. This suggestion is ruled out, however, by research on a task involving the planning of a drawing (Morra, Moizo & Scopesi, 1988). This research showed that children aged 6 to 11 can construct mental models, although limited attentional capacity prevents 6-year olds from including more than about three positions in a spatial mental model.

Morra (1989a) speculated that age (i.e, the development of working memory or attentional resources) and the number of sentences presented may interact in determining the workload imposed by each strategy. This would be consistent with Pressley, Borkowski and Johnson's (1987) suggestion that working memory capacity may be an important determinant of whether children benefit from an imagery strategy in tasks involving memory for prose. More specifically, Morra (1988) suggested that Pascual-Leone's Theory of Constructive Operators (TCO: Pascual-Leone & Goodman, 1979) might provide a theoretical framework for the study of analogal thinking, and in particular, for modelling subjects' understanding of spatial descriptions. Empirical evidence for this claim is reported in the present chapter.

A detailed account of the TCO is beyond the scope of this paper (rather, see Pascual-Leone, 1970, 1976, 1980, 1987; de Ribaupierre, 1983; for an account of language comprehension within the TCO framework, see Johnson, Fabian & Pascual-Leone, 1989). We mention briefly just two basic concepts of the theory, which are essential for understanding what follows. The first is the concept of "scheme", derived with a few modifications from Piaget's theory. Schemes are classified in various ways: for example, sensory-motor vs. representational schemes, executive vs. action, operative vs. figurative; some of these distinctions, not discussed here in detail (see Pascual-Leone, Goodman, Ammon, & Subelman, 1978) enter into our task analysis. As a rough but acceptable approximation, in the present context, the reader may understand "schemes" as "chunks" or "units of information", with the warning that both *operative* (i.e., transformation-representing, procedural) and *figurative* (i.e., state-representing, declarative) schemes are considered relevant as information units.

The TCO assumes that schemes can have various sources of activation, such as perceptual saliency, overlearning, mental attention, etc.; these sources are thought to reflect mechanisms called "silent hardware operators". One of these, labelled "M Operator" (M, for short), is the second essential concept we introduce here. It is conceived as a limited amount of mental attentional energy, that can be allocated to activate relevant schemes. The field of schemes being boosted by M, at a given moment, constitutes the current M-centration. Note that other highly activated schemes may coexist, which are not part of the M-centration. They may be boosted by other mechanisms (e.g., perceptual saliency, overlearning, affective factors). Previously activated schemes, that remain outside M-centration, tend to decay in their activation due to attentional interruption: In each act of attention (M allocation) schemes that are not being centrated tend to be inhibited automatically (Burtis, 1982; Pascual-Leone, 1983, 1984). Because of the attentional interruption mechanism, attentional effort (M allocation) often must be applied to keep activated the task relevant schemes.

The formula **Mp = e+k** represents the **power** (**p**) or capacity of a subject's **M** operator. Here **e** stands for a constant amount of energy used to keep active an executive scheme for the task; **k** represents a quantitative developmental parameter, namely the additional number of representational (figurative or operative) schemes the subject can simultaneously activate with attentional energy. For an account of how the capacity of this attentional mechanism increases with age, see Pascual-Leone (1970). The formula **Md = e+k** means that, for successfully performing a mental operation, an **M** capacity of power **e+k** is **demanded** (**d**).

Two issues motivated the current research. The first was to answer a question raised by Morra's (1989a) results: that is, when children aged 9 to 14 perform the spatial task with both two- and three-sentence descriptions, which strategy will they follow? The second issue was of a more theoretical nature: Can the TCO successfully model subjects' understanding of spatial descriptions?

We describe one experiment in the next section, followed by a detailed analysis (based on the TCO) of the mental model strategy. Finally, results are reported, keeping in mind the two questions posed above.

METHOD

Subjects were 72 children from a primary school in a suburb of Toronto. There were 25 children from each of grades four and six, and

22 from grade eight; mean ages were 9:6, 11:11, and 13:11 respectively. All subjects were fluent in English.

Testing was conducted over three sessions. During the first session, subjects received the Figural Intersections Test (FIT), a group administered paper-and-pencil measure of **M** capacity. Each FIT item requires the subject to locate the one area of intersection of a set of overlapping geometric shapes. The number of shapes in an item varies from two to eight, and defines the class of the item. The **M** capacity score corresponds to the highest class the subject passes reliably (i.e. 75% correct), given a reliable level of passing on lower class items (for more details see Pascual-Leone & Ijaz, 1989; Johnson & Pascual-Leone, 1989).

During the second session, subjects performed the spatial descriptions task (preliminary items and brief descriptions) and a task of memory for pairs of sentences. The apparatus for the spatial task consisted of a four-layer set of shelves, 31 cm high, with a dark grey plastic back and four transparent shelves. Each shelf (17.5 cm wide and 32 cm long) was divided into four zones by means of parallel coloured stripes.

There were eight *preliminary items*, each consisting of a single sentence. For each item, the experimenter placed one object on the apparatus, then handed another object to the child, and read out a sentence referring to the two objects (e.g., "The shell over the coin"). Subjects were given feedback on their performance and were allowed to make corrections, in order to make clear to them how to interpret the sentences. However, only responses prior to feedback on each item were scored (1 point per correct item).

The *brief descriptions* of the spatial task consisted of eight pairs of sentences (e.g., "The light over the fork. The light in front of the eraser."). Sentences were read at the rate of about one per 3.5 sec; then the objects were spread in front of the child, whose task was to place them in the correct spatial relationship on the apparatus. The score was the number of correct items, out of 8.

The *short-term memory* task consisted of eight pairs of sentences, quite similar to those that constitute the items of the main task. The experimenter read each pair of sentences, and the subject was required to repeat them (but not to place the objects). A subject's score was a triplet of digits, that is, the number of items on which two, one, or no sentences were correctly recalled. This scoring was used, because the probabilities of recalling two, one, or zero sentences enter as parameters in Morra's (1989a) probabilistic model.

In the third session, after a brief warm-up on the spatial task, subjects received 12 three-sentence items of the spatial descriptions task.

There were four of each of three types of item: *Continuous, Semicontinuous* and *Discontinuous. Continuous* refers to the order of appearance of the object-terms in the sentences: A and B in the first, B and C in the second, C and D in the third; for example, "The paperclip (A) over the candle (B). The stone (C) in front of the candle (B). The dice (D) under the stone (C)." In *Semicontinuous* descriptions, the first two sentences are shifted, and the order becomes B-C, A-B, C-D (with no common object between the last two sentences). In *Discontinuous* descriptions, the last two sentences are shifted, and the order becomes A-B, C-D, B-C (with no common element between the first two sentences). Each type of description was scored separately by counting the number of correct items, out of four.

Finally, three-sentence descriptions were presented for a *short term memory* task. There were four Continuous and four Semicontinuous items; Discontinuous descriptions were not included, since Morra's (1989a) probabilistic model cannot be extended to them. For each type of item, a subject's score was a vector representing the number of cases in which three, two, one or zero sentences were recalled.

Two subjects (one from each of grades six and eight) did not complete session three due to school absence. The analyses of data from this session are based on 70 subjects.

TASK ANALYSIS

In the main task used in our experiment, subjects listen to descriptions consisting of either two or three sentences, referring to the positions of either three or four objects, such as "X is under Y. Z is behind X." Then, they receive those objects and place them on the shelves, according to the description. We present here a step-by-step task analysis, listing the schemes that should be activated by the **M** Operator at each step, if a subject is trying to construct a mental model.

Let us suppose that the first sentence is "The shell behind the eraser". At step I, a subject has to parse the sentence into three constituents, and no more than three linguistic schemes (presumably, articulatory representations of the words) must be activated by the **M** Operator, that is, 'the shell', 'behind', 'the eraser'[1]. The load on the **M** Operator may actually be less than **e+3**, because of activation from the perceptual input.

[1] Single quotes are used in this section to refer to distinct linguistic (articulatory) schemes in the subject's mind. Double quotes are used just to quote the descriptions.

At step II, the relational term 'behind' is interpreted by the subject, who activates a figurative scheme for the subject of the sentence, "the shell", and an operative scheme (i.e., an analogue representation of "behind"). It is not essential for our analysis whether the figurative scheme activated for the shell is a detailed image, a generic mental model, or a linguistic label. What is important is that the operative scheme is applied to the figurative scheme; that is, a token representing the shell is placed "behind" in the mental model. At this point, a linguistic scheme 'the eraser', which the subject must not forget, is still activated by the M operator: thus, **Md = e+3**.

At step III, the load is still **e+3**: the eraser is included in the mental model, i.e. the operative scheme for "behind" is applied onto the figurative schemes representing (in the appropriate order) the shell and the eraser. There are hints in our experiments which suggest this order of steps II and III. However, if a subject interprets 'behind', 'the eraser' at step II and 'the shell' at step III, the demands placed on the M Operator remain unchanged.

At step IV, the subject must hold active the positions of the shell and the eraser (in the TCO, this is described as having learned a temporary structure for the meaning of "behind" in the item context). At this point, **Md = e+2**.

At step V, the subject receives another sentence, such as "the eraser over the knife." This also must be parsed into three constituents, as the first sentence in step I. The load or demand on the **M** Operator is no more than **e+5**; namely the step executive, the figurative schemes for the positions of the shell and the eraser in the spatial model, and the three constituents of the new sentence.

At step VI, the subject must detect which constituent of the new sentence refers to which token in the mental model: for example, that 'the eraser' in the second sentence refers to the token already placed in the front position of the mental model. The **M** demand remains at **e+5**: the representations of the *shell* and the *eraser* in the spatial model, a mental *pointer* to the eraser (as the first term of a new relation, that is as the startpoint of the path for adding new elements to the model), and a string of constituents of the second sentence, such as Ø, 'over', 'the knife', where Ø stands for the subject of the sentence (already interpreted as a pointer to the given token and, therefore, no longer needing to be held active as a linguistic scheme).

At step VII 'over' is interpreted (in the same way as in step II). The mental pointer introduced at step VI is used to keep track that the eraser is over something (rather than something over the eraser). The load at this point becomes **e+4**, that is, the step executive, the positions of the shell and the eraser, an analogue representation of "over", and a linguistic scheme for 'the knife'.

At step VIII a figurative scheme for the knife is included in the mental model, in the same way as at step III; **Md** remains at **e+4**.

At step IX the subject merely holds active the positions of three objects (cf. step IV), and **Md = e+3**.

If the description is a *brief* one, there is nothing more to do in order to understand it. The highest workload placed on the **M** Operator was **Md = e+5** (at step VI); it thus follows that **e+5** is the **M** capacity required to perform a strategy of mental model construction with a *brief description*.

However, if there is a third sentence in the description, for example "The shell under the screw", then more processing steps are required, conceptually equivalent to steps V to IX, with the difference that **Md** is increased by one unit, because at this point there are already three tokens instead of two in the mental model. Thus, in order to construct a spatial mental model of a *continuous* or *semicontinuous* description, **Md = e+6**.

If the description is a *discontinuous* one, however, two distinct mental models (each consisting of two tokens) must be constructed for the first two sentences. When a subject listens to the third sentence, which relates one term of the first to one term of the second, the load on the **M** Operator is **Md = e+7**, for reasons similar to those given at step VI, except that there are already four tokens instead of two in the spatial models.

Having enough **M** capacity is a necessary, but not sufficient condition for successful performance in a task. It is not sufficient, therefore, that a subject have **Mp = e+5**, or **e+6**, or **e+7** (according to the different types of descriptions). Adequate executive and other relevant schemes, enough time to process each sentence, and adequate motivation are also necessary. If these conditions are met, then a subject can successfully follow the strategy of mental model construction.

As mentioned above, an alternative strategy for task solution is based on verbatim storage of the sentences. We do not present a detailed task analysis of this strategy here (for TCO analyses of verbal short-term memory tasks, see Burtis, 1982; Morra, 1989b). Our analysis suggests, however, that *consistent* successful implementation of the verbatim strategy carries **Md = e+6** for *brief* descriptions and **Md = e+9** for the three-sentence descriptions. That is, the **M** demand of the verbatim strategy exceeds that of the mental model strategy. The verbatim strategy has one important advantage, however. Subjects may occasionally succeed with less **M** capacity, if they can retrieve verbal information which is not being M-centered, but has remained sufficiently activated after attentional interruption.

It follows that the spatial mental model strategy is generally preferable. Subjects who do not have enough M capacity to follow it, however, might resort to a less sophisticated strategy based on verbatim storage. Fallback to the verbatim strategy should be more likely in the *brief* descriptions, where such a strategy occasionally could be successful, depending on one's ability to retrieve partially decayed sentences.

RESULTS AND DISCUSSION

Initial analyses were aimed at estimating the strategies used by subjects at various ages. The probabilistic model (Morra, 1989a), which embodies the verbatim storage strategy, was fitted to the observed scores in 3-sentence descriptions (*continuous* and *semicontinuous*). The fit was not good (see Table 1).

Table 1. Fit of the probabilistic model of verbatim strategy to subjects' performance with continuous and semicontinuous descriptions.

	Observed	Expected	
Grade 4			
Mean	0.560	0.393	$t(24)=0.81$ n.s.
Variance	0.966	0.029	$X^2(24)=829.2$ $p<.001$
			$r(23)=-.10$ n.s.
Grade 6			
Mean	1.125	0.554	$t(23)=1.92$ $p<.07$
Variance	1.776	0.095	$X^2(23)=450.2$ $p<.001$
			$r(22)=-.20$ n.s.
Grade 8			
Mean	1.190	0.674	$t(20)=2.34$ $p<.03$
Variance	1.107	0.307	$X^2(20)=75.6$ $p<.001$
			$r(19)=.37$ $p<.05$

Note. Expected scores were computed for each subject from a probabilistic model (Morra, 1989a) which embodies the verbatim storage strategy. If the observed and expected means are not significantly different, a non-significant value of t should be obtained. If the observed and expected variances are not significantly different, a non-significant Chi-square should be obtained. If the individual expected and observed scores are correlated, a positive and significant value of r should be obtained.

Results indicated that subjects at all ages performed better than would be expected from the verbatim strategy (although only in the oldest age group was the difference significant). The variance of scores was much larger than expected, and expected and observed scores were generally uncorrelated. These results rule out the hypothesis that subjects used a verbatim storage strategy to solve the three-sentence problems.

What sort of evidence would indicate use of spatial mental models to understand the 3-sentence descriptions? The performance pattern expected and found by Ehrlich and Johnson-Laird (1982), and taken by them as evidence of mental model construction by adult subjects, was as follows: *discontinuous* < *continuous* = *semicontinuous*. We evaluated this pattern for our young subjects.

The mean numbers of passed items per grade and type of description are reported in Table 2. A mixed two-way analysis of variance yielded main effects for age ($F(2,67)=4.02$, $p<.03$) and descriptions ($F(2,134)=3.16$, $p<.05$); the interaction was not significant ($F(4,134)=1.02$, $p>.40$).A planned comparison between *continuous* and *semicontinuous* problems, yielded ($F(1,69)<1$); another planned comparison, orthogonal to the first, between *discontinuous* problems and the average of *continuous* and *semicontinuous*, yielded ($F(1,69)=6.50$, $p<.02$). The pattern observed in our subjects is thus the same as expected and observed in previous studies with adult subjects. These results provide evidence for mental model construction in children. Further, the absence of an age x description interaction suggests use of mental models in all age groups.

Table 2. Mean scores for continuous, semicontinuous and discontinuous three-sentence descriptions.

	Continuous	Semicontinuous	Discontinuous
Grade 4	0.28	0.28	0.08
Grade 6	0.67	0.46	0.25
Grade 8	0.52	0.67	0.52

This finding seems striking; but perhaps it is not surprising, given the differential in **M** demands for the two strategies: **Md = e+6** or **e+7** for the mental model strategy, versus **e+9** for verbatim storage. It appears that young subjects try to understand longer descriptions "the spatial way", and that although their performance is poor, the pattern is as it should be in the case of following a mental model strategy.

Table 3. Fit of the probabilistic model of verbatim strategy to subjects' performance with Brief descriptions.

	Observed	Expected	
FIT = 2 or 3			
Mean	2.154	2.241	$t(12)=-.031$ n.s.
Variance	3.976	1.915	$X^2(12)=27.0$ $p<.02$
			$r(11) = .90$ $p<.001$
FIT = 4			
Mean	3.333	2.904	$t(17)=1.03$ n.s.
Variance	2.333	2.682	$X^2(17)=15.6$ n.s.
			$r(16)=.41$ $p<.05$
FIT = 5			
Mean	4.687	3.663	$t(15)=1.99$ $p<.06$
Variance	2.590	2.426	$X^2(15)=17.1$ n.s
			$r(14)=.21$ n.s.
FIT = 6			
Mean	5.412	4.231	$t(16)=2.95$ $p<.01$
Variance	2.242	4.844	$X^2(16)=7.9$ $p<.10$
			$r(15)=.69$ $p<.002$
FIT = 7 or 8			
Mean	5.500	4.892	$t(7)=0.74$ n.s.
Variance	2.500	5.753	$X^2(7)=3.5$ n.s.
			$r(6)=.47$ n.s.

Note. The meaning of the statistical tests is the same as in Table 1.

Our task analysis suggested that the capacity demand of the mental model strategy in *brief* descriptions is **Md = e+5**. To test the hypothesis of a transition into the mental model strategy at **e+5**, we grouped the subjects according to their **M** Capacity (measured by the FIT score). We then fitted the probabilistic model to the observed *brief* descriptions scores within each group of subjects[2]. Results are shown in Table 3.

The model seems to fit well the data of the two groups with a FIT score of 4 or less (except for the variance of the first group), whereas the two groups with a FIT score of 5 or 6 show several signs of a bad fit. The small sample size in the last group (FIT of 7 or 8) may be insufficient for doing a proper test of fit, or these subjects may have used a verbatim strategy (i.e., they have sufficient **Mp** for doing so successfully).

Further evidence for a crucial change in how subjects with **M = e+5** (or more) understand the *brief* descriptions comes from a contingency table, in which subjects were classified according to FIT score and success in *brief* problems. It was reasoned that good performance may be possible for subjects with **M < e+5**, if they are able to retrieve decaying sentences; but that a large improvement should be found at **Mp = e+5**, because this is the **M** demand of the mental models strategy for *brief* problems.

We took six problems out of eight as the criterion of success. The number of subjects who pass or fail the task according to this criterion is reported in Table 4.

If subjects are divided in only two categories of FIT score (i.e., less than 5 versus 5 or more), the comparison of pass/fail incidence between these categories is significant: $X^2 = 16.27$, $p<.001$. But, more important, the Chi-squares or Exact Probability tests on the adjacent FIT groups (i.e., 2-3 vs. 4, 4 vs. 5, etc.) are all non-significant, with the single exception of the comparison between subjects with **M = e+4** and **M = e+5**, in which case $X^2 = 4.91$ (with Yates' correction), $N=34$, $p<.04$. These results support our hypothesis of a performance transition at **Mp = e+5**.

As expected, performance in the Spatial Descriptions Tasks and in the FIT were positively correlated. The correlation between FIT and Brief descriptions was .56 (.52 with age partialled out, $p<.001$). The correlation between FIT and total score in the three-sentence descriptions was .43 (.35 with age partialled out, $p<.001$). One might argue that these

[2] We also did a similar analysis grouping the subjects by grades, just as we did for the three-sentence descriptions (Table 1), instead of grouping them by M capacity. The fit of the probabilistic model was good at grades 4 and 8, but not at grade 6. This is also consistent with theoretical predictions, since the normal M capacity at the age of eleven is e+5.

relations could be due to reasons other than **M** capacity being necessary for the Spatial Descriptions Task (e.g., they could be due to the spatial content of all the tasks). As always, correlational data do not provide conclusive evidence. Although such an argument cannot be ruled out completely, an interesting detail helps with the interpretation of the correlational data. We observed that the relationship was non-linear between three-sentence descriptions (in which subjects tended to follow the mental model strategy) and FIT. On the contrary, the average scores in the spatial task remained quite low for the groups of subjects with FIT-score < 6 (i.e., 0.82 for FIT = 2 or 3; 0.39 for FIT = 4; 1.00 for FIT = 5) and suddenly increased among subjects with **M** Capacity of **e+6** or more (i.e., 2.19 for FIT = 6; 2.13 for FIT = 7 or 8). This detail is relevant, because **e+6** is the **M** capacity assumed necessary for successful implementation of the mental model strategy in three-sentence descriptions. Although the scores in this spatial task are generally low, these means suggest a genuine relation between subjects' **M** capacity and their performance on the spatial task.

Table 4. Number of subjects, in each FIT-score group, who pass or fail the Brief descriptions (criterion for success: at least six passed items out of eight)

	FIT score				
	2-3	4	5	6	7-8
pass	1	1	7	9	5
fail	12	17	9	8	3

We intended to run statistical analyses on contingency tables, similar to those reported in Table 4, in order to verify whether there was a significant improvement in *continuous* and *semicontinuous* problems at **M = e+6** and in *discontinuous* problems at **M = e+7**. But performance was at floor (i.e., too few correct responses) thus preventing our doing such analyses. From this point of view, it would be advisable to include an adult sample, and a slower rate of presentation, in a replication study.

Pending replication on adult subjects, we conclude at this point that the experiment gives promising evidence in support of our model, based on the Theory of Constructive Operators, of performance in the Spatial Descriptions Task.

REFERENCES

Burtis, P.J. (1982). Capacity increase and chunking in the development of short-term memory. *Journal of Experimental Child Psychology* **34**, 387-413.

Ehrlich, K., & Johnson-Laird, P.N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior 21*, 296-306.

Johnson, J., Fabian, V., & Pascual-Leone, J. (1989). Quantitative hardware-stages that constrain language development. *Human Development* **32**, 245-271.

Johnson, J., & Pascual-Leone, J. (1989). Developmental levels of processing in metaphor interpretation. *Journal of Experimental Child Psychology* **48**, 1-31.

Johnson-Laird, P.N. (1983). *Mental models.* London: Cambridge University Press.

Morra, S. (1988). A quantitative approach to the complexity of analog thinking. *In* C. Cornoldi (Ed.) *Imagery and cognition* (Proceedings of the Second Workshop on Imagery and Cognition, pp. 257-266). Padova: Dipartimento di Psicologia Generale.

Morra, S. (1989a). Developmental differences in the use of verbatim versus spatial representations in the recall of spatial descriptions: A probabilistic model and an experimental analysis. *Journal of Memory and Language* **28**, 37-55.

Morra, S. (1989b). Verbal short term memory and the word- length effect: A neo-Piagetian reappraisal. *In* J. Johnson (chair), *Mental capacity constraints on cognition: Studies from the Theory of Constructive Operators.* Symposium conducted at the tenth meeting of the International Society for the Study of Behavioral Development, Jyvaskyla, Finland.

Morra, S., Moizo, C., & Scopesi, A. (1988). Working memory (or the M operator) and the planning of children's drawings. *Journal of Experimental Child Psychology* **46**, 41-73.

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologica* **63**, 301-345.

Pascual-Leone, J. (1976). On learning and development, Piagetian style. *Canadian Psychological Review* **17**, 270-297.
Pascual-Leone, J. (1980). Constructive problems for constructive theories: the current relevance of Piaget's work and a critique of information-processing simulation psychology. *In* R.Kluwe & H.Spada (Eds.), *Developmental models of thinking*, pp.263-296. NY: Academic Press.
Pascual-Leone, J. (1983) Growing into human maturity: Towards a metasubjective theory of adulthood stages. *In* Baltes P. & Brim O. (Eds.), *Life-span development and behavior vol.5*, pp. 118-156. NY: Academic Press.
Pascual-Leone, J. (1984) Attention, dialectic, and mental effort: Towards an organismic theory of life-stages. *In* Commons M., Richards F. & Armon C. (Eds.), *Beyond formal operations: late adolescent and adult cognitive development*, pp. 182-215). NY: Prager.
Pascual-Leone, J. (1987). Organismic processes for neo- Piagetian theories: A dialectical causal account of cognitive development. *International Journal of Psychology* **22**, 531-570.
Pascual-Leone, J., & Goodman, D. (1979). Intelligence and experience: A neo-Piagetian approach. *Instructional Science* **8**, 301-367.
Pascual-Leone, J., Goodman, D., Ammon, P., & Subelman, I. (1978). Piagetian theory and neo-Piagetian analysis as psychological guides in education. *In* J. Gallagher & J.A. Easley (Eds.), *Knowledge and development vol.2*, pp. 243-289. NY: Plenum.
Pascual-Leone, J., & Ijaz, H. (1989). Mental capacity testing as a form of intellectual-developmental assessment. *In* R. Samuda, S. Kong, J. Cummins, J. Pascual-Leone & J. Lewis (Eds.), *Assessment and placement of minority students*, pp.143-171. Toronto: C.J. Hogrefe.
Pressley, M., Borkowski, J.G, & Johnson, C. (1987). The development of good strategy use: Imagery and related mnemonic strategies. *In* M.A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications*, pp. 274-297. NY: Springer-Verlag.
de Ribaupierre, A. (1983). Un modele neo-Piagetien du developpment: la theorie des operateurs constructifs de Pascual-Leone. *Cahiers de Psychologie Cognitive* **3**, 327-356.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 18

What goes on in the mind when we solve syllogisms?

Norman E. Wetherick
University of Aberdeen, United Kingdom

References to Aristotle are not commonly found in contemporary studies of the psychology of thinking, yet his formulations correspond quite closely to the position taken today by many cognitive psychologists. In De Memoria he writes:

> "We cannot think without imagery, for the same thing occurs in thinking as is found in the construction of geometrical figures. There though we do not employ as a supplementary requirement of our proof a determinateness in the size of the triangle, yet when we draw it we make it of a determinate size. Similarly in thinking also, though we do not think of the size, yet we present the object visually to ourselves as a quantum, though we do not think of its quantum." (quoted in Humphrey, 1951, p.31.)

It appears that Aristotle like most of us (not all) experienced visual imagery whilst thinking, but he did not confuse the imagery with the thinking. His Associationism was more sophisticated than many later varieties. Aristotle was also the first to formulate a theory of syllogistic reasoning (in the Prior Analytics) and it is clear that here too he started from a consideration of reasoning in mathematics. For those unfamiliar with the syllogism, definitions of the technical terms employed are given in Table 1.

Each pair of premises constitute the axioms of a miniature deductive system from which valid consequences may follow, as in geometry. Aristotle would presumably have taken it for granted that imagery accompanied syllogistic thinking (since in his view it accompanied all thinking). Cognitive psychology has taken a similar line. It postulates "cognitive representations", usually in the form of "models", on which "cognitive operations" may be performed. The representations may be thought to correspond to Aristotle's "imagery" and the mental operations, to his "thinking". Aristotle did not however set himself the task of explaining "thinking". Cognitive psychologists tend to insist that they have explained "thinking" once suitable cognitive representations have been proposed though, to my mind, the "thinking" is done by (e.g.) the "manipulator" of Johnson-Laird's mental models (Johnson-Laird, 1983) or by Baddeley's "central executive" (Baddeley, 1986), about which the theories have little to say.

Table 1. The structure of the syllogism

Propositional Forms

Universal Affirmative	"A"	"All X is Y"
Particular Affirmative	"I"	"Some Y is Z"
Universal Negative	"E"	"No Z is X"
Particular Negative	"O"	"Some Y is not Z"

The Four Figures (f) of the Syllogism

	f.1	f.2	f.3	f.4
Premises	Y-Z	Z-Y	Y-Z	Z-Y
	X-Y	X-Y	Y-X	Y-X
Conclusions	X-Z	X-Z	X-Z	X-Z

f.1: X appears as subject of one of the premises and Z as predicate of the other.
f.2: Both X and Z appear as subject of their premise.
f.3: Both X and Z appear as predicate of their premise.
f.4: X appears as predicate of one of the premises and Z as subject of the other.

N.B. Y is always the middle term and does not appear in the conclusion. X is always the subject and Z the predicate term of the conclusion but X, Y and Z may appear as either subject or predicate term of a premise. The conclusion may or may not follow logically from the premises. The order of premises is logically immaterial.

What is clear is that it is now once again permissible to suppose the existence of potentially introspectible mental contents, though theorists usually cover themselves by allowing that the models etc. may be outside consciousness. Johnson-Laird, for example, writes that "Reasoners construct mental models of the states of affairs described by premises. These models may take the form of vivid images or they may be largely outside conscious awareness." (Johnson-Laird & Bara, 1984, p.29.)

It is however difficult to see how mental operations like those postulated by Johnson-Laird ("construction" of composite mental models; "evaluation" of putative conclusions against the models etc.) could be performed, except in consciousness.

In Soar (his "unified theory of human cognition"), Newell takes a similar line. On the subject of "comprehension" he writes "Comprehension (construction of the situation model) occurs in the 'comprehend' problem space by applying a 'comprehension operator' to each incoming word." (Lewis, Newell & Polk, 1989). Whatever we are supposed to understand by this, it is clear that Newell postulates a problem space and operators which must in part at least be accessible to introspection. Though he does not mention consciousness, the "incoming words" must surely be in consciousness as they are heard or read; the "situation model" must also be in consciousness since information not contained in the premises can be derived from it.

There would seem to be scope for direct introspective studies of the cognitive process to see whether the mental objects and events postulated in cognitive psychology can be detected in conscious awareness. If not, then the psychologists can take refuge in a (strictly non-Freudian) unconscious; but they may then legitimately be required to postulate only mental operations that do not require intervention by a more remote "manipulator" or "executive" acting like a conscious decision-taker.

For a period of about five years at the beginning of this century introspective investigations of this kind were carried out in Würzburg. The moving spirit was the philosopher Külpe and the work appears to have ceased when he moved to a more prestigious chair in Bonn (in 1909). Külpe himself participated as subject in many of the published studies. (Almost none of this work is available in English. I rely on Chapters II, III, IV of Humphrey, 1951). The Wurzburgers had no immediate intellectual progeny. On the continent of Europe their credit was undermined by a vicious attack from Wundt who held very strongly to the view that introspection was not applicable to the study of the higher mental processes. In the USA their principal finding was hotly contested by Titchener. The finding was that there existed "imageless" thoughts, thoughts without associated quasi-sensory content. This would have come as no surprise to Aristotle but, at the end of the nineteenth century, Associationism insisted that thought must depend on images that had quasi-sensory content. In Titchener's view the Wurzburgers had done their introspecting badly, they had simply failed to notice the images associated with what they called "imageless" thoughts, usually because they were looking for visual or auditory images and the images that were present were kinaesthetic in nature. The controversy was however overtaken by events. Within a few years, behaviourism had achieved dominance in academic psychology. Introspection was abandoned as an experimental technique but the Wurzburgers had made their contribution. They had established that although, when performing a mental operation, there are often (relevant or irrelevant) images present to our awareness, there is usually also awareness of a mental event that has no sensory content (an imageless thought) and sometimes an operation may be performed with nothing present to our awareness except the imageless thought!

The Wurzburgers employed a range of tasks and in each case the subject performed the task and then reported what his mental contents had been whilst doing so. The Würzburg procedure must be distinguished from verbal protocol analysis. In protocol analysis the object is to enable the experimenter to follow the chain of reasoning of the subject over a period, in a task such as crypt-arithmetic (Newell & Simon, 1972) that can only be performed by a sequence of operations. It must also be distinguished from the philosopher's technique of phenomenological analysis. This is applied by the phenomenologist to his own thought processes as a means of coming to understand "things in themselves", i.e. "things as they really are". (Or, in American phenomenological psychology, to delineate states of consciousness "as they really are".) Würzburg introspection was intended to answer psychological questions about the nature of thinking. The intellectual climate of the time assumed that potentially introspectible mental contents might be involved in thinking; such a climate now exists again.

It surprised me to discover that the Wurzburgers had not employed problems of syllogistic reasoning. There is no obvious reason why a pair of premises should not be presented and subjects required to draw a conclusion (or decide that no conclusion can be drawn) and then report on their mental contents whilst doing so. That is what was done in the two studies to be reported in this paper.

It was Würzburg practice to employ a very few expert subjects, on the assumption that evidence gained from one or two such subjects would be immediately generalisable. There may be a sense in which this is true but Galton (1883; see also Hadamard, 1945) had already shown

Table 2. The twenty syllogisms employed in the studies. Subjects in the first study did all twenty in the order indicated by the numbers in parentheses, subjects in the second study did either Set a. or Set b.

			Set a				Set b
(1)		A	All Y's are Z	(5)		A	All Z's are Y
	f.1	I	Some X's are Y		f.2	A	All X's are Y
		I	Some X's are Z			-	NVC
(2)		A	All Y's are Z	(6)		I	Some Z's are Y
	f.3	E	No Y's are X		f.2	A	All X's are Y
		-	NVC			-	NVC
(3)		A	All Z's are Y	(7)		I	Some Z's are Y
	f.2	I	Some X's are Y		f.4	A	All Y's are X
		-	NVC			I	Some X's are Z
(4)		E	No Z's are Y	(10)		A	All Z's are Y
	f.4	A	All Y's are X		f.4	I	Some Y's are X
		O	Some X's are not Z			-	NVC
(8)		A	All Y's are Z	(11)		A	All Y's are Z
	f.3	I	Some Y's are X		f.1	A	All X's are Y
		I	Some X's are Z			A	All X's are Z
(9)		A	All Z's are Y	(12)		E	No Y's are Z
	f.2	E	No X's are Y		f.1	A	All X's are Y
		E	No X's are Z			E	No X's are Z
(15)		A	All Z's are Y	(13)		I	Some Y's are Z
	f.4	E	No Y's are X		f.3	A	All Y's are X
		E	No X's are Z			I	Some X's are Z
(16)		I	Some Y's are Z	(14)		E	No Z's are Y
	f.1	E	No X's are Y		f.2	A	All X's areY
		-	NVC			E	No X's are Z
(18)		I	Some Y's are Z	(17)		E	No Y's are Z
	f.1	A	All X's are Y		f.3	A	All Y's are X
		-	NVC			O	Some X's are not Z
(19)		E	No Y's are Z	(20)		A	All Y's are Z
	f.1	I	Some X's are Y		f.1	E	No X's are Y
		O	Some X's are not Z			-	NVC

NVC=No Valid Conclusion

that introspectible mental contents vary markedly from person to person and it follows that theories asserting that mental contents of one kind or another play a crucial role in thinking can only be tested satisfactorily over a group of subjects; one possible outcome is that such contents play a crucial role in some subjects but not in others. The studies reported here used groups of subjects who were not expert in any sense.

STUDY 1

A group of thirty third-year undergraduate students of psychology participated in the first study (21 female, 9 male; aged 19 to 25 years). They had not yet taken any course covering contemporary theories of syllogistic reasoning and had not done any private reading in the area; they appeared to have no preconceptions about what results might be obtained or about the bearing that any result might have on a particular theory.

Materials

A set of twenty pairs of syllogistic premises was prepared using X, Y and Z as terms. In each case Y was the middle term and the subject was instructed that his conclusion (if any) should have X as subject. (See Wetherick (1989) and Wetherick & Gilhooly (1990) for discussions of the basis of choice of the term to be used as subject of a conclusion.) Table 2 shows that eight pairs of premises presented an A and an I premise: of which four had a valid I conclusion and four had no valid conclusion. Eight pairs presented an A and an E premise; of which four had a valid E conclusion, two had a valid O conclusion, and two had no valid conclusion. Two pairs presented two A premises, one of which (in syllogistic figure 1) had a valid A conclusion and the other (in syllogistic figure 2) had no valid conclusion. Two pairs presented an I and an E premise, one of which (EI in syllogistic figure 1) had a valid O conclusion and the other (IE in syllogistic figure 1) had no valid conclusion. Altogether, twelve pairs of premises out of twenty had a valid conclusion. (Definitions of A, E, I, and O and of "figure" may be found in Table 1.)

Method

Each pair of premises was written at the top of a blank sheet of paper. The subject was instructed to consider each pair, write underneath what conclusion s/he thought followed (or that s/he thought no valid conclusion followed) and, underneath that, record what had gone on in his/her mind while s/he was considering what the conclusion might be. The instructions emphasised that the subject should record everything that came into his/her mind whether or not it was relevant to the problem and that what was not required was the subject's theory of how such problems were or should be solved (unless such a theory had in fact come into his/her mind during the solution period). What was to be recorded was anything that came into the subject's mind between reading the premises for the first time and deciding on and writing down the conclusion. When s/he had recorded his/her introspections the subject was to draw a line under what had been written. Having gone through the series in this way, subjects were given access to the correct solutions and went through the series again. Where their solution was incorrect they were asked to convince themselves that the correct solution was correct and record what had gone on in their minds whilst doing so, as before.

Results

The subjects' performance was typical of Aberdeen University undergraduate subject groups. We have data from several such groups; using X, Y and Z as terms and using concrete terms (plumbers, bricklayers etc.) Where an A and an I premise are presented, about 90% of subjects draw a valid I conclusion if there is one. But about 40% also do so where there is no valid conclusion. Where an A and an E premise are presented about 90% draw a valid E conclusion. But 40% do so where there is only a valid O conclusion or where there is no valid conclusion at all (about 25% draw the valid O conclusion). Where two A premises are presented about 90% draw a valid A conclusion and 40% do so where there is no valid conclusion. Where an I and an E premise are presented about 60% draw a valid O conclusion when there is one and 60% say that there is no valid conclusion when there is none. I shall comment on the significance of this general pattern of results at a later stage. Here my purpose is only to point out that the results of the present group were typical, before going on to consider the introspections of individual subjects.

The subjects had been asked to report their mental contents even if they were irrelevant to the task in hand. Nineteen reported that they had no irrelevant mental contents, six that they had some and five that they had a great many. I find it difficult to believe that the majority had such great powers of concentration but there is no reference group with which to compare them. My own experience is that thinking always involves a desperate struggle to keep irrelevancies at bay. It is possible that my subjects had no such problem but more likely, I think, that they were not convinced by my assurances that I was interested in irrelevancies, or that they had not really grasped the idea of introspection at all. Several of them claimed to have used Venn diagrams. Something children are taught to call "Venn diagrams" is now part of the primary school syllabus in Great Britain but whenever I have asked a student to describe what s/he had been taught, the description has been of Euler circles; none of them appears to have encountered a genuine Venn diagram. I am at a loss to account for the fact that they have been taught Euler Circles but taught to call them Venn diagrams. Two things are however clear; most of them see the relevance of whatever it was they were taught to syllogistic reasoning and few of them understand the method well enough to derive any benefit from using it. Three subjects stated that their answers came "intuitionally" and were subsequently checked by "Venn diagrams". One subject said that he did not use "pictures" at all (I take him to have been referring to visual images) and complained that he ought to have been told that it might be helpful to do so. Several said that they had read the premises over and over again and repeated them "in their heads". Several complained that they found it "difficult to think this way" but they were referring to the difficulty of the logical task, not to the difficulty of recording their introspections. One subject wrote "It is hard to write thoughts that are taken for granted and difficult to express", she really appears to have been trying to introspect and may have been trying to describe an "imageless thought". One subject said he had "visualised populations of X's, Y's and Z's".

The subject population in this study was young and relatively unsophisticated. They were accustomed to using their minds but not to making their minds objects of study in their own right, not to the second order "thinking about thinking" called for by introspection. So not much can be learned from their observations.

STUDY 2

The twenty-eight subjects in the second study (6 female, 22 male; aged 22 to 40 years) were students on a postgraduate M.Sc. course in Artificial Intelligence. All were graduates (sometimes of several years standing) in computing science or a related discipline. The hope was that their greater sophistication would make the task of introspection easier for them than it had been for the undergraduate subjects. None of them had more than a passing acquaintance with syllogistic reasoning; most had studied some aspects of modern formal logic but contemporary textbooks tend to devote little if any space to the syllogism. They undertook the task as a mind-broadening exercise and were asked to do it themselves, persuade one other person (preferably from a different educational background) to do it, and compare the respective "mental contents". There was time to use only ten of the twenty pairs of premises, so each subject did either Set a or Set b (see Table 2).

Three subjects in the first study had reported that the conclusion came immediately, as if by intuition, although it was subsequently checked by reference to what they referred to as "Venn diagrams". More than half the subjects in the second study made similar reports. Virtually all of them reported visual imagery (often of intersecting regular or irregular surfaces) but most supposed, with Aristotle, that although the thought process might use images or be checked for accuracy against them it was essentially different from them. One or two reports are worth quoting at greater length.

One subject realised that a pair of premises is a conjunction of propositions that can be said to entail both premises. He repeated as his conclusion whichever premise contained X as its subject term. (If neither premise did, he said "No valid conclusion"!) The possibility of a syllogistic conclusion did not occur to him until his own subject produced them. Another noted that "One's experience during reasoning may be like a floating polystyrene block one minute, and the tip of an iceberg the next" which is an accurate characterisation of my own experience. A third subject worked verbally, repeating the premises over and over. She experienced an "inner voice which comments upon what I do, approving, disapproving, sometimes just repeating the words I have said or read". When she was not sure of her conclusion she used images of sets of X's, Y's and Z's. Her subject used "Venn diagrams" and she was so impressed that she repeated the task herself and found the diagram method much more effective. One subject used visual imagery throughout; illustrative Figure 1 is a computer graphic depiction of her mental contents while solving AA (in syllogistic figure 2) which has no valid conclusion. Her subject however substituted "big", "red" and "dogs" for X, Y and Z. But she did not visualise "big, red, dogs", she relied on an image of the sound of the words spoken out loud. One of the two commented "Pam thought in words, I thought in vague pictures. We understood, but found that we could not reason reliably using each other's method."

Most of the subjects in the second study achieved a high level of accuracy (7 to 10 correct out of 10) but the nature of their mental contents varied from person to person. The immediate objective (to show them that there are more ways of killing a cat than beating it to death with a computer terminal) was certainly achieved. For my present purpose, their maturity and their willingness to record what happened, however odd it might appear, paid handsome dividends. I obtained the kind of introspective data that had not been forthcoming from the undergraduate subjects in the first study. Maturity, consequent on being several years older (on average), was certainly a factor. Another may have been that they had never studied psychology. Psychology, at an elementary level, still encourages the feeling that there is something odd, probably illicit, about studying the mind by introspective methods and the undergraduate subjects had to try to overcome this feeling.

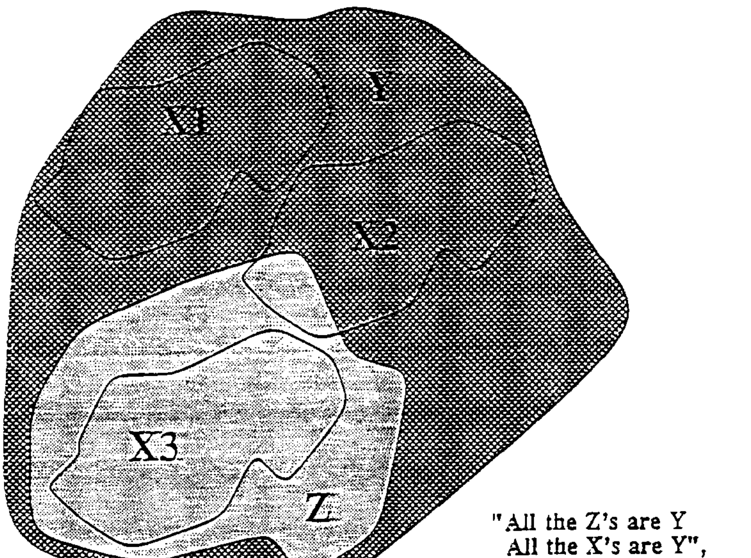

Figure 1. Quasi-visual mental contents experienced while solving syllogism No.5, AA- (See syllogistic figure 2 in Table 2). Both X and Z are wholly inside Y but X may or may not overlap with Z; three possible locations for X are illustrated - X1, X2 and X3. It follows that the syllogism has no valid conclusion. Computer graphic by Stephanie Swaine, reproduced by permission.

What goes on in the mind when we solve syllogisms is different in different people and sometimes it appears that nothing goes on at all! Two subjects at least, reported that their subject showed signs of irritation when pressed to report mental contents; there were none and why should there be?

GENERAL DISCUSSION

Earlier in this paper I remarked on the fact that although 90% of subjects draw the correct conclusion from pairs of premises having valid A, I or E conclusions, 40% draw the same type of conclusion from the same type of premises where no valid conclusion follows. From "All Y's are Z" and "Some X's are Y", 90% conclude correctly that "Some X's are Z". But from "Some Y's are Z" and "All X's are Y", 40% conclude incorrectly that "Some X's are Z"! This leads to the (wholly justified) suspicion that the 40% may not be drawing all their correct conclusions by processes of strict logic. Elsewhere I have described these subjects as "matchers" and characterised their behaviour in the experimental situation as choice of a conclusion matching in logical form the more conservative premise (i.e. the premise making an assertion about "some" rather than about "all" or "no" individuals of a certain kind) without reference to the underlying logical relationships between the terms. It happens that for many premise pairs (including most of those with valid A, I or E conclusions) matching gives the correct answer. But for some (including most of those with valid O conclusions) it does not. Examples may be seen in Table 2. Gilhooly and Wetherick (in preparation) have shown that two sub-groups of subjects may be distinguished that do equally well where matching gives the correct conclusion; however only one of them also does well where matching would have given the wrong conclusion. In the sub-group that does well on both, performance is highly correlated with ability to construct pairs of premises having a given valid conclusion ($r = 0.72$, $p<.01$) but in the other sub-group (where performance is poor on the second type of problem because many matching conclusions are chosen) there is also no correlation ($r = 0.10$, ns). I interpret this finding as evidence that the two sub-groups (who both do well on the first type of problem) are nevertheless succeeding for different reasons. The sub-group that was using logic also does well on the second type of problem, the sub-group using "matching" does not.

It is only possible to distinguish matchers from logicians by indirect methods and while it may be that subjects in the present study who claimed to be using verbal methods were in fact "matching", the data are insufficient to demonstrate the fact. (The syllogisms on which matching fails are also more difficult for subjects who are genuinely trying to use logic.) But the variety of introspective reports and the fact that absence of introspectible content is sometimes reported must cast doubt on any type of theory that relies on conscious mental representations. At most two subjects, for example, reported imaging sets of discrete individuals, although this is regarded by Johnson-Laird as an essential characteristic of a "mental model".

The hallmark of cognitive psychology is precisely that it insists on the existence of mental representations on which mental operations are performed. It would appear to me to be necessary to demonstrate the presence (and indeed the universality) of such representations rather than merely to postulate them but I have not seen any attempt to do so. Velmans (1989) has summarised experimental evidence on the function of consciousness in human information processing and provisionally concludes that it has not yet been shown to have a function. Either the information processing can be shown to proceed in the absence of consciousness or the consciousness can be shown to supervene after the information has been processed. Cognitive psychologists cover their retreat by supposing that the hypothesised models etc. may be outside conscious awareness but it is then legitimate to insist that no mental operations are postulated which require the intervention of an independent decision taker (a "manipulator" or "executive"). Theories that do not postulate such operations are beginning to emerge. Wetherick (1989) has proposed a theory of syllogistic reasoning that does not call for any kind of internal decision taker, supposing that valid conclusions resemble geometrical theorems in that they cannot be denied without self-contradiction if the premises (axioms) from which they follow are asserted. The verbal formulation of the conclusion may of course be in error. (It may in the extreme case be a deliberate lie!)

Counter-intuitive though is may appear, the conscious awarenesses that we so obviously possess may have no real function. The idea is not new. When mind was equated with consciousness, epiphenomenalism was always one of the proposed solutions to the mind/body problem. If it should happen that Velmans never needs to revise his provisional conclusion that "consciousness has not been shown to have a function", many mysteries will be resolved. Or it may turn out that consciousness can be shown to have only an adjuvant function and only in some people, some of the time. The function may be to reinstate an approximation to sensory input, in the absence of the real-world object that originally triggered the input. The resulting "consciousnesses" would be in principle inessential to the thought process involving the object (or some derivative of the object) and might be absent altogether or vary from person to person although the underlying (imageless) thought did not. That possibility would be consistent with the findings of many introspective studies and would not have surprised either the Wurzburgers or Aristotle.

REFERENCES

Baddeley, A. (1986) *Working Memory*. Oxford:Oxford University Press.

Galton, F. (1883) *Enquiries into Human Faculty*. London.

Gilhooly, K.J. & Wetherick, N.E. (in preparation) Drawing syllogistic conclusions and constructing syllogistic premises.

Hadamard, J. (1945) *The Psychology of Invention in the Mathematical Field*. Princeton University Press.

Humphrey, G. (1951) *Thinking*. London: Methuen and Co. Ltd.

Johnson-Laird, P.N. (1983) *Mental Models*. Cambridge: Cambridge University Press.

Johnson-Laird, P.N. & Bara, B. (1984) Syllogistic inference. *Cognition* **16**, 1-61.

Lewis, R.L., Newell, A. & Polk, T.A. (1989) Toward a Soar theory of taking instructions. *In* the *Proceedings of the Annual Conference of the Cognitive Science Society* 1989.

Newell, A. & Simon, H.A. (1972) *Human Problem Solving*. Englewood Cliffs: Prentice-Hall.

Velmans, M. (1989) The function of consciousness in human information processing. Paper presented at the Annual Conference of the British Psychological Society, History and Philosophy of Psychology Section, Lincoln, 1989.

Wetherick, N.E. (1989) Psychology and syllogistic reasoning. *Philosophical Psychology* **2**, 111-124.

Wetherick, N.E. & Gilhooly, K.J. (1990) Syllogistic reasoning: effects of premise order. *In* K.J. Gilhooly, M.T. Keane, R.H. Logie & G. Erdos, *Lines of Thinking*. Vol 1. Chichester: Wiley, pp 99-108.

This Page Intentionally Left Blank

Part 6

Individual Differences in Imagery

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 19

Gender differences in imagery, cognition, and memory

John T. E. Richardson
Brunel University, United Kingdom

Men and women probably have been alleged to differ from each other in all areas of psychological functioning at some time or another, and mental imagery is no exception to this. I shall begin this chapter by describing some of the major research issues that are raised when seeking to compare men and women in terms of their cognitive function. I shall then review the main theoretical positions that tend to be adopted in this field. In the body of the chapter, in order to try and achieve some sort of coherent synoptic view of the relevant literature, I shall consider four different ways of conceptualizing mental imagery: (a) as a phenomenal experience; (b) as an internal representation; (c) as a stimulus attribute; and (d) as a mnemonic strategy. I should emphasize that this is simply a didactic device that is not intended to have any specific theoretical implications.

METHODOLOGICAL ISSUES

It is most important at the outset to clarify the distinction between sex differences and gender differences. "Sex" refers to the biological distinction between men and women that may be based upon their anatomical, physiological, or chromosomal features. "Gender" marks the sociocultural distinction between men and women on the basis of the traits and behaviour that are conventionally regarded as characteristic of and appropriate to the two groups of people. Feminist theory argues that gender is a social construction which is linked by society to each sex in a wholly arbitrary way and which is learned quite independently of the underlying biological information (e.g., Humm, 1989, pp. 84, 203; Tresemer, 1975; Unger, 1979). In the present context, it is appropriate to talk of "gender differences" rather than "sex differences", because in the vast majority of studies comparing the use of mental imagery in men and women the participants are categorized on the basis of their outward appearance and behaviour rather than on the basis of their biological characteristics.

While there are borderline examples in both cases, "sex" and "gender" refer to distinctions which tend towards strict dichotomies. In contrast, most actual or putative differences between men and women in both physical characteristics and behaviour constitute overlapping distributions on some continuous dimension rather than any strict dimorphism (see Morgan, 1980; Nyborg, 1983). Moreover, a number of authors have remarked that in many cases the magnitude of the difference in question (that is, the degree of separation between the two relevant distributions) is really very small, and yet to talk of a "difference" emphasizes the contrast between the two groups at the expense of the (potentially quite considerable) within-group variation (Anderson, 1987; Caplan, MacPherson, & Tobin, 1985; Hyde, 1981; Jacklin, 1981; Plomin & Foch, 1981; but cf. Rosenthal & Rubin, 1982).

Both sex and gender are also classification variables determined in advance of a research study rather than treatment variables that are under experimental control. Since it is not possible to assign the participants at random to different experimental groups in comparisons between men and women, it follows that the findings will only be correlational in nature, and that any differences may well be due to the effects of other variables which are confounded with sex or gender (Ferguson & Takane, 1989, pp. 238, 246). The difficulty of obtaining matched samples of males and females is compounded because for convenience most studies obtain their participants from within particular institutions, and yet these institutions recruit, select, or otherwise receive individuals according to criteria that are biased towards one gender or the other. In common with most other areas of psychological research, studies into gender differences overwhelmingly are conducted using students in higher education, especially those taking introductory courses in psychology. Male and female students do not have the same opportunities to enter higher education or the same motivations to follow particular academic disciplines, and they may even be subject to different admissions criteria (Grady, 1981; Hyde, Geiringer, & Yen, 1975).

Another major problem in evaluating research on gender differences is that of publication bias. Jacklin (1981) expressed this point as follows: "If a positive instance is found, it is much easier to publish; it is more likely to be reprinted; it gets into the abstracts. In short, it becomes a part of the literature" (p. 267). In contrast, research studies which show no significant difference between men and women are much less likely to be accepted for publication (Rosenthal, 1979). Consequently, as Grady (1981) remarked, there is no research literature of gender similarities, only a research literature of gender differences. Moreover, any genuine differences in psychological functioning are likely to be exaggerated in their significance, and any spurious findings are likely to remain in the literature, even when they have proved impossible to replicate (Caplan et al., 1985; Eichler, 1988; Maccoby & Jacklin, 1974, pp. 3-5).

Nevertheless, most researchers totally ignore gender as an important social variable (see Peeck, 1973). It may be omitted from the design of the original experiment, overlooked in the data analysis, or eliminated from the final report; Maccoby and Jacklin (1974, p. 5) cited a case where findings of significant gender differences had been removed from a paper before publication under editorial instructions. Eichler (1988, chap. 4) declaimed against such "gender insensitivity" as a profound methodological problem that was inherently sexist in its probable consequences. However, other feminist writers have criticized the practice of routinely testing for gender differences without a coherent theoretical justification, thus generating a literature that is replete with "inconsequential, accidental, and incidental findings" (Grady, 1981, p. 632; see also Ussher, in press).

In any field of experimental research, the possibility has at least to be considered that the findings may result from the participants' being subtly cued by the researcher or being influenced by unintended aspects of the researcher's behaviour. Rosenthal (1976, pp. 48-51) found that both male and female experimenters smiled at female subjects more than at male subjects both before and during the administration of instructions, and they also glanced at female subjects more than at male subjects during the intervals between successive presentation trials. Moreover, both male and female experimenters exchanged fewer glances with subjects of the opposite gender while reading the formal instructions, and also tended to prolong the intervals between successive trials. Rosenthal concluded that the behaviour of researchers was related sometimes to the gender of the subjects, sometimes to the gender of the experimenter, and sometimes to both these variables. As a result of this, "male and female subjects may, psychologically, simply not be in the same experiment at all" (p. 56).

However, "passive" experimenter effects that seem to result from the mere presence of an experimenter of one gender or the other have also been found. One such effect that is particularly pertinent here is a tendency for participants to report more irrelevant images and thoughts when they are tested by an experimenter of the opposite gender (see

Algom & Singer, 1984-85; Giambra, 1988-89). Nevertheless, several experiments have found that subjects tend to achieve higher levels of objective performance under these circumstances (Archer, Cejka, & Thompson, 1961; Littig & Waddell, 1967; Stevenson & Allen, 1964). Stevenson and Allen suggested that being tested by an experimenter of the opposite gender "may result in increased competitiveness, higher anxiety, a greater desire to please, or a change in some other psychological process" (p. 216). Similar results were found in an experiment by Gralton, Hayes, and Richardson (1979), but only in the case of extraverts: introverts produced fewer correct responses and more errors when tested by experimenters of the opposite gender.

Richardson (1982) commented that research into pupillary dilation had suggested that being tested by an experimenter of the opposite gender induced an increased level of arousal in heterosexual subjects, and that effects of individual differences in introversion-extraversion were typically attributed to variations in arousal or emotional responsiveness. He then pointed out that the results of Gralton et al. (1979) were wholly consistent with the view that introverts naturally approach optimal levels of arousal (and so would be impaired by any further increase in arousal), but that extraverts are naturally at suboptimal levels of arousal (and so would benefit from an increase in arousal). Certain manipulations would compound these effects of being tested by an experimenter of the opposite gender. For instance, in eye-blink conditioning experiments, subjects are typically tested in a nearly prone position in an isolated, dark room that contains elaborate technical equipment. It is known that female subjects tend to condition more readily than males, but also that anxiety produces faster conditioning. It seems likely that the apparent gender difference is largely the result of female subjects' greater anxiety when tested by male researchers in such a setting (Jacklin, 1981; Sherman, 1978, p. 25).

THEORIES OF GENDER DIFFERENCES

A consequence of talking about "sex differences" is that researchers' attention is inevitably drawn towards biological factors in endeavouring to explain observed differences in performance between males and females. The very use of the term "sex" is taken to imply a biological mechanism, and in some cases a vague appeal to unspecified biological factors may be put forward as a sufficient explanation of the findings (see Grady, 1981; Unger, 1979). This is often in turn an occasion for wild speculation about the evolutionary pressures that might have led to such differences in performance. Of course, explaining how an aspect of human performance might be prescribed for each individual as a portion of the genetic code is a major problem that has yet to be solved in any domain of behaviour.

Another problem for biological accounts is to explain why it is that apparent differences between males and females in significant domains of expertise (such as language or mathematics) are often not manifested until adolescence or early adulthood (Unger, 1979). One obvious strategy would be to try to link the development of such functions with puberty and the consequent increase in the circulating levels of reproductive hormones. A theoretical paper by Broverman, Klaiber, Kobayashi, and Vogel (1968) cited the effects of gonadal steroids in order to explain the "well-established differences between males and females in cognitive abilities" (p. 24). It is an implication of this sort of account that any such differences would be modulated by the normal cyclical process of ovulation and menstruation (see Broverman, Vogel, Klaiber, Majcher, Shea, & Paul, 1981).

Broverman et al. (1968) claimed that available psychopharmacological research broadly supported their theory, though Parlee (1973) argued that there was little or no direct evidence to support a number of fundamental assumptions in that theory. More seriously, Parlee (1972) questioned the accuracy of Broverman et al.'s account of differences between males and females in their cognitive abilities, and she concluded that the putative findings which their model was supposed to explain were simply unreliable. A similar assessment was given by Maccoby and Jacklin (1974, pp. 99-102). The theory proposed by Broverman et al. is however the only articulated analysis of the effects of reproductive hormones upon cognitive function.

Biological accounts also have the problem of explaining how genetic, physiological, or hormonal differences between men and women come to be implicated in the mechanisms controlling human cognition. One solution is to appeal to the apparent representation of linguistic and spatial skills within the brain. Buffery and Gray (1972) argued that the lateralization of language function within the left cerebral hemisphere (in other words, cerebral "dominance") occurred at an earlier stage in the brains of girls than in the brains of boys. They inferred that this tended to constrain the development of spatial skills to the residual cortical regions within the right cerebral hemisphere, whereas these skills were able to develop in a bilateral manner among boys. Buffery and Gray then hypothesized that a bilateral representation of spatial skills would be more efficient than a unilateral one, and they concluded that this would explain the supposed superiority of males in tests of spatial ability (see below).

An alternative theory was put forward by Levy (1972), who suggested that for normal individuals spatial skills were represented in the right cerebral hemisphere, but that females and left-handed males were somewhat more likely than right-handed males to have a bilateral representation of language function. Consequently, she argued, the encroachment of verbal processes into the right cerebral hemisphere would lead to poorer spatial performance. However, both this theory and that of Buffery and Gray were subjected to considerable criticism, and neither is well-supported by the available empirical data (see, e.g., Fairweather, 1976; Marshall, 1973; Sherman, 1978, pp. 112-118). Indeed, it can be seriously doubted whether the existing evidence is sufficiently sound or coherent to motivate any specific theory concerning gender differences in the representation of cognitive faculties within the brain (cf. Bryden, 1982; McGlone, 1980).

In contrast to these notions, a consequence of talking about "gender differences" is that the attention of researchers is directed towards sociocultural factors in seeking to explain observed differences between males and females (Unger, 1979). Fairweather (1976) emphasized that the differences between men and women in cognitive performance were sensitive to such patently nonbiological variables as the birth order and culture of the subject as well as the gender of the experimenter. Other researchers have suggested that apparent differences between males and females may be artefacts arising from the use of materials or procedures in psychological experiments that elicit stereotypical behaviour amongst the participants.

A fairly decisive sort of evidence comes the study of changes in the magnitude of gender differences over the course of time. Hyde and Linn (1988) carried out a meta-analysis of previous investigations of gender differences in verbal ability. For studies published up to 1973, they found that there was a modest gender difference indicating a superiority of females over males that corresponded to 23 per cent of the population standard deviation. Nevertheless, for the studies published after 1973, the gender difference was significantly smaller and corresponded to only 10 per cent of the population standard deviation. A sociocultural account would have no difficulty in accommodating such findings, since the period in question has seen many changes in cultural attitudes and educational opportunities, but it is not at all clear how a biological account would handle a shift within a single generation (see Rosenthal & Rubin, 1982).

Finally, it is worth noting that none of the theories that have been discussed offers a distinctively *cognitive* account of putative differences between men and women in their psychological functioning. Anderson (1987) argued that cognitive theorists should consider the basic information processes that might underlie gender differences in cognition instead of concerning themselves unduly with biological or sociological explanations.

IMAGERY AS A PHENOMENAL EXPERIENCE

Mental imagery is essentially a "private" or "subjective" experience in the sense that we cannot directly observe other people's mental images. Moreover, we cannot infer whether other individuals are having any mental images (nor, indeed, *what* they are having images of) on the basis of their observable nonverbal behaviour. This means that scientific investigations of mental imagery have to depend mainly upon other people's verbal reports of their phenomenal experience. The earliest formal research of this sort was carried out by Galton (1883), who asked subjects about the quality of the mental imagery that was elicited when they tried to visualize specific objects or scenes, such as their breakfast table that morning. Galton concluded from his inquiries that "the power of visualising is higher in the female sex than in the male" (p. 99), but his published accounts were based solely upon the responses of men and boys (see also Galton, 1880). More recently, Anastasi (1958) stated that women "seem to have more vivid mental imagery than men in every sense modality" (p. 474), although she too provided no empirical data in support of this claim. The question thus arises as to whether it has been borne out in subsequent research.

The appropriate instrument for evaluating Anastasi's claim and by far the most widely used technique for evaluating the subjective vividness of experienced mental imagery is the Questionnaire Upon Mental Imagery (QMI). This was developed by Betts (1909) on the basis of some of the questions used in Galton's (1883) study, and in its original form included 150 items covering the seven major sensory modalities: visual, auditory, cutaneous, kinaesthetic, gustatory, olfactory, and organic. Sheehan (1967) found that female subjects tended to report more vivid imagery than males in most of the sensory modalities, but that the differences were both small and nonsignificant. Sheehan then developed a shortened form of the QMI containing five questions for each of the seven sensory modalities, and this is the version that is used in current research. Some studies have found that women report more vivid imagery than men on this version of the QMI (Durndell & Wetherick, 1975; Ernest, 1983; White, Ashton, & Brown, 1977). Nevertheless, it should been added that the effect of gender is relatively small and in many other studies has proved to be nonsignificant (Beech & Leslie, 1978; Di Vesta, Ingersoll, & Sunshine, 1971; Forisha, 1981; Guy and McCarter, 1978; Hiscock & Cohen, 1973; Lane, 1977; Short, 1975; Smith, 1987; Van Dyne & Stava, 1981; see also Hiscock, 1978).

Marks (1973) suggested that it was more appropriate to focus on the visual modality as the one that was most likely to be aroused in cognitive tasks. He developed a 16-item Vividness of Visual Imagery Questionnaire (VVIQ) on the basis of the visual questions in the QMI. Females tend to produce ratings on this instrument that are higher than those of males, but not significantly so (Hiscock & Cohen, 1973; Phillips, 1978). In its original form, the VVIQ is supposed to be completed twice, with the eyes closed and with the eyes open. Narchal and Broota (1988) suggested that females gave higher ratings than males when their eyes were open, whereas males gave higher ratings than females when their eyes were closed. This descriptive account which they provided does not however entirely accord with the results of their own statistical analysis.

Another questionnaire that should be mentioned in this connection is the Test of Visual Imagery Control (TVIC) which was originally devised by Gordon (1949) and revised by Richardson (1969). In its current form, it contains 12 items concerned with the extent to which visual imagery can be controlled and manipulated by the individual. In a study that was carried out by Durndell and Wetherick (1975) females produced significantly higher ratings on this instrument than males, but other investigators have found merely a nonsignificant tendency in this direction (Di Vesta et al., 1971; Forisha, 1981; Lane, 1975, 1977; Phillips, 1978) or occasionally no gender difference at all (Ernest, 1983; Hiscock, 1978; see also Hiscock & Cohen, 1973). Lane (1975, 1977) produced an extended Questionnaire on Imagery Control that was intended to measure the controllability of mental imagery in all seven of the modalities covered by the QMI. This test similarly found a small but nonsignificant tendency for females to produce higher ratings of controllability than males.

Phillips (1978) devised a short questionnaire in which the subjects were asked to image the face of a close friend and to rate the vividness of the resulting image. Women were found to report significantly more vivid imagery than men on this instrument. Subsequently, McKelvie (1984) developed a longer questionnaire on visual imagery for faces in two parts which were modelled upon the VVIQ and the TVIC, respectively. When results were combined from 10 samples totalling 435 subjects, the women were found to have rated their visual imagery for familiar faces as significantly more vivid and significantly more controllable than the men. However, the magnitude of these effects corresponded to just 8 per cent and 5 per cent of the population variance, respectively. It might also be mentioned that female subjects tend to produce higher ratings than males when normative data are collected on the imageability of individual verbal stimuli (see Bellezza, Greenwald, & Banaji, 1986).

White, Sheehan, and Ashton (1977) suggested that gender differences in subjective reports on the vividness of experienced imagery might result from differences in the degree of hemispheric specialization between males and females. However, this notion was predicated on the assumption that "imagery is a right hemisphere function" (p. 159), whereas the available evidence suggests that the phenomenal experience of imagery depends mainly upon the integrity of structures within the left cerebral hemisphere (see Richardson, 1991). By way of an alternative explanation, White et al. put forward the hypothesis that "females may have a more lax criterion for evaluating the vividness of the evoked image whereas males may have a much stricter criterion" (p. 159). To test this idea, Ashton and White (1980) re-analysed data which had been obtained by White, Ashton, and Law (1978) using a modified version of Sheehan's (1967) QMI in which the 35 questions had been reworded into a common format and reordered so that successive items elicited responses from different sensory modalities. White et al. had found that this resulted in a collapsed factor structure in which the chemical modalities (olfactory and gustatory) loaded on one factor and the mechanical modalities (auditory and cutaneous) loaded on another factor; they claimed that the grouping together of similar items in the original QMI "leads to a response bias that results in the subjects' giving similar ratings to contiguous items within the same modality" (pp. 76-77).

Ashton and White found that on this modified version of Sheehan's QMI there were significant gender differences only on the visual and organic modalities, and no difference between males and females at all in terms of their overall vividness ratings. They concluded that gender differences that were obtained using the original version of the QMI were an artefact of blocking the questions by sensory modality, in that this led to a bias amongst female subjects to use a lenient decision criterion in evaluating their experienced mental imagery (see also Sheehan, Ashton, & White, 1983). However, it might equally be claimed that the failure of White et al. (1978) to obtain a coherent factor structure and the failure of Ashton and White (1980) to obtain a significant overall gender difference on the modified version of the QMI were artefacts of randomizing the order of the items and thus making it more difficult for subjects to employ consistent decision criteria from one item to the next.

Harshman and Paivio (1987) argued that this matter could be settled by considering the responses of male and female subjects to the Individual Differences Questionnaire (IDQ). This instrument was originally compiled by Paivio (1971, p. 495) to measure the methods of thinking that were used in various situations, and it yields separate scores on imagery and verbal thinking. Shortened versions were subsequently developed by Richardson (1977) and Hiscock (1978). The IDQ requires subjects to make true-false judgements about 86 statements which are presented in a random order and in both positively and negatively worded forms in order to control for any bias towards either an acquiescent or a defensive response style. Paivio and Harshman (1983) had found no significant difference between males and females with regard to their tendency to endorse the items on the imagery scale. However, Hiscock (1978) and Ernest (1983) had found that females achieved significantly higher scores than males on this scale. Harshman and Paivio (1987) therefore concluded that gender differences in ratings on the QMI of the subjective vividness of experienced imagery constituted a genuine phenomenon and were not merely a measurement artefact.

There are also certain suggestions in the literature that the mental images produced by males and females differ in qualitative characteristics such as shape, size, and colour (Lindauer, 1973; Narchal & Broota, 1988; Palmer & Field, 1968). Harshman and Paivio (1987) found that females were more likely to report the use of mental imagery to remember, the evocation of vivid images of previously experienced scenes, and the spontaneous arousal of imagery by verbal stimuli; on the other hand, males were more likely to report the use of mental imagery in problem solving and the ability to visualize moving objects. They suggested "that females may excel more frequently at static memory imagery, whereas males may excel more frequently at dynamic transformation or manipulation of images" (p. 287). Paivio and Clark (1991) also found that men produced faster responses than women when they were asked to generate dynamic images, but that women produced faster response times than men when they were asked to generate static images. It is interesting that a functional dichotomy between a static visual system and a dynamic spatial system is envisaged in recent experimental research into visual short-term memory (Logie, 1989).

In short, there is a general tendency for females to generate higher ratings than males when asked to evaluate the vividness or controllability of their experienced imagery. Nevertheless, the difference in question is relatively slight and in many studies it turns out not to be statistically nonsignificant (cf. Davis, 1932). As one final thought, since each person can only experience his or her own mental imagery, subjects who respond to the QMI or similar instruments have no absolute or objective criteria for evaluating the vividness of their experienced imagery. It is thus not at all clear what sense it can make to say *either* that females "really" do experience more vivid imagery than males *or* that they "really" do not.

IMAGERY AS AN INTERNAL REPRESENTATION

Mental imagery is not merely a phenomenal experience, but a medium or a variety of internal representation in which information about the visual appearance of physical objects can be depicted and manipulated within some "visual buffer" or working memory (Kosslyn, 1980). Such a representation possesses emergent properties that could not be easily computed or deduced from the propositional knowledge upon which it is based (Richardson, 1980, pp. 40-41; Rollins, 1989, chap. 5). As a result, mental imagery is able to make a distinctive contribution to objective performance in certain cognitive tasks, and these are normally characterized as tests of "spatial ability". There exists a vast research literature concerned with gender differences in spatial ability (see McGee, 1979; Nyborg, 1983; Sherman, 1978), but I shall merely draw out some central themes that are especially relevant to research on mental imagery.

One problem in this field is that of defining the concept of spatial ability itself (Caplan et al., 1985; Fairweather, 1976; Maccoby & Jacklin, 1974, p. 91). A common solution is to argue that this is not a unitary function, but one that consists of two or more distinct components. For instance, Maccoby and Jacklin (1974, pp. 93-105) distinguished between: (a) visual-spatial tasks that demanded analytic processing, such as the Embedded Figures Test, the Rod and Frame Test, and Block Design; (b) visual-spatial tasks that did not demand analytic processing, such as the spatial subtests of the Differential Aptitudes Test and the Primary Mental Abilities Test as well as mazes, form boards, and block counting; and (c) nonvisual spatial tasks. Maccoby and Jacklin concluded that analytic and nonanalytic visual-spatial tasks showed consistent differences in favour of male subjects from early adolescence onwards (see also pp. 351-352). In contrast, the relatively few studies using nonvisual tasks (including some tactual analogues of analytic visual-spatial tasks) had consistently failed to produce such differences (see also McGee, 1979; Sherman, 1978,

p. 45). This led Maccoby and Jacklin to conclude that any gender differences in analytic processing were specific to visual-spatial tasks (p. 97).

"Analytic" ability in this sense was supposed to be reflected in the style or trait described as "field independence", the extent to which an individual's perceptions are independent of environmental cues. This is paradigmatically measured by the Rod and Frame Test (RFT), in which the subject is presented with a moveable luminous rod contained in a square, luminous frame in an otherwise dark room; both the rod and the frame are presented in a tilted position, and the subject's task is to instruct the experimenter to rotate the rod until it is vertical. Field independence is manifested in the accurate positioning of the rod (presumably on the basis of kinaesthetic cues) despite the discrepant information from the frame (Witkin, Lewis, Hertzman, Machover, Meissner, & Wapner, 1954, pp. 25-27). Several investigations have found that males are significantly more field independent than females from early adolescence onwards, but others have failed to show this (see Maccoby & Jacklin, 1974, pp. 95-97).

Maccoby and Jacklin (1974, pp. 93-94, 105) noted that the emergence of gender differences on the RFT and other measures of field independence exhibited a similar pattern to the developmental changes in nonanalytic visual abilities, which suggested that they could be attributed chiefly to the visual-spatial nature of the relevant tasks rather than to analytic processing per se. Subsequently, Hyde et al. (1975) showed that gender differences on the RFT disappeared when the effects of verbal and spatial ability were statistically controlled by the use of multiple regression techniques. It follows that any putative gender difference on the RFT is not due to any lack of analytic ability amongst women (Sherman, 1978, p. 48). In addition, Keogh (1971) found that boys were significantly more accurate than girls in walking out patterns on a floor, but only when additional visual cues were provided in the form of a wooden frame or a footprint trace in sand. In other words, boys proved to be rather *more* responsive to environmental cues than girls when these served to support rather than disrupt visual cognition, and there was no gender difference at all in the least field-dependent condition (see Caplan et al., 1985).

Nevertheless, Sherman (1978, pp. 25-26) also pointed out that gender differences on the RFT might well be artefactual. First, she noted that the performance of female subjects on the RFT might be disrupted by their unease at being tested by a male experimenter under conditions of nearly total darkness. Indeed, she implied that there were anecdotal reports of actual sexual harrassment in such circumstances. (The testing situation is in fact still more bizarre than was noted by Sherman. The subject has to be brought into the test room blindfolded and to remain thus for 4 min while the experimenter adjusts a headrest. Even after the removal of the blindfold, the subject is required to sit with eyes closed for periods of up to 1 min between successive trials: Witkin et al., 1954, pp. 26-27.) As Sherman acknowledged, this artefact could be eliminated by use of the portable version of the RFT. Oltman (1968), who designed this apparatus, found gender differences of similar magnitude on the standard and portable versions of the RFT, but in neither case was the difference statistically significant. However, Bogo, Winget, and Gleser (1970) found that men were significantly more field independent than women even on the portable RFT. Second, Sherman also noted that in the RFT the rod had to be adjusted by the experimenter under instructions from the subject and to that extent accurate responding demanded assertive social behaviour. As she observed, those subjects who were either not assertive in character or not used to being assertive as part of their social role might be less likely to bother the experimenter to ensure that the rod was accurately adjusted.

Other researchers have differentiated among tests of spatial ability on the basis of the information-processing demands of the tasks involved rather than the presence or absence of irrelevant cues. McGee (1979) felt that a distinction between the factors of visualization and orientation was especially well founded. The former supposedly involves "the ability to mentally manipulate, rotate, twist, or invert a pictorially presented stimulus object", whereas the latter involves "the comprehension of the arrangement of elements within a visual stimulus pattern and the aptitude to remain unconfused by the changing orientation in which a spatial configuration may be presented" (p. 893). Nevertheless, McGee noted that males tended to produce better performance than females on both of these factors from early adolescence, and that tests of spatial visualization and orientation were sometimes positively correlated with one another. It follows that this distinction may not be a very useful one in seeking to understand the nature of gender differences in spatial ability.

Sherman (1978, p. 53) suggested that the absolute magnitude of these differences was in fact very small. Hyde (1981) conducted a meta-analysis of the studies of visual-spatial ability and of field dependence which had been cited by Maccoby and Jacklin (1974). Amongst the former, the median gender difference was 0.45 of a sample standard deviation, and the median proportion of variance explained was 4.3 per cent; amongst the latter, the median gender difference was 0.51 of a sample standard deviation, and the median proportion of variance explained was 2.5 per cent. Subsequent meta-analyses performed by Linn and Petersen (1985) confirmed that gender differences in spatial ability were typically associated with relatively modest effects. Some researchers have nevertheless emphasized that even modest gender differences can have important practical consequences (e.g., Rosenthal & Rubin, 1982), particularly when individuals are being sampled from one end of a distribution, as in the case of selection for higher education, training, or employment (Burnett, 1986; Johnson & Meade, 1987).

There is however an important exception to this pattern, which is the family of tasks concerned with "mental rotation". While there are some obvious precursors within conventional psychometric test batteries, work on this topic essentially began with an experiment by Shepard and Metzler (1971). This demonstrated that the time which subjects took to match two different views of the same three-dimensional object was linearly related to the angle between the two views. Shepard and Metzler concluded that they were mentally rotating one or both stimuli at a constant rate until they had the same orientation, and all of their subjects reported using mental imagery in order to carry out this process of mental rotation. In a subsequent paper, Metzler and Shepard (1974) described further studies on this topic, and they stated informally that the slope of the function relating the response time to the angle between the two views tended to be somewhat greater (indicating a slower rate of mental rotation) in female subjects than in male subjects. Intriguingly, this evidence on potential gender differences was regarded as "relatively inconclusive" and omitted when the paper was reprinted elsewhere (Shepard & Cooper, 1982, p. 342).

Tapley and Bryden (1977) carried out a formal investigation of this issue, and found that male subjects were both faster and more accurate in comparing computer-generated representations of three-dimensional objects than females. The difference was especially pronounced in the slope of the function relating response time to angular rotation, though there was a tendency for the intercept to be lower in men than in women. Similar data were presented by Kail, Carter, and Pellegrino (1979) and by Blough and Slavin (1987), whereas Waber, Carlson, and Mann (1982) found no gender difference among schoolchildren.

Vandenberg and Kuse (1978) developed an untimed pencil-and-paper test based upon the stimulus materials which had been employed in Shepard and Metzler's (1971) original study. Using a very strict scoring criterion, measures of correct performance on this test yielded a substantial gender difference in favour of male subjects. Sanders, Soares, and D'Aquila (1982) found that this difference was 0.83 of a sample standard deviation in magnitude and explained 16 per cent of the variance, an effect which statisticians would regard as large (see also Linn & Petersen, 1985, 1986; but cf. Herman & Bruce, 1983). Indeed, in a survey which eventually encompassed more than 5,000 people from the general population of Hawaii, the average score obtained by males was roughly twice that obtained by females across the entire age range from 14 to 60 years (Wilson, DeFries, McClearn, Vandenberg, Johnson, & Rashad, 1975; Wilson & Vandenberg, 1978; see also McGee, 1978). Phillips and Rawles (1979) devised a timed pencil- and-paper test based on Shepard and Metzler's (1971) materials, and this also produced a substantial gender difference in favour of male subjects.

Explanations of gender differences in spatial ability extend across the full range of theoretical options described earlier (see McGee, 1979; Nyborg, 1983). Hereditarian accounts have been related to the supposed exploratory behaviour of the male, especially in subsistence environments (Dawson, 1972). Genetic accounts are supported by population studies and research on twins (Eliot and Fralley, 1976; McGee, 1979) and also by the study of individuals with chromosomal abnormalities or endocrinological disturbances (Dawson, 1972). A specific explanation that is associated with the work of Stafford (1961) but which is at least 20 years older suggests that spatial ability may be enhanced by a recessive gene linked to the X chromosome; this predicts a particular pattern of family intercorrelations which has not been confirmed in empirical research (Bouchard & McGee, 1977; cf. Yen, 1975). Genetic theories are also difficult to reconcile with the fact that differences between males and females in spatial ability are typically not observed until puberty (Fairweather, 1976; Johnson & Meade, 1987; Sylvester, 1913, p. 38). This suggests the involvement of hormonal factors (Caplan et al., 1985), and there have been suggestions of links between spatial ability and levels of both oestrogen and testosterone (Broverman et al., 1981; McKeever & Deyo, 1990). However, variations in women's performance across the menstrual cycle have proved hard to replicate (Chiarello, McMahon, & Schaefer, 1989; Hampson & Kimura, 1988; Ho, Gilger, & Brink, 1986; see also Sommer, 1982).

The theory of differential brain lateralization proposed by Buffery and Gray (1972) was intended to explain the supposed greater incidence of bilateral representation of spatial abilities in women than in men, but a number of commentators have insisted that the relevant empirical data are simply not reliable and in many cases show no gender differences at all (Annett, 1980; Caplan et al., 1985; Fairweather, 1976; Kinsbourne, 1980; McGlone, 1980; Sherman, 1978, chap. 6; Wolff, 1980). In mental rotation tasks, for instance, normal subjects tend to show no differences in speed or accuracy between stimuli presented in the two visual hemifields (Cohen & Polich, 1989; Corballis, 1982) or a right visual-field (left hemisphere) advantage that shows little difference between males and females (Masters, 1989; Ronsaville, 1986). Some writers have postulated that the efficient performance of complex cognitive tasks demands active cooperation between the two cerebral hemispheres (e.g., Caplan et al., 1985; McGuinness, 1980; Wolff, 1980). Indeed, recent evidence suggests that the neural mechanisms that are responsible for generating mental images are located within the left hemisphere, while the right hemisphere plays an important role in the transformation of such images (see Richardson, 1991b).

The role of sociocultural factors is shown by the fact that patterns of spatial ability appear to depend upon gender of rearing as much as upon biological sex when these are differentiated as the result of chromosomal abnormalities (Dawson, 1972; Unger, 1979). There is evidence of variation across cultures in the magnitude of these gender differences, at least on indices of field independence (Maccoby & Jacklin, 1974, pp. 129-133, 362), and they can also be modified or eliminated by means of extended practice on the relevant tasks (Caplan et al., 1985; Sherman, 1978, pp. 168-169). Within Western cultures spatial tasks tend to be perceived as masculine in nature (Newcombe, Bandura, & Taylor, 1983; Stein & Smithells, 1969), a phenomenon which Markus and Oyserman (1989) relate to the development of the self concept; and older women in particular seem to find these tasks difficult and stressful (Wilson et al., 1975). Moreover, Rosenthal and Rubin (1982) found that the magnitude of gender differences in spatial ability had become significantly smaller during the 1960s and 1970s (see also Feingold, 1988), which they concluded should make one "reluctant to make strong biogenic interpretations" of such differences (p. 711).

In addition, two studies have found a gender difference in the opposite direction when the tasks in question are perceived to be more appropriate to females. First, Birkett (1976) noted that females were supposedly more interested in family matters than males, and he gave his subjects problems about kinship relations. When the problems were presented visually, fewer than half of the subjects reported the use of mental imagery and there was no significant gender difference in the solution times. However, when the problems were presented auditorily, nearly all the subjects reported using a mental image in the shape of a schematic family tree, and the females produced significantly faster solution times than males. Second, Naditch (1976; abstracted in Sherman, 1978, pp. 229-230) administered the RFT to males and females in either a standard or a modified form. The standard version was described as a test of perceptual abilities related to spatial aspects of intelligence; as expected males achieved higher scores on field independence than females. In the modified version, the rod was replaced by a human figure, and the task was described as a test of empathy; under these circumstances, females produced higher scores than males.

In short, the notion of "spatial ability" has been taken to encompass a variety of skills. Harris (1978) suggested that "each characterization implies mental imagery, but of a distinctly kinetic rather than static kind" (p. 405). In formal tests of spatial ability, males almost always achieve higher levels of performance than females, and Paivio and Clark (1991) noted that this was consistent with the tendency of males to be more likely than females to endorse items in the IDQ relating to the use of dynamic imagery (Harshman & Paivio, 1987). In fact, Tapley and Bryden (1977) found that imagery scores on the IDQ were significantly related to the rate of mental rotation in men but not in women, so that the gender difference in mental rotation was chiefly caused by a subset of men with high visual imagery. Nevertheless, scores on the IDQ measure the habitual use of imagery and verbal thinking, and not imagery or verbal ability per se. It should be emphasized that a number of researchers have suggested that any putative gender differences in spatial tasks may reflect the disposition of males and females to use imaginal and verbal strategies (e.g., Coltheart, Hull, & Slater, 1975; McGuinness, 1980), and that variations in the organization of brain function result in differences in the use of cognitive strategies rather than differences in spatial ability (Sherman, 1980; Wolff, 1980).

IMAGERY AS A STIMULUS ATTRIBUTE

The stimulus material that subjects encounter in the course of formal laboratory experiments or psychometric testing varies in terms of the ease and the speed with which it arouses mental imagery. Paivio, Yuille, and Madigan (1968) sought to measure this quality by asking large numbers of subjects to rate verbal items in terms of their imageability along 7-point scales. By comparing the resulting ratings with an appropriate index of objective performance, Paivio (1971) argued that one could test whether mental imagery was involved in a specific cognitive task. There has been some research on the extent to which stimulus imageability influences the normal tendency for individual words to be better recognized in the right visual hemifield than in the left visual hemifield (see, e.g., Lambert & Beaumont, 1981). However, it is not clear whether such an interaction is in fact attributable to the image-arousing quality of verbal material (see Bruyer & Racquez, 1985). The specific involvement of mental imagery as a cognitive process is rather better established in the case of learning and memory. The remainder of this paper will therefore be concerned with the existence of gender differences in mental imagery and human memory.

Although Maccoby and Jacklin (1974, pp. 44-45) suggested that there were no gender differences in paired associate learning, only one of the studies that they described had employed subjects over the age of 14. In other aspects of verbal memory, females tended to achieve somewhat higher scores than males, but more than half of the studies reviewed had found no significant gender difference. Maccoby and Jacklin argued that there were no gender differences in mnemonic skills or strategies, but that there was some advantage to females by virtue of their superior language skills when tasks involved verbal materials (p. 59). A number of other studies have found that females are superior to males in the recognition and recall of pictures or names of concrete objects (Ernest, 1983; Ernest & Paivio, 1971; Harshman, Hampson, & Berenbaum, 1983; Marks, 1973). Females are also superior to males in the recognition of abstract nouns, while males appear to be superior to females in the recognition of pictorial forms (Chiswell, 1981). McGuinness and McLaughlin (1982) presented geographical scenes for subsequent recognition and recall, and found that females were superior in terms of their ability to generate verbal descriptors of the scenes and were selectively better than males in their recall of details.

To return to the role of stimulus attributes, formal experiments have reliably demonstrated that concrete material which readily evokes mental imagery is more easily remembered in a variety of situations than abstract material which evokes mental imagery only with difficulty (see Richardson, 1980, chap. 7). Sheehan (1971) tested recognition memory for concrete and abstract words under both intentional and incidental learning conditions, and found no significant difference between males and females either in terms of their absolute level of performance or in terms of the difference in performance between concrete and abstract words. Using a free-recall task, Suter (1979) found that females were superior to males under several different mnemonic instructions, but there was no significant interaction between this gender difference and the effect of stimulus concreteness. Papineau and Lohr (1981) obtained similar results using a paired-associate learning task to compare words rated high in visual or auditory imagery.

However, these investigations did not consider the possibility that memory performance might be affected by the gender of the experimenters rather than by that of the subjects (cf. Ernest, 1983). Gralton et al. (1979) carried out a formal investigation of this notion in which male and female experimenters administered a test of free recall to male and female subjects. The results are shown in Table 1. The female subjects produced significantly more correct responses than the male subjects, regardless of the gender of the experimenter; the female subjects also tended to produce fewer intrusion errors than the male subjects, which shows that the gender difference in question reflected a genuine superiority in verbal retention and not merely a more lenient response style during the recall test. The magnitude of this gender difference was identical in the case of concrete and abstract material, confirming the findings of Sheehan (1971) and Suter (1979), and indicating that there is a global difference between males and females in the retention of verbal information in long-term memory, rather than a selective difference that is specific to imaginal encoding.

Table 1. Mean percentage correct in recall of concrete and abstract words and mean number of intrusion errors by sex of subject and sex of experimenter.

	Concrete	Abstract	Overall	Intrusions
Male experimenters				
Male subjects	33	16	25	3.8
Female subjects	48	25	37	1.9
Female experimenters				
Male subjects	31	12	21	3.1
Female subjects	40	27	33	2.5
All experimenters				
Male subjects	32	14	23	3.5
Female subjects	44	26	35	2.2

Source: Gralton, Hayes, and Richardson (1979)

More recently, I conducted a further experiment to determine whether the effect of stimulus concreteness varied in magnitude across the phases of the normal menstrual cycle (Richardson, 1991a). The subjects were given parallel forms of a paired-associate learning task at each of four weekly sessions. At the end of the fourth session, the female subjects were given a biographical questionnaire which enquired into their recent menstrual history, and their responses were used to classify their recall scores with regard to four phases of the menstrual cycle. The results are shown in Table 2. There was no sign of any difference in the magnitude of the concreteness effect between the 34 male and the 68 female subjects, and no sign of any variation in the magnitude of this effect among the female subjects across the different phases of the menstrual cycle.

Table 2. Mean percentage correct recall in paired associate learning of concrete and abstract material by female subjects during the Menstruum (M), Early Intermenstruum (EIM), Late Intermenstruum (LIM), and Premenstruum (PM), and by a control group of male subjects

	Females				
	M	EIM	LIM	PM	Males
Concrete	64.73	61.30	64.24	68.95	61.84
Abstract	37.48	34.99	39.11	36.99	34.78

Source: Richardson (1991a).

In short, in so far as effects of stimulus imageability or concreteness can be taken to be diagnostic of the subjects' use of mental imagery as a mnemonic code (cf. Paivio, 1986, pp. 169-171; Richardson, 1980, chap. 7), the evidence implies that males and females do not differ in this regard.

IMAGERY AS A MNEMONIC STRATEGY

The classic paradigm for investigating the function of mental imagery in memory encoding has been to compare the retention of subjects learning under standard instructions which do not specify any particular technique for carrying out the task or instructions to engage in rote rehearsal with the performance of subjects who have been instructed to use mental imagery in their learning. Experimental research has generally found consistent, reliable, and substantial improvements in retention resulting from the administration of imagery instructions (see Richardson, 1980, chap. 6).

Rimm, Alexander, and Eiles (1969) compared the effectiveness of four different instructional sets on the paired-associate learning of concrete nouns. Interactive imagery instructions were found to yield significantly better performance than standard instructions, rote rehearsal, and verbal mediation in terms of the performance on the first trial, the total number of errors to criterion, and the number of trials to criterion. Men tended to achieve slightly (though not significantly) worse performance than women on all three variables and under all instructional sets with the exception of verbal mediation, and there was no sign of any interaction between the effects of gender and instructional set on any of the three measures. A very similar pattern of results was obtained by Short (1975), by Treat and Reese (1976), and by Suter (1979). In short, there seems to be no gender difference in the advantage gained from imagery mnemonic instructions.

However, two studies have found that males benefit more than females from imagery instructions in tasks designed to tap the different coding processes used by the two cerebral hemispheres. Metzger and Antes (1976) presented pairs of concrete nouns under instructions to engage in imagery or rote rehearsal. When subsequently presented with a word probe within either the left or right visual hemifield, male subjects responded faster than female subjects in the imagery condition, whereas female subjects responded faster than male subjects in the rehearsal condition. No such interaction was obtained in a second experiment using picture probes. Lenhart and Schwartz (1983) used an analogous experimental procedure to present abstract tactual stimuli to the left and right hands. Subjects who received imagery instructions were significantly more accurate in making same/different judgements about probe stimuli than subjects who received verbal instructions, but only when male subjects were asked to palpate the probe stimulus with the left hand. This was interpreted to mean that imaginal encoding was more lateralized to the right cerebral hemisphere in men than in women.

It will be recalled that Harshman and Paivio (1987) found that male subjects reported more vivid images of moving objects than females, but that female subjects reported more vivid images of static scenes. These subjective reports do not accord with the results of a study by Joseph, McKay, and Joseph (1982), who found that female subjects benefited more from mnemonic instructions to make up moving images, while male subjects benefited more from instructions to make up static images; however, this pattern emerged only in a test of free recall, and no gender differences were found in a test of recognition memory. In their own study, Harshman and Paivio (1987) remarked upon "the widespread tendency for females more frequently to report the occurrence and use of memory imagery" (p. 296). However, rather than ask subjects to endorse or reject general statements of the sort employed in the IDQ, Sheehan et al. (1983) suggested that it might be more valid to obtain concurrent reports of subjective experiences during the criterion task itself (cf. Davis, 1932).

In the case of human learning and memory, the appropriate methodology is the use of postlearning questionnaires in which subjects report on the use of different sorts of learning strategies or mediators. Richardson (1978) showed that the proportion of items for which subjects reported the use of imaginal mediators was an excellent predictor of their performance in learning concrete material. More recently, Richardson (1985) showed that the administration of imagery instructions increased the reported use of imaginal mediators, but had no effect upon the conditional probability that a particular item would be recalled given that imagery had been used. This was taken to mean that imagery instructions enhanced the availability of imaginal mediators but not their efficacy. Given that mediator reports seem to be a sensitive way of monitoring the use of imagery as a mnemonic strategy, it is pertinent to ask whether they exhibit gender differences.

I therefore carried out an experiment similar to those of Richardson (1985) involving four cohorts of students (consisting of 50 males and 123 females). They were presented with 120 high-frequency nouns assigned at random to two lists of 30 paired associates; in each case, the first and last five pairs were regarded as filler items to remove the primacy and recency effects in recall, and only the remaining 20 pairs were tested at the time of recall. The first of the two lists was learned under standard instructions, whereas the second was learned under imagery instructions. Immediately after the recall test on each list, the subjects were given a postlearning questionnaire in which they were asked to indicate the type of mediator they had employed ("imagery", "verbal", "repetition", "other", or "none") for each of the critical paired associates.

The results are shown in Table 3. The proportion of pairs correctly recalled showed a highly significant increase from List 1 to List 2. The female subjects produced slightly better retention than the male subjects, but neither the effect of gender nor its interaction with the effect of instructions even approached statistical significance. These results are entirely consistent with those of previous research. The proportion of pairs for which an imaginal mediator was reported also showed a highly significant increase from List 1 to List 2. In this case, there was also a significant difference between the male and the female subjects, but no interaction between these two effects. Contrary to the data of Harshman and Paivio (1987), female subjects were actually less likely than male subjects to report the use of imaginal mediators in paired-associate learning. Finally, the conditional probability of correct recall based upon the items for which imaginal mediators had been reported showed no significant effect of either gender or instructions, though the female subjects tended to show more efficient performance than the males.

Table 3. Mean percentage correct recall in paired associate learning, mean percentage use of imaginal mediators, and mean percentage correct for imaged pairs for male and female subjects before (List 1) and after (List 2) administration of imagery mnemonic instructions

	P(R)		P(I)		P(R | I)	
	List 1	List 2	List 1	List 2	List 1	List 2
Males	60.0	73.5	67.0	91.0	73.5	73.5
Females	60.3	78.7	55.8	88.4	77.4	79.0
Overall	60.2	77.3	59.2	89.2	76.3	77.5

CONCLUDING SUMMARY

1. Females produce higher ratings than males on imagery vividness and controllability, though the effect is small and often nonsignificant.
2. Males beyond early adolescence achieve higher scores than females on tests of spatial ability; the effect is usually small and sometimes nonsignificant, but tests of mental rotation often yield substantial effects. In at least some cases the direction of the effect can be reversed by using tasks that are more appropriate to females.
3. There is no reliable difference between males and females in the effects of stimulus imageability in tests of learning and memory.
4. There is no reliable difference between males and females in the effects of imagery instructions in tests of learning and memory.
5. Females are less likely than males to use mental imagery as a mnemonic device, but they tend to be more likely to recall those items where mental imagery has been so used. As a result, any gender difference in recall tends to be in the favour of females rather than of males.

REFERENCES

Algom, D., & Singer, J. L. (1984-85). Interpersonal influences on task-irrelevant thought and imagery in a signal detection experiment. *Imagination, Cognition, and Personality* **4**, 69-83.

Anastasi, A. (1958). *Differential Psychology: Individual and Group Differences in Behavior, 3rd ed.* New York: Macmillan.

Anderson, N.S. (1987). Cognition, learning, and memory. *In* M.A. Baker (Ed.), *Sex Differences in Human Performance.* Chichester: Wiley. Pp. 37- 54.

Annett, M. (1980). Sex differences in laterality: Meaningfulness versus reliability. *Behavioral and Brain Sciences* **3**, 227-228.

Archer, E.J., Cejka, J.A., & Thompson, C.P. (1961). Serial-trigram learning as a function of differential meaningfulness and sex of subjects and experimenters. *Canadian Journal of Psychology* **15**, 148-153.

Ashton, R., & White, K.D. (1980). Sex differences in imagery vividness: An artifact of the test. *British Journal of Psychology* **71**, 35-38.

Beech, J.R., & Leslie, J.C. (1978). Monitoring the effects of increasing resource demands on the clarity of visual images. *British Journal of Psychology* **69**, 323-333.

Bellezza, F.S., Greenwald, A.G., & Banaji, M.R. (1986). Words high and low in pleasantness as rated by male and female college students. *Behavior Research Methods, Instruments, and Computers* **18**, 299-303.

Betts, G.H. (1909). *The Distribution and Functions of Mental Imagery.* New York: Teachers College.

Birkett, P. (1976). Sex differences and reasoning vs. imagery strategies in the solution of visually and auditorily presented family relationship problems. *Bulletin of the Psychonomic Society* **8**, 139-142.

Blough, P.M., & Slavin, L.K. (1987). Reaction time assessments of gender differences in visual-spatial performance. *Perception and Psychophysics* **41**, 276-281.

Bogo, N., Winget, C., & Gleser, G.C. (1970). Ego defenses and perceptual styles. *Perceptual and Motor Skills* **30**, 599-605.

Bouchard, T.J., Jr., & McGee, M.G. (1977). Sex differences in human spatial ability: Not an X-linked recessive gene effect. *Social Biology* **24**, 332-335.

Broverman, D.M., Klaiber, E.L., Kobayashi, Y., & Vogel, W. (1968). Roles of activation and inhibition in sex differences in cognitive abilities. *Psychological Review* **75**, 23-50.

Broverman, D.M., Vogel, W., Klaiber, E.L., Majcher, D., Shea, D., & Paul, V. (1981). Changes in cognitive task performance across the menstrual cycle. *Journal of Comparative and Physiological Psychology* **95**, 646-654.

Bryden, M.P. (1982). *Laterality: Function Asymmetry in the Intact Brain.* New York: Academic Press.

Buffery, A.W.H., & Gray, J.A. (1972). Sex differences in the development of spatial and linguistic skills. *In* C. Ounsted & D.C. Taylor (Eds.), *Gender Differences: Their Ontogency and Significance.* Edinburgh: Churchill Livingstone. Pp. 123-157.

Burnett, S.A. (1986). Sex-related differences in spatial ability: Are they trivial? *American Psychologist* **41**, 1012-1014.

Caplan, P.J., MacPherson, G.M., & Tobin, P. (1985). Do sex-related differences in spatial abilities exist? A multilevel critique with new data. *American Psychologist* **40**, 786-799.

Chiarello, C., McMahon, M.A., & Schaefer, K. (1989). Visual cerebral lateralization over phases of the menstrual cycle: A preliminary investigation. *Brain and Cognition* **11**, 18-36.

Chiswell, J.N. (1981). The effects of imagery and sex on hemispheric laterality and recognition using abstract picture and word stimuli. *Dissertation Abstracts International* **41A**, 4943.

Cohen, W., & Polich, J. (1989). No hemispheric differences for mental rotation of letters or polygons. *Bulletin of the Psychonomic Society* **27**, 25-28.

Coltheart, M., Hull, E., & Slater, D. (1975). Sex differences in imagery and reading. *Nature* **253**, 438-440.

Corballis, M.C. (1982). Mental rotation: Anatomy of a paradigm. *In* M. Potegal (Ed.), *Spatial Abilities: Developmental and Physiological Foundations.* New York: Academic Press. Pp. 173-198.

Davis, F.C. (1932). The functional significance of imagery differences. *Journal of Experimental Psychology* **15**, 630-661.

Dawson, J.L.M. (1972). Effects of sex hormones on cognitive style in rats and men. *Behavior Genetics* **2**, 21-42.

Di Vesta, F.J., Ingersoll, G., & Sunshine, P. (1971). A factor analysis of imagery tests. *Journal of Verbal Learning and Verbal Behavior* **10**, 471-479.

Durndell, A.J., & Wetherick, N.E. (1975). Reported imagery and social desirability. *Perceptual and Motor Skills* **41**, 987-992.

Eichler, M. (1988). *Nonsexist Research Methods.* Boston, MA: Allen & Unwin.

Eliot, J., & Fralley, J.S. (1976). Sex differences in spatial abilities. *Young Children* **31**, 487-498.

Ernest, C.H. (1983). Imagery and verbal ability and recognition memory for pictures and words in males and females. *Educational Psychology* **3**, 227-244.

Ernest, C.H., & Paivio, A. (1971). Imagery and sex differences in incidental recall. *British Journal of Psychology* **62**, 67-72.

Fairweather, H. (1976). Sex differences in cognition. *Cognition* **4**, 231-280.

Feingold, A. (1988). Cognitive gender differences are disappearing. *American Psychologist* **43**, 95-103.

Ferguson, G.A., & Takane, Y. (1989). *Statistical Analysis in Psychology and Education, 6th ed.* New York: McGraw-Hill.

Forisha, B.L. (1981). Patterns of creativity and mental imagery in men and women. *Journal of Mental Imagery* **5**(1), 85-96.

Galton, F. (1880). Statistics of mental imagery. *Mind* **5**, 301-318

Galton, F. (1883). *Inquiries into Human Faculty and Its Development.* London: Macmillan.

Giambra, L.M. (1988-89). The influence of subject-experimenter sexual congruence on the frequency of task-unrelated imagery and thought: Further evidence. *Imagination, Cognition, and Personality* **8**, 249-260.

Gordon, R. (1949). An investigation into some of the factors that favour the formation of stereotyped images. *British Journal of Psychology* **40**, 156-167.

Grady, K.E. (1981). Sex bias in research design. *Psychology of Women Quarterly* **5**, 628-636.

Gralton, M.A., Hayes, Y.A., & Richardson, J.T.E. (1979). Introversion-extraversion and mental imagery. *Journal of Mental Imagery* **3**, 1-10.

Guy, M.E., & McCarter, R.E. (1978). A scale to measure emotive imagery. *Perceptual and Motor Skills* **46**, 1267-1274.

Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behavioral Neuroscience* **102**, 456-459.

Harris, L.J. (1978). Sex differences in spatial ability: Possible environmental, genetic, and neurological factors. *In* M. Kinsbourne (Ed.), *Asymmetrical Function of the Brain.* Cambridge: Cambridge University Press. Pp. 405-522.

Harshman, R.A., Hampson, E., & Berenbaum, S.A. (1983). Individual differences in cognitive abilities and brain organization, Part I: Sex and handedness differences in ability. *Canadian Journal of Psychology* **37**, 144-192.

Harshman, R.A., & Paivio, A. (1987). "Paradoxical" sex differences in self-reported imagery. *Canadian Journal of Psychology* **41**, 287-302.
Herman, J.F., & Bruce, P.R. (1983). Adults' mental rotation of spatial information: Effects of age, sex and cerebral laterality. *Experimental Aging Research* **9**, 83-85.
Hiscock, M. (1978). Imagery assessment through self-report: What do imagery questionnaires measure? *Journal of Consulting and Clinical Psychology* **46**, 223-230.
Hiscock, M., & Cohen, D.B. (1973). Visual imagery and dream recall. *Journal of Research in Personality* **7**, 179-188.
Ho, H.-Z., Gilger, J.W., & Brink, T.M. (1986). Effects of menstrual cycle on spatial information-processes. *Perceptual and Motor Skills* **63**, 743-751.
Humm, M. (1989). *The Dictionary of Feminist Theory.* New York: Harvester Wheatsheaf.
Hyde, J.S. (1981). How large are cognitive gender differences? A meta-analysis using w^2 and *d. American Psychologist* **36**, 892-901.
Hyde, J.S., Geiringer, E.R., & Yen, W.M. (1975). On the empirical relation between spatial ability and sex differences in other aspects of cognitive performance. *Multivariate Behavioral Research* **10**, 289-309.
Hyde, J.S., & Linn, M.C. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin* **104**, 53-69.
Jacklin, C.N. (1981). Methodological issues in the study of sex-related differences. *Developmental Review* **1**, 266-273.
Johnson, E.S., & Meade, A.C. (1987). Developmental patterns of spatial ability. *Child Development* **58**, 725-740.
Joseph, C.A., McKay, T.D., & Joseph, C.R. (1982). Sex effects on free recall and recognition of transforming orally presented nouns into imagery form. *Journal of General Psychology* **106**, 213-217.
Kail, R., Carter, P., & Pellegrino, J. (1979). The locus of sex differences in spatial ability. *Perception and Psychophysics* **26**, 182- 186.
Keogh, B.K. (1971). Pattern copying under three conditions of an expanded spatial field. *Developmental Psychology* **4**, 25-31.
Kinsbourne, M. (1980). If sex differences in brain lateralization exist, they have yet to be discovered. *Behavioral and Brain Sciences* **3**, 241-242.
Kosslyn, S.M. (1980). *Image and Mind.* Cambridge, MA: Harvard University Press.
Lane, J.B. (1975). Imagination and personality: The multi-trait investigation of a new measure of imagery control. *Dissertation Abstracts International* **35B**, 6099.

Lane, J.B. (1977). Problems in assessment of vividness and control of imagery. *Perceptual and Motor Skills* **45**, 363-368.

Lenhart, R.E., & Schwartz, S.M. (1983). Tactile perception and the right hemisphere: A masculine superiority for imagery coding. *Brain and Cognition* **2**, 224-232.

Levy, J. (1972). Lateral specialization of the human brain: Behavioral manifestations and possible evolutionary basis. *In* J.A. Kiger, Jr. (Ed.), *The Biology of Behavior.* Corvallis, OR: Oregon State University Press. Pp. 159-180.

Lindauer, M. (1973). Form imagery to colors. *Perceptual and Motor Skills* **36**, 165-166.

Linn, M.C., & Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development* **56**, 1479-1498.

Linn, M.C., & Petersen, A.C. (1986). A meta-analysis of gender differences in spatial ability: Implications for mathematics and science achievement. *In* J.S. Hyde & M.C. Linn (Eds.), *The Psychology of Gender: Advances through Meta-analysis.* Baltimore, MD: Johns Hopkins University Press. Pp. 67-101.

Littig, L.W., & Waddell, C.M. (1967). Sex and experimenter interaction in serial learning. *Journal of Verbal Learning and Verbal Behavior* **6**, 676-678.

Logie, R.H. (1989). Characteristics of visual short-term memory. *European Journal of Cognitive Psychology* **1**, 275-284.

Maccoby, E.E., & Jacklin, C.N. (1974), *The Psychology of Sex Differences.* Stanford, CA: Stanford University Press.

Marks, D.F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology* **64**, 17-24.

Markus, H., & Oyserman, D. (1989). Gender and thought: The role of the self-concept. *In* M. Crawford & M. Gentry (Eds.), *Gender and Thought: Psychological Perspectives.* New York: Springer-Verlag. Pp. 100-127.

Marshall, J.C. (1973). Some problems and paradoxes associated with recent accounts of hemispheric specialization. *Neuropsychologia* **11**, 463-470.

Masters, M.P. (1989). Visual half field studies of mental rotation. *Dissertation Abstracts International* **50B**, 2179.

McGee, M.G. (1978). Handedness and mental rotation. *Perceptual and Motor Skills* 47, 641-642.

McGee, M.G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin* **86**, 889-918.

McGlone, J. (1980). Sex differences in human brain asymmetry: A critical survey. *Behavioral and Brain Sciences* **3**, 215-263.

McGuinness, D. (1980). Strategies, demands, and lateralized sex differences. *Behavioral and Brain Sciences* **3**, 244.

McGuinness, D., & McLaughlin, L. (1982). An investigation of sex differences in visual recognition and recall. *Journal of Mental Imagery* **6**, 203-212.

McKeever, W.F., & Deyo, R.A. (1990). Testosterone, dihydrotestosterone, and spatial task performances of males. *Bulletin of the Psychonomic Society* **28**, 305-308.

McKelvie, S.J. (1984). Reported visual imagery for faces and facial recognition memory. *Perceptual and Motor Skills* **59**, 825-826.

Metzger, R.L., & Antes, J.R. (1976). Sex and coding strategy effects on reaction time to hemispheric probes. *Memory and Cognition* **4**, 167-171.

Metzler, J., & Shepard, R.N. (1974). Transformational studies of the internal representation of three-dimensional objects. *In* R.L. Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium.* Potomac, MD: Erlbaum. Pp. 147-201.

Morgan, M.J. (1980). Influences of sex on variation in human brain asymmetry. *Behavioral and Brain Sciences* **3**, 244-245.

Naditch, S.F. (1976). Sex differences in field dependence: The role of social influence. *Paper presented at the convention of the American Psychological Association*, Washington, D.C.

Narchal, R., & Broota, K.D. (1988). Sex differences in vividness of visual imagery under eyes open and eyes closed conditions. *Journal of Mental Imagery* **12**, 81-87.

Newcombe, N., Bandura, M.M., & Taylor, D.G. (1983). Sex differences in spatial ability and spatial activities. *Sex Roles* **9**, 377-386.

Nyborg, H. (1983). Spatial ability in men and women: Review and new theory. *Advances in Behaviour Research and Therapy* **5**, 89-140.

Oltman, P.K. (1968). A portable rod-and-frame apparatus. *Perceptual and Motor Skills* **26**, 503-506.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart, & Winston.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach.* New York: Oxford University Press.

Paivio, A. & Clark, J.M. (1991). Static versus dynamic imagery. *In* C. Cornoldi & M.A. McDaniel (Eds.), *Imagery and Cognition.* New York: Springer-Verlag. pp 221-245.

Paivio, A., & Harshman, R. (1983). Factor analysis of a questionnaire on imagery and verbal habits and skills. *Canadian Journal of Psychology* **37**, 461-483.

Paivio, A., Yuille, J.C., & Madigan, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, Monograph Supplement* **76**, 1 (Part 2).

Palmer, R.D., & Field, P.B. (1968). Visual imagery and susceptibility to hypnosis. *Journal of Consulting and Clinical Psychology* **32**, 456-461.

Papineau, W., & Lohr, J.M. (1981). Effect of sensory modality for presentation and word imagery on paired-associate learning. *Perceptual and Motor Skills* **52**, 747-750.

Parlee, M.B. (1972). Comments on "Roles of activation and inhibition in sex differences in cognitive abilities" by D.M. Broverman, E.L. Klaiber, Y. Kobayashi, and W. Vogel. *Psychological Review* **79**, 180-184.

Parlee, M.B. (1973). The premenstrual syndrome. *Psychological Bulletin* **80**, 454-465.

Peeck, J. (1973). Sexeverschillen in leerprocessen met plaatjes. *Nederlands Tijdschrift voor de Psychologie* **28**, 343-349.

Phillips, R.J. (1978). Recognition, recall and imagery of faces. *In* M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), *Practical Aspects of Memory*. London: Academic Press. Pp. 270-277.

Phillips, R.J., & Rawles, R.E. (1979). Recognition of upright and inverted faces: A correlational study. *Perception* **8**, 577-583.

Plomin, R., & Foch, T.T. (1981). Sex differences and individual differences. *Child Development* **52**, 383-385.

Richardson, A. (1969). *Mental Imagery*. New York: Springer.

Richardson, A. (1977). Verbalizer-visualizer: A cognitive style dimension. *Journal of Mental Imagery* **1**, 109-125.

Richardson, J.T.E. (1978). Reported mediators and individual differences in mental imagery. *Memory and Cognition* **6**, 376-378.

Richardson, J.T.E. (1980). *Mental Imagery and Human Memory*. London: Macmillan.

Richardson, J.T.E. (1982). Introversion-extraversion and experimenter effects in memory tasks. *Personality and Individual Differences* **3**, 327-328.

Richardson, J.T.E. (1985). Converging operations and reported mediators in the investigation of mental imagery. *British Journal of Psychology* **76**, 205-214.

Richardson, J.T.E. (1991a). Cognition, memory, and the menstrual cycle. *Cahiers de Psychologie Cognitive* **11**, 3-26.

Richardson, J.T.E. (1991b). Imagery and the brain. *In* C. Cornoldi & M.A. McDaniel (Eds.), *Imagery and Cognition*. New York: Springer-Verlag. pp 1-45.

Rimm, D.C., Alexander, R.A., & Eiles, R.R. (1969). Effects of different mediational instructions and sex of subject on paired-associate learning of concrete nouns. *Psychological Reports* **25**, 935-940.

Rollins, M. (1989). *Mental Imagery: On the Limits of Cognitive Science*. New Haven, CT: Yale University Press.

Ronsaville, D.L.S. (1986). Sex differences in processing spatial rotations as related to brain organization and strategy. *Dissertation Abstracts International* **47B**, 1764.

Rosenthal, R. (1976). *Experimenter Effects in Behavioral Research*, enlarged ed. New York: Irvington.

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin* **86**, 638-640.

Rosenthal, R., & Rubin, D.B. (1982). Further meta-analytic procedures for assessing cognitive gender differences. *Journal of Educational Psychology* **74**, 708-712.

Sanders, B., Soares, M.P., & D'Aquila, J.M. (1982). The sex difference on one test of spatial visualization: A nontrivial difference. *Child Development* **53**, 1106-1110.

Sheehan, P.W. (1967). A shortened form of Betts' Questionnaire Upon Mental Imagery. *Journal of Clinical Psychology* **23**, 386-389.

Sheehan, P.W. (1971). Individual differences in vividness of imagery and the function of imagery in incidental learning. *Australian Journal of Psychology* **23**, 279-288.

Sheehan, P.W., Ashton, R., & White, K. (1983). Assessment of mental imagery. *In* A.A. Sheikh (Ed.), *Imagery: Current Theory, Research, and Application*. New York: Wiley. Pp. 189-221.

Shepard, R.N., & Cooper, L.A. (1982). *Mental Images and Their Transformations*. Cambridge, MA: MIT Press.

Shepard, R.N., & Metzler, J. (1971). Mental rotation of three- dimensional objects. *Science* **171**, 701-703.

Sherman, J.A. (1978). *Sex-Related Cognitive Differences: An Essay on Theory and Evidence*. Springfield, IL: Thomas.

Sherman, J. (1980). Sex-related differences in functional human brain asymmetry: Verbal function--no; spatial function--maybe. *Behavioral and Brain Sciences* **3**, 248-249.

Short, R.H. (1975). Imaginal processes in learning and memory. *Dissertation Abstracts International* **36A**, 1403-1404.

Smith, M.C. (1987). An investigation of the effects of different sound frequencies on vividness and creativity of imagery. *Dissertation Abstracts International* **47B**, 3708.

Sommer, B. (1982). Cognitive behavior and the menstrual cycle. *In* R.C. Friedman (Ed.), *Behavior and the Menstrual Cycle.* New York: Marcel Dekker. Pp. 101-127.

Stafford, R.E. (1961). Sex differences in spatial visualization as evidence of sex-linked inheritance. *Perceptual and Motor Skills* **13**, 428.

Stein, A.H., & Smithells, J. (1969). Age and sex differences in children's sex-role standards about achievement. *Developmental Psychology* **1**, 252-259.

Stevenson, H.W., & Allen, S. (1964). Adult performance as a function of sex of experimenter and sex of subject. *Journal of Abnormal and Social Psychology* **68**, 214-216.

Suter, W.N. (1979). Free recall as a function of instructional set, imagery level, and sex. *Perceptual and Motor Skills* **48**, 535-538.

Sylvester, R.H. (1913). The form board test. *Psychological Monographs* **15** (4, Whole No. 65).

Tapley, S.M., & Bryden, M.P. (1977). An investigation of sex differences in spatial ability: Mental rotation of three-dimensional objects. *Canadian Journal of Psychology* **31**, 122-130.

Treat, N.J., & Reese, H.W. (1976). Age, pacing, and imagery in paired-associate learning. *Developmental Psychology* **12**, 119-124.

Tresemer, D. (1975). Assumptions made about gender roles. *In* M. Millman & R.M. Kanter (Eds.), *Another Voice: Feminist Perspectives on Social Life and Social Science.* Garden City, NY: Anchor Books. Pp. 308-339.

Unger, R.K. (1979). Toward a redefinition of sex and gender. *American Psychologist* **34**, 1085-1094.

Ussher, J.M. (in press). A diathesis-stressor approach to menstruation. *In* J. T. E. Richardson (Ed.), *Cognition and the Menstrual Cycle.* New York: Springer-Verlag.

Vandenberg, S.G., & Kuse, A.R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills* **47**, 599-604.

Van Dyne, W.T., & Stava, L.J. (1981). Analysis of relationships among hypnotic susceptibility, personality type, and vividness of mental imagery. *Psychological Reports* **48**, 23-26.

Waber, D.P., Carlson, D., & Mann, M. (1982). Developmental and differential aspects of mental rotation in early adolescence. *Child Development* **53**, 1614-1621.

White, K.D., Ashton, R., & Brown, R.M.D. (1977). The measurement of imagery vividness: Normative data and their relationship to sex, age, and modality differences. *British Journal of Psychology* **68**, 203-211.

White, K.D., Ashton, R., & Law, H. (1978). The measurement of imagery vividness: Effects of format and order on the Betts' Questionnaire Upon Mental Imagery. *Canadian Journal of Behavioral Science* **10**, 68-78.

White, K.D., Sheehan, P.W., & Ashton, R. (1977). Imagery assessment: A survey of self-report measures. *Journal of Mental Imagery* **1**, 145-170.

Witkin, H.A., Lewis, H.B., Hertzman, M., Machover, K., Meissner, P.B., & Wapner, S. (1954). *Personality Through Perception: An Experimental and Clinical Study.* New York: Harper & Brothers.

Wilson, J.R., DeFries, J.C., McClearn, G.E., Vandenberg, S.G., Johnson, R.C., & Rashad, M.N. (1975). Cognitive abilities: Use of family data as a control to assess sex and age differences in two ethnic groups. *International Journal of Aging and Human Development* **6**, 261-276.

Wilson, J.R., & Vandenberg, S.G. (1978). Sex differences in cognition: Evidence from the Hawaii Family Study. *In* T.E. McGill, D.A. Dewsbury, & B.D. Sachs (Eds.), *Sex and Behavior: Status and Prospectus.* New York: Plenum Press. Pp. 317-335.

Wolff, P.H. (1980). A difference that may make no difference. *Behavioral and Brain Sciences* **3**, 250-251.

Yen, W.M. (1975). Sex-linked major-gene influences on selected types of spatial performance. *Behavior Genetics* **5**, 281-298.

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 20

The study of vividness of images

Cesare Cornoldi[1], Rossana De Beni[1], Fiorella Giusberti[2], Francesco Marucci[3], M. Massironi[4] and Giuliana Mazzoni[1]
Universities of [1]Padova, [2]Bologna, [3]Roma, [4]Verona, Italy

This paper summarises some studies we have carried out on vividness of images. In a first series of experiments we examined which variables may affect vividness ratings. In a second series we used the vividness rating procedure to contrast different conditions and groups of subjects.

By focusing on vividness qualities of images we intend to suggest that the study of the structural properties of mental images is no less important than is the study of their functional properties. Furthermore, the phenomenological consideration of images appears to have the advantage that it does not require any legitimating proof concerning the existence of mental imagery, as the experience of having an image is central to the subjective mental life of the majority of people. Vividness, in particular, is one of the main qualities of such experience and has been the subject of research, especially in connection with the more or less explicit idea that those who have more vivid images are generally better at imaging and thus perform better in tasks requiring the use of imagery (Betts, 1909; Denis, 1982; Marks, 1972, etc).

Image vividness has been defined only intuitively as if the rating instructions given to subjects were based on a primitive dimension that is immediately comprehensible though not wholly definable. These intuitive definitions sometimes refer to two aspects defining image "vividness": a) the extent to which the image approaches actual visual experience; b) image luminosity-clarity. These two characteristics (a and b) are not necessarily present together. Furthermore, they do not seem to involve all the possible variables affecting ratings of vividness. Indeed, on the basis of a series of interviews we observed that other aspects were often mentioned as being related to vividness.

Our first study (Cornoldi et al., in press) was aimed at identifying factors influencing vividness judgements. We hypothesized that it is possible to identify some characteristics of an image which may play a critical role in determining the assessment of vividness.

In a small group of experts we explored the definitions of a vivid image given in an informal interview. Our re-analysis and discussion of their answers brought us to focus on the six following characteristics of an image: presence of colours, presence of a rich context, emergence of salient features, richness of details, well-defined shape and contour, generality of the represented object. For example an image of an house could be either coloured, or black and white (colour), within a landscape (context), with an emergent architectural element, such as an arch, (saliency), quite detailed in its elements, such as windows, doors, chimney etc. (details), well defined in its shape and contour, (shape and contour), and not referred to a particular example of a specific house (generality).

In the first experiment, we asked 18 University students to form images specifically related to one of the six above mentioned characteristics. For example, when based on colour, subjects had to create a good coloured image of a named object, and to concentrate on colours as a fundamental element. Thirty object names (five for each characteristic) were orally presented at a rate of one every 20 sec, during which time subjects heard the word, wrote it down, formed an image of the required kind (the six characteristics were randomly balanced within the list) and concentrated on it until they were asked to rate it in vividness by putting a mark at some point along a 20 cm line. After forming all 30 images and rating their vividness, subjects were asked to perform a 90 sec interpolated task (counting backwards) and then to write down all the words they remembered from the 30 item list.

The various types of image (i.e. the stimuli imagined with one of the six characteristics) produced significant differences in recall (coloured images were recalled better than detailed and generic ones) but not in vividness ratings.

Nevertheless, this failure to find differential effects on vividness ratings could have been due to the procedure adopted and, in particular, to the fact that the subjects' concentration had continuously to shift from one characteristic to another. It was surprising to find that the characteristic of generality (where the instructions requested subjects to form an image which did not represent any particular object, but a generic representation of the object) affected vividness ratings in the same way as the other characteristics. This is especially surprising when compared with the results of a previous study where generality appeared to negatively influence vividness (Cornoldi, De Beni & Pra Baldi, 1989). It is possible that carry-over from one instruction to the next occurred, equalizing the effects of the different instructions.

For this reason we divided our second into six sessions, in each of which 45 subjects concentrated on only one of the six characteristics for the same thirty object names. Inviting subjects to form different images for the same objects minimized the influence of the stimuli and gave us the possibility of exploring which characteristic was more emergent when an image was retrieved from memory.

To examine the effects on the vividness ratings of groups of different imagery ability, tested on the basis of introspective reports, we administered the IDQ imagery test (Paivio & Harshman, 1983) to 88 University students. Of these, 23 with scores in the I scale (propensity to use images) higher than 31 were classified as high imagers, and 22 with scores lower than 23 as low imagers. High imagers gave significantly higher vividness ratings for all the six conditions. Correlations between IDQ-I scores and mean ratings in the six conditions were nevertheless significant only in the cases of detail, shape, and contour and generality (the correlations were respectively .298, .291, and .281). Furthermore, the object names that were recalled best proved to be those imaged with salient characteristics, while those imaged with the characteristic of details were recalled least well.

An Anova revealed a significant effect of characteristic: Post-hoc comparison revealed that saliency (mean recall = 3.84) was significantly recalled better (i.e. more frequently) than shape and contour (3.13), generality (3.07), and detail (2.60); context (3.57) and colour (3.33), too, obtained significantly higher recall than detail.

The main aim of experiment 3 was to assess which characteristics have greater relevance in producing vividness ratings when images are spontaneously generated and immediately rated for vividness. For this purpose subjects were asked only to create an image of the presented noun. After giving the vividness ratings, they were asked to rate to what extent the six characteristics (examined in the preceding experiments) were present in the image by putting a mark along a line.

Stepwise multiple regression analysis revealed that the dependent variable represented by the vividness ratings was significantly influenced first by shape-and-contour ($B = .42$ $p < .001$), then by detail ($B = .20$, $p < .001$), context ($B = .08$, $p = .022$), and generality ($B = .08$, $p = .024$).

We should add that items recalled at a following memory test were rated as significantly more vivid than non-recalled items.

In our opinion, it was in this experiment that the procedure adopted most closely mirrored the situation in which a person creates an image and immediately experiences its degree of vividness. However, the successive rating with regard to the presence of the six characteristics could have been influenced by the preceding overall vividness rating.

To examine the effects of the order of ratings, we carried out a fourth experiment in which the procedure for experiment 3 was slightly modified; our subjects were now required to rate, first, the presence of each of the six characteristics, and then the overall vividness of the image generated.

In this way we expected to find a still greater weight of the six characteristics on overall vividness ratings, as in fact happened. The characteristics entered the stepwise multiple regression analysis in the following order: shape-and-contour ($B = .24$, $p < .001$), colour ($B = .20$, $p < .001$), detail ($B = .16$, $p < .001$), generality ($B = -.13$, $p < .001$), and saliency ($B = .14$, $p < .001$). Only context did not significantly enter the final equation. Before processing the overall image, context probably has no influence on overall vividness ratings, since subjects base their final ratings on the properties of the critical item rather than on those of the context. The negative sign concerning generality, found in both experiments 3 and 4, was due to the fact that low values in generality (i.e. high in specificity) contributed to the high vividness ratings.

The main result concerning vividness found in this series of experiments was that all of the six identified characteristics contributed in some way to the vividness of an image. When images are generated using only one characteristic for different stimuli (experiment 1) or for the same stimuli (experiment 2), any of the six characteristics of the image (shape-and-outline, detail, context, colour, generality and salience) influences vividness ratings to a similar extent. Further, when an image generated and rated in vividness, either immediately (experiment 3) or as the outcome of a progressive construction (experiment 4), is considered at the same time for all the six characteristics, then all the characteristics seem to influence the vividness ratings in a specific way; however, some of them (first of all shape-and-contour) are more likely to influence vividness than others whose influence changes depending on the procedure adopted (see for a discussion Cornoldi et al, in press).

The second series of experiments illustrates how the consideration of the different vividness properties of the images may help in studying critical problems concerning visual imagery. In particular we focused on the relationship between visual perception and visual imagery. If we consider that vividness ratings are related to the extent to which the image approaches actual visual experience, our goal of comparing vividness ratings for perceived and for the corresponding imaged patterns could appear trivial and circular.

Therefore our hypothesis that vividness ratings may be used to demonstrate some differences between perception and imagery may be particularly challenging. Indeed, our preceding study had shown that other characteristics also influenced vividness ratings. Moreover we reasoned that the reference to visual perception which can be used during vividness ratings may be based on the extent to which properties which render a perceived object vivid are also present in the image. In this case the same properties (or others related to the specific medium) could be active in a different manner for a perceived object and for a corresponding imaged object. We formed this expectation on the basis of the idea that an image is the result of constructive cognitive processes, whereas some forms of perceptual activity are immediate and do not require the use of attention and cognition.

In particular, studies on visual perception (see e.g. Treisman, 1986) have shown that some properties of objects such as rotation, colour and movement are immediately processed by the perceiver, and that some modalities of these properties (such as high colour contrast vs low colour contrast) emerge more clearly and probably more vividly (pop-out effect). In our opinion this effect is specific to perception; in imagery, given its different psychological processes, it should tend to disappear.

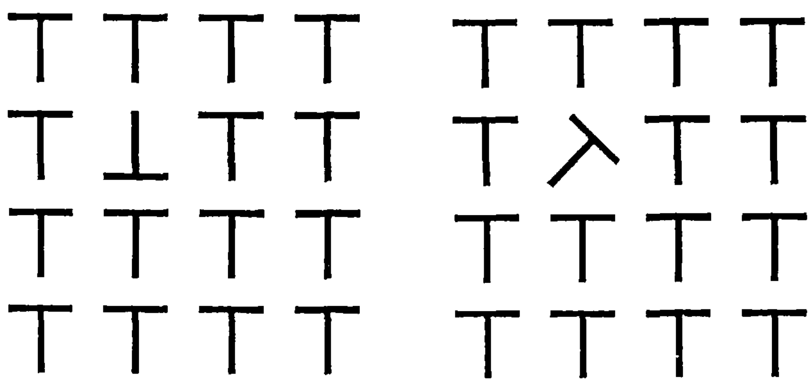

Figure 1. Example of letter matrix

To explore this hypothesis, Giusberti, Cornoldi, De Beni and Massironi (in press) (see also Rocchi, Cornoldi & Massironi, 1990) invited different groups of adult subjects to rate (giving a number between 0 and 99) the vividness of perceived patterns or of the corresponding imaged patterns. The patterns presented consisted of matrices of letters. In each matrix four rows of 4 letters each were included with the letter in the position 2,2 in a particular condition (see Figure 1 for an example).

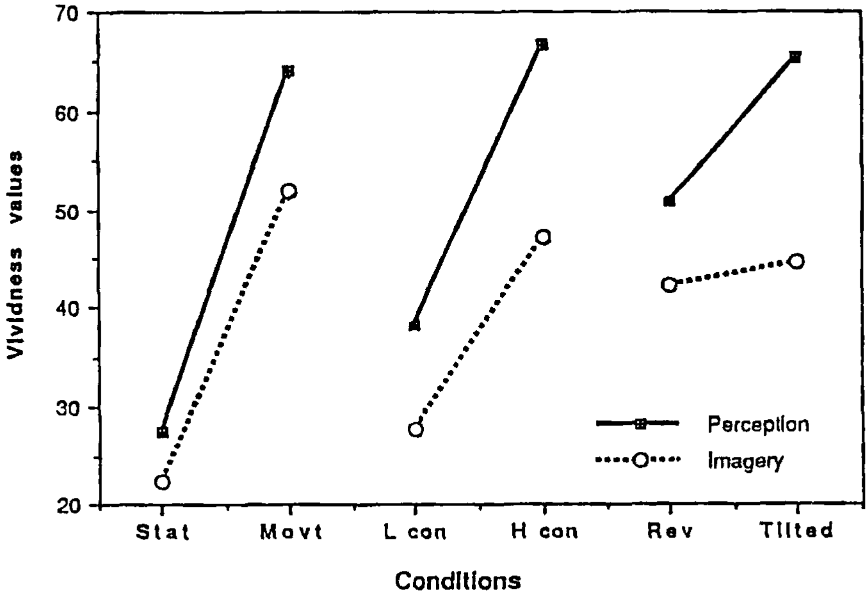

Figure 2. Trend of mean vividness ratings subdivided into three conditions: static vs. movements, low vs. high colour contrast, and reverse vs. tilted.

In a first experiment, the critical letter was presented according to three conditions. For each of these one presentation was favourable to the emergence of a perceptual pop-out effect and the other one was unfavourable. For the movement condition (actual movement of the letter during the experiment), the favourable presentation was represented by the movement of the letter along the second row from left to right and the unfavourable presentation was due to the absence of any movement (static). For the colour condition high and low contrasts were, respectively, the favourable and the unfavourable presentations. Finally, for the rotation conditions the favourable presentation was determined, for the series of the Roman alphabet letters we had selected (F, L, Q, R, T, U), by a 45° inclination of the letter, whereas the unfavourable one was determined by its rotation through 180° (reversed presentation). Subjects were asked either to observe or to imagine different matrices and to rate the phenomenological quality of the critical letter in the 2,2 position with respect to the overall matrix, giving vividness ratings.

The results obtained confirmed our expectations (see figure 2). In particular, the far greater vividness of inclined vs reversed letters in the perceptual modality tended to disappear in the imagery modality, determining a highly significant interaction between modalities and patterns.

In a follow-up experiment we tested the hypothesis that the difference between perception and imagery was due to the immediacy of the perceptual activity involved. In fact, in this experiment we introduced a third condition in which the perceptual vividness rating followed a constructive attentional phase involving drawing the required patterns (drawing modality). Due to the complications involved in the drawing modality, we limited the task to three different presentations of the critical letter: static, inclined and reversed. In all other respects the procedure was identical to that of the preceding experiment.

Again we found a clear difference between the ratings of perceived and imaged matrices; this was especially true in the case of the inclined letter presentation. Further, ratings in the drawing modality were different from those in the standard perceptual modality, thus suggesting that different processes were involved.

The similarities between the drawing and the imagery modality were conspicuous, though not complete, thus suggesting that the difference we had found between perception and imagery was at least in part due to the constructive processes involved in imagery.

REFERENCES

Betts, G.H. (1909). *The distributions and functions of mental imagery.* New York: Teacher's College, Columbia University.

Cornoldi, C., De Beni, R., Cavedon, A., Mazzoni, G., Giusberti, F., Marucci, F. (in press). How can a vivid image be described? Characteristics influencing vividness judgements and the relationship between vividness and memory. *Journal of Mental Imagery.*

Cornoldi, C., De Beni, R. & Pra Baldi, A. (1989). Generation and retrieval of general, specific and autobiographic images representing concrete nouns. *Acta Psychologica* **72**, 25-39.

Denis, M. (1982). Imagery while reading: a study of individual differences. *Memory and Cognition* **10**, 540-545.

Giusberti, F., Cornoldi, C., De Beni, R., Massironi, M., (submitted for publication). Differences in vividness ratings of perceived and imaged patterns.

Marks, D.F. (1972). Individual differences in the vividness of visual imagery and their effect on function. *In* P. Sheehan (Ed.), *The function and the nature of imagery.* New York: Academic Press.

Paivio, A., & Harshman, R.A. (1983). Factor analysis of a questionnaire on imagery and verbal habits and skills. *Canadian Journal of Psychology* **37**, 461-483.

Rocchi, P., Cornoldi, C., & Massironi, M. (1990). Il contributo della percezione visiva alla immaginazione visiva: giudizi di vividezza in ciechi e vedenti. *Contributi dei Dipartimenti e degli Istituti di Psicologia italiani* **3**.

Treisman, A. (1986). Properties, parts and objects. *In* K.Boff, L.Kaufman & J.Thomas (Eds). *Handbook of Perception and Performance* **vol.2**. New York: Wiley and Sons.

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 21

Word meaning and the links between the verbal system and modalities of perception and imagery or In verbal memory the eyes see vividly, but ears only faintly hear, fingers barely feel and the nose doesn't know

Nick Ellis
University College of North Wales, Bangor, United Kingdom

INTRODUCTION

When people are asked to learn lists of words, the greater the imageability of a word, the more likely it is to be recalled. This is a robust effect in Free Recall (FR) experiments (Paivio, 1971). It is even more reliable in Paired-Associate Learning (PAL) (Paivio, 1971; Rubin, 1980). It has withstood many attempts to demonstrate that its association with recall is spuriously attributable to attributes such as meaningfulness (Paivio, Yuille & Smythe, 1966; Dukes & Bastian, 1966; Christian, Bickley, Tarka & Clayton, 1978; Rubin, 1983), concreteness (Christian et al., 1978), familiarity (Paivio, 1968; Frinke, 1968), or age-of-acquisition (Gilhooly & Gilhooly, 1979).

In Paivio's 'Dual Coding Theory' abstract words (of low imageability) have only verbal semantic representations in memory and only these representations and those for concepts associated in meaning are accessed following exposure to the word. In contrast, words with high imageability are represented not only in this semantic system but also in an imagery code, as "sensory images awakened" (James, 1890). "Concrete terms such as *house* readily evoke both images and words as associative (meaning) reactions, whereas abstract words such as *truth* more readily arouse only verbal associations. The meaning of the latter is primarily intraverbal." (Paivio, 1971, p. 85).

Thus images and verbal processes are alternative coding systems, or modes of symbolic representation, which are developmentally linked to experiences with concrete objects and events as well as with language.

In verbal learning tasks the stimuli must be represented in memory and associations forged with other items, these inter-item associations serving as retrieval cues making other items more accessible (Tulving & Pearlstone, 1966): "If S can discover or learn a simple rule or principle which characterizes the items on a list and which relates them to one another, then he uses that rule as a retrieval plan in reconstructing the items from memory, with a consequent improvement in his performance" (Bower et al., 1969, p. 340). The effects of imagery in such experiments may result from imageable words having richer representation as a result of their associations in the modalities of perception, and/or these imagery representations may allow greater associative linkages with other items, these facilitating retrieval.

Prior analyses of imagery have concentrated on the visual system and Paivio has proffered different explanations of visual imageability effects, at times emphasising the number of different codes (*coding redundancy*), at times the richness of representation in vision (*coding richness*) and at times the parallel nature of the visual system (*visual parallelism*). These options are not mutually exclusive, but it would be well to assess their separate effects by investigating the effects of different imagery systems. That such different modalities of imagery exist is demonstrated both by phenomenal reports and by clinical cases of loss: Betts (1909) showed that 95% of the college students that he studied could invoke mental images corresponding to visual, auditory, cutaneous, kinaesthetic, gustatory, olfactory, and organic sensory fields, and Rubens (1979) reviews evidence for the dissociation of visual, auditory and tactile agnosias. The visual system differs from the others both in terms of its massive parallelism and its richness of representation, and thus a word's imageability in these other modalities may not operate in verbal learning tasks in the same way as does its visual imageability.

In this paper we therefore derive word norms for imageability in visual, auditory, olfactory and touch modalities and investigate the effects of these separable factors in (i) incidental learning whilst the subjects were rating the words, (ii) intentional learning followed by either null, picture imagery, or sound imagery interference, (iii) PAL, and (iv) the Stroop effect. We then ask subjects to consider prototypically visual, auditory, smell and touch words for their meaningfulness and their attributes, and to imagine the referents in their 'home' sensory modality and report (i) the attributes that they perceive through this home sensory modality, (ii) all other percepts which are simultaneously available in that imagery episode, and (iii), as many items as possible that are perceptually similar to the referent in that modality. Their generated responses are then analysed for modality differences and for the degree to which the different modalities and aspects (similarity, contemporaneity, etc.) contribute to meaningfulness and memorability.

NORM DERIVATION

Five lists of words were prepared: words with high visual, auditory, olfactory, or tactile associated activity, along with a set of words with none of these associations.

Five raters listed the nouns which represented the objects which most readily came to mind as receiving one modality of sensory analysis. From these initial lists four sets of 40 nouns were selected. An additional set of 40 more abstract words was compiled where each word referred to a concept which the subject used fairly frequently but where the object referent did not typically receive sensory analysis. The five sets were matched for word frequency The four sense modalities were, as far as possible, unassociated - each word predominating in one sensory modality only.

Sixteen volunteers rated the 200 words first for familiarity of conscious recognition of use of the concept, then the four modality ratings followed in a Latin square ordering. After rating them the subjects recalled as many of the words as possible.

The norms for these ratings on each of the five dimensions are available from the author, as are fuller descriptions of all subsequent methods and results. This corpus is used in subsequent experiments.

To assess the weight of the contribution of the different sensory modalities to traditional judgements of overall imagery value, the present ratings for vision, olfaction, audition, touch and familiarity along with word frequency were used as predictors of the Gilhooly & Logie (1980) and Paivio, Yuille & Madigan (1968) imagery and concreteness ratings in full multiple regressions. In this and subsequent regression analyses we report the standardised regression coefficient, ß, for the significant predictors from full multiple regressions where the full set of predictors was forced into the regression equation; ß thus reflects the *independent* contribution of each predictor, controlling for covariance with other effects. Vision was by far the greatest contributor to Imagery ratings ($ß=0.80$, $p<0.001$ and 0.98, $p<0.001$ respectively), along with a smaller, but significant, effect of olfaction ($ß=0.17$, $p=0.01$ for Gilhooly &

Logie). Audition was a significant predictor for the Gilhooly & Logie set (ß=0.28, $p<0.01$), but not for the Paivio, Yuille & Madigan overlap words. In neither case did touch show any significant predictive power. Overall these ratings explain most of the variability in imagery ratings (R^2s of 78% and 95%).

The ratings also explained much of the variance in concreteness (R^2s of 83% and 90%). The relative contributions of the different modalities were much the same as for imagery, except touch was a stronger predictor of concreteness.

There is, understandably, somewhat less explanation of Paivio, Yuille & Madigan meaningfulness (R^2=51%), and again there is a significant effect of the visual modality alone (ß=0.62, $p<0.001$) .

FREE RECALL - INCIDENTAL LEARNING

The rating exercise encouraged subjects to concentrate on associations in a particular perceptual domain, and it is likely that this emphasis at encoding would affect later recall.

Table 1. Full multiple regression solutions for prediction of word recall from incidental learning during ratings of word sensory-association.

Predictor Variables	Dependent variable: Recall from rating conditions (200 words)				
	Vision beta	Olfaction beta	Audition beta	Touch beta	Familiarity beta
Word associate					
Visual imagery	0.50**	0.32**	0.25**	0.09	0.13
Olfactory imagery	0.09	0.50**	-0.05	0.12	0.06
Auditory imagery	-0.01	-0.02	0.46**	-0.02	0.02
Touch imagery	-0.25**	-0.19*	-0.02	0.43**	0.05
Familiarity rating	-0.02	-0.14	0.01	-0.12	0.24**
T&L Frequency	0.26**	0.12	0.14	0.17*	0.03
R^2	0.24	0.32	0.30	0.23	0.14

*$p<0.05$ **$p<0.01$

The words recalled after incidental learning during each of the 5 different rating conditions were analysed using full multiple regressions to investigate the predictive power of each of visual, olfactory, auditory, tactile imagery, concept familiarity and frequency on recall. There was an effect of visual sensory-association on word recall from three of the rating conditions (visual, auditory and olfactory). Furthermore, as is shown in Table 1, there were modality of processing effects whereby the dimension being used by the subjects in the rating task determines that words high on that dimension were better recalled after the completion of the rating task.

When subjects are oriented towards modality-specific attributes of words, it is these attributes which determine later recall. At the time of recall it is associations within and seeded from this modality which cue retrieval for recall. However, it is not the case, except perhaps for visual imagery associations, that the word's attributes in all modalities are automatically accessed.

FREE RECALL - INTENTIONAL LEARNING

Do the strong effects of imagery on FR occur for perceptual modalities other than vision?

273 school children participated in one of the three conditions - control, picture or sound interference. They saw a 100 word subset of the corpus in a FR experiment. In the control condition the retention interval was unfilled; in the picture interference condition it was filled by the presentation of 15 pictures of famous scenes which were to be learnt and recalled; in the sound interference condition the subjects learnt and recalled 15 environmental sounds.

The average number of words recalled was 25 in the control condition, 21 under picture interference, and 15 under sound interference. The interference treatments were thus decreasing overall recall.

Control Condition. The recall scores were entered into a full multiple regression with the 4 sensory modality ratings, familiarity, frequency and serial position (SP) (as a U shaped transform) as independent variables. The only significant predictors were SP ($\beta=0.53$, $p<0.001$) and visual imagery rating ($\beta=0.33$, $p<0.01$).

Picture Interference. The same multiple regression analyses were performed for the word recall data after picture interference. Although the significant effect of SP remained ($\beta=0.53$, $p<0.001$), visual imagery rating now failed to be a significant predictor ($\beta=0.13$, n.s.). None of the betas for the other imagery rating scales changed markedly in their magnitudes from the control condition. Thus the picture interference task specifically interfered with visual imagery processes.

Prior attempts to selectively interfere with the imageability effect in FR have been largely unsuccessful. Thus there was no differential effect of pursuit-rotor tracking on recall of concrete and abstract nouns whether the tracking was concurrent with learning (Baddeley, Grant, Wight & Thomson, 1974) or recall (Warren, 1977). We believe this occurs because, although these interference tasks are using that spatial modality and disrupting its active control processes, unlike the present picture interference task they do not interfere with (i) the relevant associative pathways within that modality, and/or (ii) the referential pathways between the visual representations and semantic associates, and this finding is more consistent with models where the effect of imagery in FR is due to mediation between items rather than 'raw perceptual traces'.

Sound Interference. Here the significant predictors of recall were SP ($\beta=0.46$, $p<0.01$), familiarity ($\beta=0.36$, $p<0.01$) and visual imagery rating ($\beta=0.24$, $p<0.05$).

Recall of Words with Low Visual Imageability. The large contribution of visual imageability may be swamping the potential effects of these other dimensions. To investigate the effects of these other modalities for words which have a negligible visual imageability rating yet which are high on these other dimensions the item pool was restricted to words with low visual imagery ratings. The 41 items included most words from the auditory set (16), and all of the abstract words (20) yet there was still no effect of auditory imageability in FR. Furthermore, even when the visual imagery variance is severely restricted, it remains a significant predictor of FR in all three conditions (control $\beta=0.34$, $p<0.05$; picture interference $\beta=0.34$, $p<0.05$; sound interference $\beta=0.53$, $p<0.001$). Thus the effects of visual imageability are pervasive and remain even when its variance is tightly constrained to low values.

The only imagery dimension which predicts FR is visual imagery. Neither olfaction, audition nor touch imagery seems to play any role. The incidental recall findings demonstrate that people can use modality specific imagery associations beside vision. In contrast the FR results show that under normal circumstances either they do not, or that those that automatically come to mind are somewhat impoverished and lacking in extent of interconnectivity compared to those within the visual system.

PAIRED ASSOCIATE LEARNING (PAL)

PAL depends on the subject finding and remembering some kind of relation between the two words and the most robust visual imagery effects are found in PAL. It is thus an ideal medium for investigation the mediational role of different modalities of imagery in verbal learning.

72 undergraduates saw 100 paired-associates. These were subdivided into 25 different categories where the stimulus word was from either the visual, olfactory, auditory, touch or abstract sets and these were crossed with response words from the same range of categories. There were thus 4 examples of each of the 25 categories (V-V, V-O, ..,O-V, O-O, ...A-A).

The paired-associate recall data were analysed using a full multiple regression with the set of independent variables being visual, olfactory, touch and auditory imagery and familiarity ratings for each of the pair, stimulus and response, and SP.

As in the FR experiments, the only significant predictors were visual imageability (stimulus $\beta=0.35$, $p<0.05$, response $\beta=0.40$, $p<0.01$) and SP - there was no effect of imageability in the other modalities.

How might visual imagery aid mediation in PAL? There are a number of possibilities: (i) The concreteness of the items might allow a greater number of meaningful associative linkages, the image of a complex object including its many parts and attributes and thus awakening semantic associations from all of these parts and those which are similar either in image or meaning *(CODING RICHNESS)*, (ii) Any two images can be juxtaposed in the same 'still life imaginal frame' perhaps interacting in some vivid way in an integrative scene, the stimulus and response being thus relationally associated by capitalising on the parallelism of vision *(VISUAL PARALLELISM)*, (iii) If the two images are linked in a unique context then the awakening of the stimulus image at recall will involve implicit retrieval of contextual information that will serve as a maximally functional retrieval cue for the response image (*cue specificity*, Watkins & Watkins, 1975). These options are by not mutually exclusive and there is evidence for a separate contribution from each. Thus Bower (1970) demonstrated better recall following PAL where the subjects imagined the two objects denoted by the words interacting in some vivid way in an integrative scene than when they were imagined non-interacting, far separated in the "left versus right sides of the imaginary visual field". Yet Winograd and Lynn (1979) later showed that when separation imagery is used but each image is made in a distinctive context (the first pair in an imagined movie theatre, the next in a soccer field, etc.) then performance improves to approximate to interactive imagery levels.

In PAL the major determinant of success is thus the degree to which the stimulus and response words are strongly yet uniquely associated. The studies of Bower (1970) and Winograd and Lynn (1979) which held item imageability constant yet demonstrated clear effects of strategies of interactive imagery identify a large amount of the effect of visual imagery in PAL to be dependent on the parallel nature of the visual imagery system where two representations can be integrated interactively. The null effects of imagery in the essentially non-parallel auditory, olfactory and touch modalities in our present PAL experiment support this interpretation. However, it is also the case that the effect of visual imagery on PAL continues to increase across its full range (as opposed to being a simple step function where the effect cuts in at some threshold level which would allow image construction, but thereafter holds level), and this is more consistent with the *coding richness* interpretations where the extra associations in a distributed memory system afforded by rich representation in the visual imaginal system allows stronger association between the S and R in PAL, and thus better recall.

The present results therefore demonstrate a unique contribution of visual imageability in PAL and lend support to both the *visual parallelism* and *coding richness* explanations of this effect.

STROOP EFFECTS

The 'Stroop effect' originally concerned the large disruption and delay in the naming of the ink colours of written words which are the names of different colours. However, the phenomenon has been shown to generalise to other words and the imagery value of a word is a strong determinant (Davelaar & Besner, 1988).

The Stroop effect demonstrates automatic access to, and spreading activation of, the associations of the stimulus word. Warren (1972) provides a clear demonstration of this. Subjects first repeated three instances of a category which were presented aurally (e.g. *robin, canary, sparrow*), and then named the colour of ink of another word (the Stroop component). The coloured word was either a word from the spoken list (e.g. *robin*), the category of the word in the spoken list (e.g. *bird*), or a semantically unrelated word (e.g. *pencil*). Category words and list members slowed the colour naming by about 100 milliseconds in comparison with unrelated words. Such results clearly demonstrate automatic semantic access, and it appears from Davelaar and Besner (1988) that imagery associations are also obligatorily activated, hence the imagery Stroop effect. This effect must be attributable to *coding richness* since each trial in traditional Stroop tasks is entirely independent of all others - unlike the above FR and PAL experiments there is no requirement for the subjects to relate any of the items together. Therefore any imagery Stroop effects cannot be explained in terms of *visual parallelism*. The Stroop effect thus allows us to investigate *coding richness* in the absence of any contribution from *parallelism*. Hence we here compare the different imagery modalities for their effects in a Stroop colour-naming task.

Twenty one volunteers named the ink colours in eight trials where the stimulus was stars in each of four colours and the 200 words from our norms.

A stepwise regression with median Stroop effect as the dependent variable and visual, olfactory, touch, auditory imagery, familiarity, word length and frequency as the predictors stopped after just one block when it had entered visual imagery as the only significant independent variable ($ß=0.16$, $p<0.05$).

As with FR, so with PAL and thus with Stroop, only visual imagery has any effect on performance. However, in the Stroop task, in contrast to the learning experiments, each word is independent and there is no premium in relating them across trials, *visual parallelism* explanations are irrelevant here, and the visual imagery Stroop effect must be attributable to the *richness* of representation afforded by vision, these associations being automatically accessed and thus slowing ink colour naming. That there are no such effects of either auditory, touch or smell representation again indicates that vision is special and primary in the representations and associations that it affords.

DISCUSSION

The null effects of imagery modalities other than vision in this wide range of tasks makes the *coding redundancy* hypothesis (more codes = more memorable) untenable: it is only representation in visual imagery which affords greater memorability and greater interference in Stroop tasks. What then is special about visual imagery that might underlie these effects? In PAL there is the relational association afforded by the *parallelism* of vision: Any two images can be juxtaposed in the same 'still life imaginal frame' perhaps interacting in some vivid way in an integrative scene. The null effects of imagery in the essentially non-parallel auditory, olfactory and touch modalities in our present PAL experiment support this interpretation. However, in the PAL experiments it was also the case that the higher the visual imagery (even above the threshold where image generation was possible), the greater the recall. Thus the richer the representation in vision, the greater the recall. This richness effect is also present in the FR and Stroop effects where parallel representations in the same 'still frame' are unlikely mediators. Thus we have the further range of *coding richness* explanations whereby the visual imageability of an item allows the possibility of a greater number of meaningful associative linkages, the image of a complex object including its many parts and attributes and thus awakening semantic associations from all of these parts and those which are similar either in image or meaning.

Samuel Taylor Coleridge (1891) introspected on some of these possibilities:

> "Contemporaneity then, being the common condition of all the laws of association, and a component element in all the materia subjecta, the parts of which are to be associated, must needs be co-present with all. ... But if we appeal to our own consciousness, we shall find that even time itself, as the cause of a particular act of association, is distinct from contemporaneity, as the condition of all association. Seeing a mackerel it may happen that I immediately think of gooseberries, I at the same time ate mackerel with gooseberries as the sauce. The first syllable of the latter word being that which had co-existed with the image of the bird so called, I may then think of a goose. In the next moment the image of a swan may arise before me, though I have never seen the two birds together. In the two former instances, I am conscious that their co-existence in time was the circumstance that enabled me to recollect them; and equally conscious am I, that the latter was recalled to me by the joint operation of likeness and contrast." (Coleridge, 1891, pp. 60-61).

Here the mackerel->gooseberries association rests upon the parallelism of vision, the contemporaneity of episodic imagery; the gooseberries->goose association starts with lexical/semantic associations and traverses referentially to the image of the bird; the goose->swan association is discussed in terms of likeness and contrast, such similarity potentially being both semantic and imaginal. Just as these different factors may serve in Coleridge's reminding so they may serve in learners' subjective organisation and relational mediation in memory

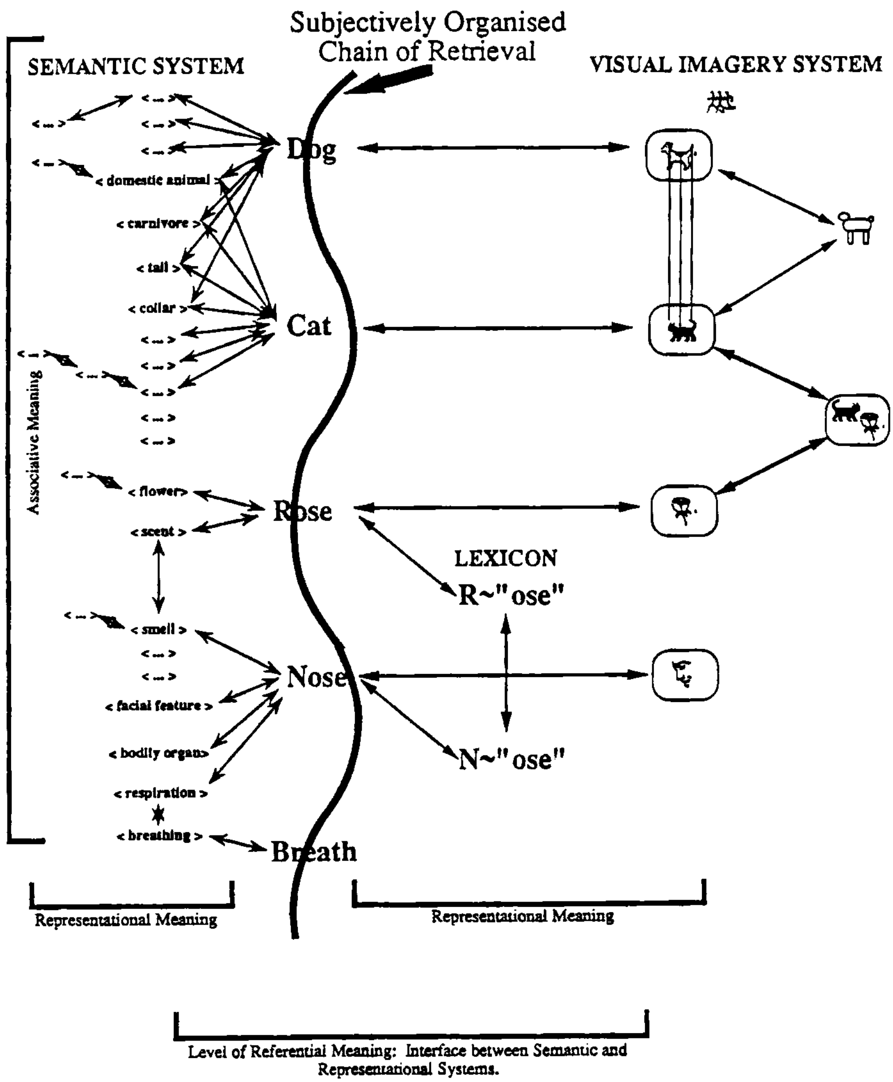

Figure 1. Tangled hierarchies of memory representations

experiments, using "any strategy they can devise to give meaning to an item or pair" (Baddeley, 1976, p.273). We summarise some of these potential associative paths in Figure 1. Imageable items, *dog, cat, rose, nose* are represented in both the semantic and visual systems whereas more abstract *breath* is not. The association between *breath* and *nose* is semantically mediated. *Nose* and *rose,* however, whilst sharing less meaning, may also be related by lexical similarities based on rhyme &/or orthography. *Rose* and *cat* may be linked by relational imagery mnemonics capitalising on the parallelism of vision, as indeed may any visually imageable words. The associations between *dog* and *cat* occur in both semantic and visual systems, in both of which they are exceptionally richly represented with many shared features. Thus at the semantic level they share the more formal connotations of, e.g. *carnivore* or *domestic,* at the visual level they share gross similarities like shape and size (there would be considerable overlap in their stored 3-D model descriptions [Marr, 1982]) as well as fine detail in many attributes down to the level of, e.g., their black noses and the way these shine with spittle, or their paws and the characteristic movements that these allow.

Just as the semantic associations spread and allow subjective organisation in memory so the visual associations prompt visually-based reminding, as Schank (1982, p.26) observes: "Sometimes one thing just looks like another. Since our minds organize perceptual cues and find items in memory based on such cues, it is hardly surprising that such reminding should occur.".

Similar factors operate in perception: just as there is semantic priming whereby recognising a word like *knife* is made faster and more accurate if one has just recognised a related word like *fork* or *spoon* (Meyer & Schvaneveldt, 1971), so the same holds for object recognition where a picture of a knife is recognised more quickly and accurately having just seen a picture of a fork or spoon (Guenther, Klatzky & Putnam, 1980). Importantly, semantic priming also occurs between words and objects: recognising the word *knife* is primed by preceding it with a picture of a fork, and vice versa (Guenther et al., 1980).

Notice that such spreading activation in perceptual priming is *not* tied to a particular interpretation of word meaning: e.g. Swinney (1979) demonstrates the faster visual lexical decision of both the 'insects' and 'microphones' connotations of 'bugs' when primed by the heard context 'Because he was afraid of electronic surveillance, the spy carefully searched the room for bugs'. Fodor (1983), in discussing the Swinney results, states: "associations are the means whereby stupid processing systems manage to behave as though they were smart ones. In particular, interlexical associations are the means whereby the language processor is enabled to act as though it knows that spies have to do with bugs (whereas, in fact, it knows no such thing). The idea is that, just as the [associationist] tradition supposed, terms for things frequently connected in experience become themselves connected in the lexicon. Such connection is not knowledge; it is not even judgement. It is simply the mechanism of the contextual adjustment of response thresholds." (Fodor, 1983, pp. 81-82). It is to be expected that this same unresolved spread of associations that is found in these perception experiments would also occur in the present Stroop experiment, since the same input modules are involved. However, we must show caution in extrapolating to the memory experiments. There are indeed many contrary demonstrations (e.g. encoding specificity, Tulving & Thomson, 1983) that it is the *particular* interpretation of a stimulus which determines its encoding and recall (e.g. the situation where recall is better than recognition - having learned the associated pair *air-port* the subject may fail to recognise *port*, yet will successfully give it as the appropriate PA response to the stimulus *air-*). In these situations where there is context, as is universally the case in the real world, the spread of activation is seriously confined to one particular resolution of meaning. However, all of the memory experiments reported in this paper stem from the unnatural world of the verbal learning tradition where individual words are presented out of context and the subject can choose which of many possible meanings makes for good organisation - the potential spread of activation is unconfined, and remembering, like perceiving, comes from a 'conspiracy of individual memory traces' (McClelland & Rumelhart, 1986) in "tangled hierarchies" of representations in memory (Anderson, 1983). Such experiments, like Coleridge's free associations and Shereshevskii's synaesthesia (Luria, 1968), tell us much about the connotations of words, but potentially less about knowledge.

In this unnatural world of possible meanings rather than those constrained by context, we find these possibilities engendered by semantic associations, experiential contemporaneity, the parts or attributes of the word referent, and its similarity to other concepts. This interpretation of imagery effects in terms of inter-item relational processing rather than the retention of images in some modality-specific form accords both with recent theoretical developments by Marschark and Surian (1989) and with Ryle's (1949) analysis of imagery where he dismisses the existence of mental pictures: "Roughly, imaging occurs, but images are not seen. ... a person picturing his nursery ... is not being a spectator of a resemblance of his nursery, but he is resembling a spectator of his nursery." (Ryle, 1978, p.234).

But we must also ask how much each perceptual modality affords on each of these dimensions of semantic association, experiential contemporaneity, and information regarding the parts or attributes of the word referent, and its similarity to other concepts. The above experiments suggest that Vision affords much more than the other senses with respect experiential contemporaneity and gives more information concerning the parts and attributes of a concept. We next validated these conclusions by studying responses in perhaps the only other situation where individual words are unbounded by context - the free association task. These experiments (details available from the author) involved subjects imagining word referents in particular sense modalities and using the image to generate its attributes, to describe perceptually similar concepts, and to describe other concepts that are contemporaneous in the same imagery episode.

These association generation exercises demonstrated that visually imageable words are more meaningful (in the sense of prompting more free associates), this confirming the oft reported correlation between imageability and meaning, but qualifying it with respect to modality, i.e. it is visual (rather than auditory, smell or touch) imageability that is making the contribution. Quantitative and qualitative analyses demonstrated that vision also predominates in the number of attributes of a word referent that are invoked through its being imaged, and also in the number of associated concepts that are contemporaneously imaged, thus confirming the *richness of representation* and *visual parallelism* explanations of the advantage of visual imagery in memory and Stroop experiments: the connotations of a word, which contribute to both its memorability and its meaningfulness as presently measured, in part result from (i) the parts of its referent and their interrelations, and, (ii) contemporaneous associates in time and space. Vision, in comparison with the other senses, affords much information about the former and more opportunity to perceive the latter.

REFERENCES

Anderson, J.R. (1983) *The Architecture of Cognition.* Cambridge, Massachusetts: Harvard University Press.

Baddeley, A.D. (1976) *The Psychology of Memory.* New York: Harper and Row.

Baddeley, A.D., Grant, S., Wight, E. & Thomson, N. (1974) Imagery and visual working memory. pp. 205-217 in P.M. Rabbitt & S. Dornic

(Eds.) *Attention and Performance* (Volume 5). New York: Academic Press.

Betts, G.H. (1909) *The Distribution and Functions of Mental Imagery.* New York: Teachers College.

Bower, G.H. (1970) Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behaviour* **9**, 529-533.

Bower, G.H., Clark, M.C., Lesgold, A.M. & Winzenz, D. (1969) Hierarchical retrieval schemes in recall of categorised word lists. *Journal of Verbal Learning and Verbal Behaviour* **8**, 481-493.

Christian, J., Bickley, W., Tarka, M. & Clayton, K. (1978) Measures of free recall of 900 English nouns: Correlations with imagery, concreteness, meaningfulness, and frequency. *Memory and Cognition* **6** (4), 379-390.

Coleridge, S.T. (1891) *Biographia Literaria; or, Biographical Sketches of My Literary Life and Opinions; and Two Lay Sermons.* London: George Bell and Sons.

Davelaar, E. & Besner, D. (1988) Word identification: imageability, semantics, and the content-functor distinction. *Quarterly Journal of Experimental Psychology* **40A** (4), 789-799.

Dukes, W.F. & Bastian, J. (1966) Recall of abstract and concrete words equated for meaningfulness. *Journal of Verbal Learning and Verbal Behaviour* **5**, 455-458.

Fodor, J.A. (1983) *The Modularity of Mind.* Cambridge, Mass.: MIT Press.

Frinke, G. (1968) Word characteristics, associative-relatedness and the free-recall of nouns. *Journal of Verbal Learning and Verbal Behaviour* **7**, 366-372.

Gilhooly, K.J. & Gilhooly, M.L. (1979) Age-of-acquisition effects in lexical and episodic memory tasks. *Memory and Cognition* **7** (3), 214-223.

Gilhooly, K. & Logie, R. (1980) Age of Acquisition, Imagery, Concreteness, Familiarity and Ambiguity Measures for 1944 words. *Behaviour Research Methods and Instrumentation* **12** (4), 395-427.

Guenther, R.K., Klatzky, R.L. & Putnam, W. (1980) Commonalities and differences in semantic decisions about pictures and words. *Journal of Verbal Learning and Verbal Behaviour* **19**, 54-74.

James, W. (1890) *The Principles of Psychology.* New York: Holt.

Luria, A.R. (1968) *The Mind of a Mnemonist.* New York: Basic Books.

Marr, D. (1982) *Vision.* San Francisco: W.H. Freeman & Co.

Marschark, M. & Surian, L. (1989) Why does imagery improve memory? *European Journal of Cognitive Psychology* **1**, 251-263.

McClelland, J.L. & Rumelhart, D.E. (1986) A distributed model of human learning and memory. In J. L. McClelland, D.E. Rumelhart and the

PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume 2: Psychological and Biological Models.* Cambridge, Mass.: MIT Press.

Meyer, D.E. & Schvaneveldt, R.W. (1971) Facilitation in recognising pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology* **90**, 227-234.

Paivio, A. (1968) A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behaviour* **7**, 41-49.

Paivio, A. (1971) *Imagery and Verbal Processes.* New York: Holt, Rinehart and Winston.

Paivio, A., Yuille, J.C. & Madigan, S.A. (1968) Concreteness, Imagery and Meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement* **76** (1,2), 1-25.

Paivio, A., Yuille, J.C. & Smythe, P. (1966) Stimulus and response abstractness, imagery and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology* **20**, 362-367.

Rubens, A.B. (1979) Agnosia. In K. M. Heilman & E. Valenstein (Eds.) *Clinical Neuropsychology.* New York: Oxford University Press.

Rubin, D.C. (1980) 51 properties of 125 words: a unit analysis of verbal behaviour. *Journal of Verbal Learning and Verbal Behaviour* **19**, 736-755.

Rubin, D.C. (1983) Associative asymmetry, availability and retrieval. *Memory and Cognition* **11** (1), 83-92.

Ryle, G. (1978) *The Concept of Mind.* Harmondsworth: Penguin. First published by Hutchinson, 1949.

Schank, R.C. (1982) *Dynamic Memory: A Theory of Reminding and Learning in Computers and People.* Cambridge: Cambridge University Press.

Swinney, D. (1979) Lexical access during sentence comprehension (Re)consideration of context effects. *Journal of Verbal Behaviour and Verbal Behaviour* **18**, 645-660.

Tulving, E. & Pearlstone, Z. (1966) Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behaviour* **5**, 381-391.

Tulving, E. & Thomson, D.M. (1973) Encoding specificity and retrieval processes in episodic memory. *Psychological Review* **82**, 261-275.

Warren, M.W. (1977) The effects of concurrent visual-motor distraction on picture and word recall. *Memory and Cognition* **5**, 362-370.

Warren, R.E. (1972) Stimulus encoding and memory. *Journal of Experimental Psychology* **94**, 90-100.

Watkins, O.C. & Watkins, M.J. (1975) Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory* **104**, 442-452.

Winograd, E. & Lynn, D.S. (1979) Role of contextual imagery in free recall. *Memory and Cognition* **7** (1), 29-34.

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 22

Imagery and spatial ability: When introspective reports predict performance

Graham Dean and Peter E Morris
Lancaster University, United Kingdom

INTRODUCTION

Historically, research into the function of imagery in cognition began with the study of individual differences in introspective reports (e.g. Galton, 1883). There has been a commonly held belief that imagery is involved in a variety of cognitive tasks, and in the resulting research, subjective report measures have played a central role. However, research from an individual differences perspective has found very little relation between introspective reports of imagery ability and performance on "objective" tests that are thought to require imagery (e.g. Richardson, 1977; Durndell & Wetherick, 1976). Part of the problem may be that individual differences research has been unsystematic, frequently ignoring empirical advances in other areas of imagery research. As Paivio (1988) points out "the research has consisted mainly of piecemeal attempts to predict memory or some other cognitive skill using scores on a single test that is supposed to measure imagery ability".

When investigating mental imagery by subjective report, researchers still tend to rely on a small set of overvalued questionnaires, the content of which is derived mainly from Galton's original study. Indeed, the content of many of the existing questionnaires seems to have been selected with little empirical or theoretical basis. An examination of existing imagery questionnaires taking into account the large amount of research into the nature of the imagery system reveals several weaknesses.

Several studies have suggested that differences in the content of items imagined affect overall measures of imagery ability. Bat-Zion (1986) found that rated vividness decreased the more transformation was required, the more complex a stimuli and the more imaginative an image was. White and Ashton (1977) found four factors in a factor analysis of Gordon's test of visual imagery control, movement, misfortune, colour and stationary. In studies of our own (Dean & Morris, 1990) we have found that the VVIQ does not appear to be unidimensional, finding four underlying factors corresponding to the four groups of items in the questionnaire. The fact that differences in the type and content of items affects an overall measure of imagery ability, is of obvious concern when one considers the large difference between items contained in established imagery questionnaires and those in the spatial tests with which rated imagery ability is often compared. The items on established questionnaires (e.g. VVIQ, Gordon's) are scenes retrieved from long term memory of everyday objects or events undergoing a wide variety of transformations and containing varying amounts of information. In comparison, items from spatial tests are frequently abstract geometrical shapes not previously encountered and placed in short term memory/the visual buffer by the processes of perception, where the transformation is frequently limited to one specific type.

A second major concern surrounds the properties of images subjects are required to rate on existing questionnaires. Since Galton, vividness has been the most popular measure of individual differences. It is supposed to represent the unique quality of mental imagery, its resemblance to actually perceived stimuli and events. Reliance on vividness as a measure has led many researchers to treat imagery as a single ability and to dichotomise performance in terms of "good" and "bad" imagery ability. However, the findings of Kosslyn and his co-workers (e.g. Kosslyn 1980, 1990) indicate that imagery involves a number of underlying processing components rather than being a single skill. If imagery is used on certain cognitive tasks it is likely that the study of imagery dimensions which can be linked to underlying functional processes, will be of more use in predicting performance on those tasks and in investigating the function of imagery in cognition than will the study of dimensions that bear little or no such relation.

The following study was an attempt to take an initial step in constructing an introspective measure of imagery ability based on current theories and findings on the structure of the imagery system; one that will be of practical use in investigating the role of imagery in cognition.

Two factors were considered in the construction of the questionnaire; the items and dimensions rated. Five items were selected which were either the same as or very similar to the type of items used in the five spatial tests with which the questionnaire was to be compared. The presentation of these items was also in a similar style to that of the items on the spatial tests.

Unlike the VVIQ and the Gordon tests with which the questionnaire was to be compared, a variety of ratings were made for each item imagined. These were of two kinds, ratings of the subjective ease of a hypothesised process (ease of evocation of the image, its maintenance and transformation) and ratings of pictorial aspects of an image that could reflect underlying processes either on their own or by interaction with each other. Throughout the questionnaire the type of transformation required was of a single type, image rotation.

METHOD

Subjects

Forty university undergraduates (18 male, 22 female) aged 18-20 completed five spatial tests and three questionnaires under group conditions in three separate sessions with a one week interval between each testing session. The series of tests were part of a larger study investigating the function of imagery in cognition. All subjects completed the tests in the following order.

First session

CABS: Comprehensive ability battery spatial test (mental rotation 2-D)

Space Relations: selected from the battery of Differential Aptitude Tests Form T.

Vandenberg mental rotation test:(Vandenberg 1978) Based on stimuli from Shepard and Metzler's (1971) experiment.

Paper Folding: (VZ-2) and **Cube Comparison** (S-2) both selected from the Kit of Factor-Referenced cognitive tests (Ekstrom et al, 1976).

Second session

Vividness of Visual Imagery Questionnaire (VVIQ) (Marks, 1973)

Gordon's Test of Visual Imagery Control (Richardson, 1969)

Third session

This session involved the questionnaire constructed for this experiment (see Appendix 1). The questionnaire required subjects to imagine five shapes (see figure 1) that were of the same type as occurred in the spatial tests, and to rate their images on a 1-9 scale (where appropriate) on the following properties;

1)How easy is it for you to evoke this image?
2)How much detail is there in your image?
3)How clear and sharp (in pictorial terms) is your image?
4)How easily can you maintain this image now that you have evoked it?
5)How much does a) the detail, and b) the clarity of your image change when you try to maintain your image?
6)How much of the shape can you form an image of, at any one time? (scale 1-10)
7)How large is your image of the shape?
8)How vivid is your image of the shape?
9)Imagine the shape rotating in the plane of the paper; How easily can you perform the rotation?
10)Is the rotation you can imagine continuous or in discrete stages?
11)Does the size of the imagined shape (or the apparent distance at which it is imagined) change when you rotate it? If so, in what way does your image of the shape change?
12)Does rotating the shape affect a) the detail and b) the clarity of your image in any of the way?
13)How much of the shape do you form an image of, when rotating it. (1-10 scale)
14)Finally, how vivid is your image of the shape rotating?

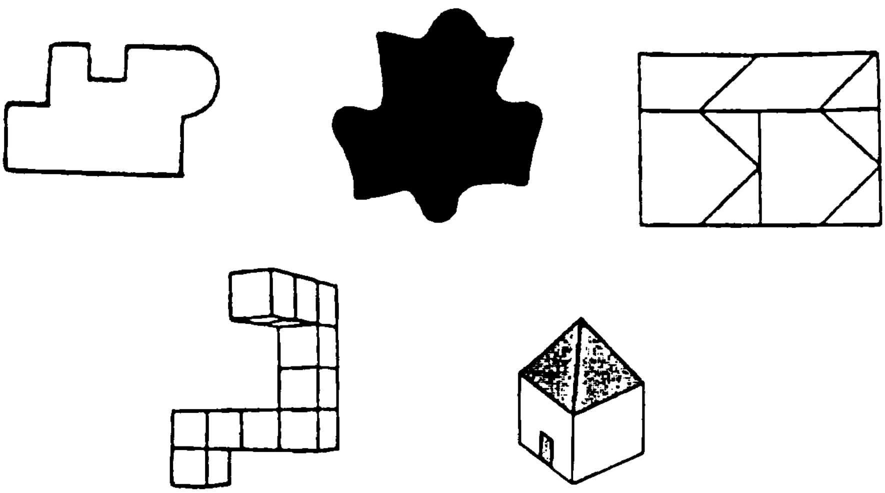

Figure 1. Shapes used for the constructed questionnaire.

Subjects were instructed not to look at the presented shape whilst imagining it, and to try to rate each property separately and independently of the others. There was no time limit for the completion of the questionnaire. After completion subjects were questioned as to whether they had any difficulty in making the ratings or in differentiating one rating from another.

RESULTS

The total scores for each rating across all the five shapes in the questionnaire were compared with the scores on the five spatial tests and the scores on VVIQ and Gordon (see correlation matrix Table 1).

Analysis of the correlations revealed that scores on the VVIQ and Gordon did not consistently predict performance on the spatial tests. Performance on Gordon's questionnaire was significantly correlated with that on The Vandenberg test ($r=.41$, $p<.01$), VVIQ scores were significantly correlated with scores on Space Relations and Paper Folding tests ($r=-.35$, $p<.05$; $r=-.39$, $p<.05$ respectively). VVIQ correlated with Gordon's significantly ($r=-.39$, $p<.01$). It should be noted that the lower the score on the VVIQ the more vividly the images were rated.

Examination of the correlations of the ratings on our questionnaire revealed that ease of maintenance was significantly correlated with scores on all the spatial tests and with scores on the VVIQ and Gordon's; CABS, $r=.39$, $p<.05$; Space Relations, $r=.32$, $p<.05$; Vandenberg, $r=.49$, $p<.01$; Paper Folding, $r=.51$, $p<.01$; Cube Comparisons, $r=.35$, $p<.05$; VVIQ, $r= -.51$, $p<.01$; and Gordon's, $r=.44$, $p<.01$. Although maintenance was the only rating to significantly predict performance on all the spatial tests, other ratings on the questionnaire significantly correlated with some of the spatial tests: Ease with Paper Folding ($r=.31$, $p<.05$); Detail with Space Relations ($r=.39$, $p<.05$) and Paper Folding ($r=.48$, $p<.01$); Clarity with Space Relations ($r=.36$, $p<.05$); Vandenberg ($r=.33$, $p<.05$); Paper Folding ($r=.40$, $p<.01$) and Cube Comparison ($r=.34$, $p<.05$). Rated Vividness during rotation correlated with CABS ($r=.32$, $p<.05$) Paper Folding ($r=.31$, $p<.05$) and Cube Comparison ($r=.31$, $p<.05$). It is interesting to note that vividness ratings of a stationary shape did not correlate significantly with any of the spatial tests.

Stepwise regression was performed to investigate which combination of variables best accounted for variance on the spatial tests. The scores from VVIQ, Gordon and each of the ratings from our questionnaire were included in each regression against each of the five spatial tests.

Table 1. Correlation Matrix

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	1.00																					
2	.41	1.00																				
3	.41	.47	1.00																			
4	.43	.70	.50	1.00																		
5	.63	.50	.57	.49	1.00																	
6	.18	.18	.41	.19	.26	1.00																
7	.09	-.35	-.18	-.39	-.09	-.39	1.00															
8	.18	.09	.26	.31	.14	.42	-.56	1.00														
9	.28	.39	.24	.48	.26	.36	-.68	.71	1.00													
10	.27	.36	.33	.40	.34	.33	-.75	.70	.82	1.00												
11	.39	.32	.49	.51	.35	.44	-.51	.69	.59	.55	1.00											
12	-.13	.02	.29	-.17	.07	.36	-.03	.11	.04	-.01	.33	1.00										
13	.01	.01	.29	.01	.16	.47	-.26	.18	.13	.14	.46	.80	1.00									
14	.07	.13	.13	.15	.24	.58	-.38	.49	.32	.31	.43	.36	.36	1.00								
15	.03	.12	-.10	.01	-.03	-.04	-.03	.18	.00	.00	.13	.05	.02	.38	1.00							
16	.26	.09	.19	.22	.21	.42	-.63	.82	.69	.80	.72	.13	.32	.52	.20	1.00						
17	.29	.22	.38	.25	.21	.44	-.47	.78	.55	.52	.58	.18	.15	.52	.17	.64	1.00					
18	.08	.06	.09	-.19	.12	.14	-.07	.18	.13	.10	.18	.30	.06	.36	.15	.21	.29	1.00				
19	.11	.02	-.10	-.14	-.16	.08	.18	-.28	-.18	-.25	-.26	.05	-.02	-.35	-.12	-.31	-.25	-.25	1.00			
20	-.12	.14	.34	.09	.08	.19	.06	.02	-.10	-.18	.17	.62	.55	.10	-.04	-.17	.20	.02	.04	1.00		
21	-.23	.02	.10	-.07	-.07	.06	-.07	.04	-.09	-.04	.18	.49	.51	.04	-.05	-.01	.19	-.03	-.04	.78	1.00	
22	.10	.22	.12	.22	.17	.28	-.22	.48	.31	.20	.45	.19	.20	.67	.46	.37	.61	.35	-.34	.23	.15	1.00
23	.32	.23	.31	.31	.14	.35	-.63	.76	.67	.72	.73	.11	.31	.33	.14	.84	.75	.18	-.27	.10	.19	.45

$p<.05$ $r=.300$: $p<.01$ $r=.394$: $N=40$

Variable Name	Questionnaire ratings
1.CABS	8.Ease of Evocation
2.Space Relations	9.Detail
3.Vandenberg	10.Clarity
4.Paper Folding	11.Maintain
5.Cube Comparisons	12.Detail Change
6.Gordon	13.Clarity Change
7.VVIQ	14.Proportion Imagined
	15.Size of Image
	16.Vividness
	17.Ease of Rotation
	18.Continuous/Discrete Rotation
	19.Size of Image During Rotation
	20.Detail Change During Rotation
	21.Clarity Change During Rotation
	22.Proportion of Image Rotated
	23.Vividness During Rotation

The significant variables in each equation are shown in Table 2.

Table 2. Results of stepwise regressions.

Against CABS

Step	Variable	Multiple R	Rsq	F(Eqn)	SigF
1	Maintain	.3923	.1539	6.366	.016
2	Clarity Change During Rotation	.4963	.2463	5.555	.008

Against Spatial Relations

Step	Variable	Multiple R	Rsq	F(Eqn)	SigF
1	Detail	.3948	.1559	6.463	.016

Against Vandenberg

Step	Variable	Multiple R	Rsq	F(Eqn)	SigF
1	Maintain	.4881	.2382	10.944	.002

Against Paper Folding

Step	Variable	Multiple R	Rsq	F(Eqn)	SigF
1	Maintain	.5141	.2643	12.572	.001
2	Ch Detail	.6284	.3949	11.094	.001
3	Vivid	.6922	.4791	10.117	.001
4	Clarity	.7783	.6057	12.291	.001

Against Cube Comparisons

Step	Variable	Multiple R	Rsq	F(Eqn)	SigF
1	Maintain	.3517	.1237	4.940	.033

A valid criticism of using stepwise regression in the absence of any particular theory is that the order of addition of the variables is somewhat arbitrary based on which correlation is the highest. In our data, some of the ratings other than for 'Maintain' correlate significantly with some of the spatial tests and with each other, and many of the ratings are significantly correlated with scores on the VVIQ and Gordon's. A further test of the importance of the variables was therefore carried out. The correlation of ratings with the spatial tests were obtained partialling out in turn the variance accounted for by scores on the VVIQ, Gordon's and the vividness and vividness during rotation ratings from the third questionnaire. The correlations obtained are displayed in Tables 3, 4 5 and 6.

Table 3. Matrix of partial correlations controlling for VVIQ scores

	CABS	Space	Vanden	Paper F	Cube Comp
Ease	.16	-.13	.20	.12	.10
Detail	.30*	.23	.16	.31*	.27
Clarity	.30*	.16	.29	.18	.41**
Maintain	.40**	.17	.47**	.40**	.36*
Detail Ch	-.13	.01	.29	-.20	.07
Clarity Ch	-.02	-.09	.26	-.10	.14
Proportion	.03	-.00	.07	.00	.23
Size	.03	.11	-.11	.00	-.03
Vividness	.26	-.17	.10	-.04	.20
Ease (Rot)	.28	.07	.34*	.08	.19
C/Discrete	.07	.04	.08	-.24	.11
Size (Rot)	.13	.09	-.07	-.08	-.14
Detail Ch R	-.12	.17	.36	.12	.08
Clarity Ch R	-.24	-.00	.09	-.10	-.07
Proportion R	.08	.16	.08	.15	.15
Vividness R	.33	.02	.26	.08	.11

On partialing out the variance contributed by the VVIQ, Maintenance ratings were still significantly correlated with scores on four of the spatial tests (CABS, Vandenberg, Paper Folding and Cube Comparison). Clarity ratings were significantly correlated with CABS and Cube Comparison, Detail ratings with CABS and Paper Folding.

Partialing out the variance contributed by Gordon's scores again had little effect. Significant correlations remained with Maintain, CABS, Vandenberg and Paper Folding, Clarity and Detail with Spatial Relations and Paper Folding.

The partial correlations obtained by removing scores on the vividness rating on the new questionnaire again had little effect. Maintain ratings still correlated significantly with all the spatial tests except Cube Comparisons, and Detail and Clarity ratings again correlated significantly with Space Relations and Paper Folding scores.

Table 4. Matrix of partial correlations controlling for Gordon's scores

	CABS	Space	Vanden	Paper F	Cube Comp
Ease	.12	.02	.11	.26	.03
Detail	.24	.36*	.11	.45**	.18
Clarity	.23	.33*	.23	.37*	.28
Maintain	.35*	.27	.38*	.49**	.28
Detail Ch	-.21	-.05	.17	-.26	-.02
Clarity Ch	-.09	-.09	.13	-.09	.05
Proportion	-.05	.03	-.14	.04	.12
Size	.04	.12	-.09	.02	-.02
Vividness	.20	.02	.03	.15	.12
Ease (Rot)	.24	.16	.25	.19	.11
C/Discrete	.06	.03	.04	-.22	.09
Size (Rot)	.10	.01	-.15	-.16	-.18
Detail Ch R	-.16	.11	.29	.05	.03
Clarity Ch R	-.24	.01	.08	-.08	-.09
Proportion R	.05	.18	.00	.18	.10
Vividness R	.28	.18	.20	.26	.06

Although the vividness during rotation ratings correlate significantly with CABS, Vandenberg and Paper Folding, partialing out the variance of this variable still leaves Detail correlating significantly with Space Relations and Paper Folding, Clarity with Cube comparisons and Maintain with Vandenberg, Paper Folding and Cube Comparisons.

The final examination of partial correlations is obviously partialing out the variance accounted for by Maintain ratings (see table 6). In this instance neither VVIQ nor Gordon's scores correlate significantly with the spatial tests neither do Detail or Clarity ratings. Interestingly, removal of the variance associated with Maintain ratings shows that Clarity change during rotation ratings correlate significantly with CABS scores ($r = -.33$, $p < .05$) and that Detail change scores correlate significantly with Paper Folding ($r = -.42$, $p < .01$).

Table 5: Matrix of partial correlations controlling for Vividness ratings

	CABS	Space	Vanden	Paper F	Cube Comp
Ease	-.05	.02	.18	.24	-.07
Detail	.15	.46**	.15	.47**	.15
Clarity	.11	.48**	.29	.39*	.28
Maintain	.31*	.36*	.51**	.53**	.29
Detail Ch	-.17	.01	.28-	.21	.05
Clarity Ch	-.08	-.02	.25-	.07	.10
Proportion	-.08	.09	.04	.04	.16
Size	-.02	.10	-.15	-.04	-.08
Ease (Rot)	.17	.21	.34*	.15	.10
C/Discrete	.03	.04	.05	-.25	.08
Size (Rot)	.21	.05	-.05	-.08	-.10
Detail Ch R	-.08	.16	.39	.13	.12
Clarity Ch R	-.23	.02	.10	-.06	-.07
Proportion R	.01	.20	.05	.16	.10
Vividness R	.20	.28	.29	.23	-.07

DISCUSSION

The above results clearly show that ratings of ease of Maintenance, Clarity and Detail measure aspects of the imagery system that are not measured by the overall imagery ratings on the two established imagery questionnaires tested, and that the imagery ability rated in this case is probably involved with the solution of the spatial tests given to subjects. Given the significant intercorrelation of Maintain, Clarity and Detail ratings, it cannot be stated with any certainty what subprocess or subprocesses of the imagery system are measured by these ratings. Rather, these ratings seem to quantify a unique feature of imagery that vividness ratings do not measure to any great extent.

The fact that scores on both the established questionnaires correlate both with the ratings from our questionnaire that predicted performance on the spatial test and those that did not, and that the intercorrelations between some of the new ratings is weak, is further evidence that these are, at best, overall measures of imagery ability; measures giving an overall impression of the interactions of some of the subprocesses of the imagery system. The interaction of subprocesses would account for the item effects found with vividness ratings in other research, different items placing different demands on the subprocesses. If performance on a given task is limited by specific components of the imagery system, an overall measure of imagery ability such as vividness can only serve to disguise possible functional relationships given the problems inherent in introspective measures to start with.

Table 6. Partial correlations controlling for Vividness during rotation ratings

	CABS	Space	Vanden	Paper F	Cube Comp
Ease	-.10	-.14	.03	.13	.04
Detail	.10	.33*	.04	.39*	.22
Clarity	.06	.28	.16	.28	.34*
Maintain	.25	.22	.40**	.45**	.37*
Detail Ch	-.18	-.01	.27	-.22	.06
Clarity Ch	-.10	-.07	.22	-.09	.12
Proportion	-.04	.06	.03	.05	.21
Size	-.01	.09	-.15-	.03	-.05
Vividness	-.02	-.19	-.14	-.07	.17
Ease(Rot)	.09	.07	.23	.03	.16
C/Discrete	.02	.02	.04	-.26	.10
Size(Rot)	.22	.09	-.02	-.06	-.12
Detail Ch R	-.16	.12	.33*	.06	.06
Clarity Ch R	-.31	-.03	.04-	.13	-.10
Proportion R	-.05	.13	-.03	.10	.12

Considering that the ease of maintenance rating predicts performance on the spatial tests and the ease of rotation rating does not, rules out the possibility that subjects are simply rating the ease with which they can manipulate the image in their cognitive workspace when rating ease of maintenance. This implies that there are further vital imagery processes involved in performing a spatial test other than just the transformation of the image. It could be that those subjects who could more easily maintain an image had more "resources" to spare to either transform the image, to keep track of the changes to the image during transformation, or for comparison purposes with other images. Such an hypothesis however, raises large numbers of questions about the nature of the spare "resources" and about the interaction between the resources distributed between the quality of the image formed and those used for the transformation itself. In this context it is interesting to note that removal of the variance associated with maintain ratings (a rating of the ease of an imagery process) in some cases reveals that ratings associated with the quality of the image, significantly predict performance on some of the spatial tests.

Table 7 Partial correlations controlling for Maintenance ratings

	CABS	Space	Vanden	Paper F	Cube Comp
Ease	-.13	-.19	-.12	-.07	-.16
Detail	.07	.27	-.06	.25	.07
Clarity	.07	.24	.09	.17	.18
Detail Ch	-.29	-.09	.16	-.42**	-.05
Clarity Ch	-.22	-.17	.09	-.30*	-.00
Proportion	-.12	-.01	-.10	-.10	.11
Size	-.02	.08	-.19	-.07	-.08
Vividness	-.04	-.21	-.27	-.26	-.06
Ease(Rot)	.09	.04	.14	-.07	.01
C/Discrete	.01	.00	.00	-.33*	.06
Size(Rot)	.25	.11	.03	-.01	-.07
Detail Ch R	-.21	.09	.30	.00	.02
Clarity Ch R	-.33*	-.04	.01	-.19	-.14
Proportion R	-.09	.09	-.13	-.01	.01
Vividness R	.05	.00	-.07	-.12	-.18
Gordon	.01	.04	.25	-.04	.12
VVIQ	.13	-.23	.09	-.18	.11

*=p<.05 **= p<.01

Concern may be expressed about accepting a significance level of 5% with such a large number of correlations. It should be noted that a repeat of this experiment with 120 subjects (Dean & Morris 1990) obtained very similar results with correlations being of the same magnitude as reported here.

Questioning subjects after completion of the new questionnaire revealed that it did not seem to fully capture some aspects of imagery. The most frequently voiced comment was that the detail and clarity of images changed but that this change could not be strictly viewed in terms of more or less. Rather, different parts of the shape imagined became clearer or more detailed whilst other parts became less clear or detailed also implying that clarity was not uniform. It seems from these comments that some measure of the amount of change occurring within the image is needed. Secondly many subjects stated that even though they were unable to imagine the whole shape at one go, they were "aware" that the rest of the shape was present and sometimes switched to imagining different portions of the shape. These comments raise questions about the proportion ratings originally intended to measure one aspect of the capacity of the visual buffer, especially since several subjects reported being able to perform transformations on the whole shape being "aware" of the transformation yet not being able to imagine it all at one time. These comments need to be the basis of a new series of investigations since if valid they limit the role the pictorial aspects of an image play in certain transformations.

No subjects reported difficulty in giving the number of ratings reported, nor were any difficulties reported in distinguishing the differences between rating types, though some subjects raised doubts as to the appropriateness of certain ratings to simple geometrical shapes. This is especially true of detail ratings which although considered different from proportion ratings, some subjects felt would be better phrased as a question about the amount of pictorial information in an image.

Although many of the ratings on the new questionnaire did not seem to predict performance on the spatial tests, it should be noted that the five tests are an extremely narrow subset of the tasks in which imagery is thought to be involved. An extension of this study is currently underway examining how items and scenes of different types and complexity affect the relationship of the ratings on this questionnaire both with each other and with a far wider variety of cognitive tasks. Therefore at the present state of enquiry none of the ratings can be considered redundant.

In conclusion these results support the idea that imagery is a collection of different abilities, some of which can be measured by introspective ratings, and that the abilities employed when solving a spatial task may not be those previously rated on established imagery questionnaires. The results provide a basis for the development of new, more functional introspective measures of imagery, ones that would be effective in investigating the role of imagery in cognition.

REFERENCES

Bat-Zion, N. (1986). The relation of imagery to other mental abilities. *In* D. Russell, D.F. Marks & J.T.E. Richardson (Eds.) *Imagery 2,* Dunedin: Human Performance Associates

Dean, G.M. & Morris, P.E. (1990, unpublished manuscript). *A factor analysis of the vividness of visual imagery questionnaire.*

Dean, G.M. & Morris, P.E. (1990, unpublished manuscript). *Imagery and spatial ability: A detailed analysis of introspective reports.*

Durndell, A.J. & Wetherick, N.E. (1976). The relation of reported imagery tests to cognitive performance. *British Journal of Psychology* **67**, 501-506.

Ekstrom, R.B. French, J.W. Harman, H.H. & Dermen, D. (1976) *Kit of factor-referenced cognitive tests.* Princeton: Educational Testing Service.

Kosslyn, S.M. (1980). *Image and Mind.* Cambridge, Mass: Harvard University Press.

Kosslyn, S.M. Van Kleeck, M.H. & Kirby, K.N. (1990). A neurologically plausible model of individual differences in visual mental imagery. *In* P.J Hampson, D.F. Marks & J.T.E. Richardson (Eds.) *Imagery Current developments.* London: Routledge.

Marks, D.F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology* **64**, 17-24.

Morris, P.E. & Hampson, P.J. (1983). *Imagery and Consciousness.* London: Academic Press.

Paivio, A. (1988). Basic puzzles in imagery research. *In* M. Denis, J. Engelkamp & J.T.E. Richardson (Eds.) *Cognitive and Neuropsychological Approaches to Mental Imagery.* Dordrecht: Martinus Nijhoff Publishers.

Richardson, A. (1969). *Mental Imagery.* London: Routledge & Kegan Paul.

Richardson, A. (1977). The meaning and measurement of memory imagery. *British Journal of Psychology* **68**, 29-43

Shepard, R.N. & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* **191**, 952-954.

Vandenberg, S.G. & Kuse, A.R. (1978). Mental rotations, a group test of three-dimensional spatial visualisation. *Perceptual and Motor Skills* **47**, 599-604.

White, K. D. & Ashton, R. (1977). Visual imagery control: One dimension or four? *Journal of Mental Imagery* **2**, 245-250.

APPENDIX 1. Questionnaire

The individual differences that have been found in people's ability to form mental images of items, have long been of interest to psychologists. Some people have very lifelike "mental pictures", whilst others only "know" that they are thinking about things.

In this questionnaire, you will be asked to imagine several shapes. The shapes will be shown at the top of each page, underneath will be a number of questions. Once you have looked at the shape, we would like you to shut your eyes and imagine it, then answer the questions underneath (It will probably help if you imagine the shape each time you answer a question, rather than trying to answer all the questions from imagining the shape only once). If you can imagine the shape more easily with your eyes open, do so, but DO NOT LOOK AT THE SHAPE WHILST IMAGINING.

The first question will always ask you how easily you can evoke an image of the shape in question, and to rate this on a scale of 1-9. What we are interested in here, is how easily you can evoke the image (or how easily you know you are thinking about it), and ONLY this. We are not interested in the properties of your image at this stage (e.g.: how vivid or detailed your image is), questions on these properties will come later.

The questions underneath will ask you to rate various properties and qualities of the image you have formed. Remember, we wish you to consider ONLY THE PROPERTY OR QUALITY BEING ASKED ABOUT, and to rate each question about your image SEPARATELY AND INDEPENDENTLY of all the others. This is very important, as you may find that your ratings of each property or quality are different.

Concentrate on imagining the actual shape rather than the property you are trying to judge. This sounds difficult, but what we mean, is that we do not wish you to change what you are imagining to improve a particular property. Rather, we wish you to rate the level of that property which occurred "automatically", when you imagined the scene.

Some of the questions will ask you to rate how easily you can perform various manipulations of the image in question. Remember, we only wish you to rate the ease of manipulation, not any aspect of the quality of the image, there are separate questions concerning this.

There is no time limit for this questionnaire, all we wish, is that you think carefully about each item, and give an accurate an answer as possible. If at any time you feel yourself getting tired or loosing concentration, take a rest! REMEMBER, your accurate and honest answers are vital for the validity of this study.

You may begin when ready, if you have any doubts concerning what to do, please ask the experimenter now.

If at any point you have difficulty in rating something, RATE IT AS BEST YOU CAN, then write by the side what you found difficult, or why you found it difficult. The same applies if you do not feel that the ratings capture something essential about your images. However, PLEASE TRY TO MAKE A RATING IN ALL CASES, then write down what is wrong.

IMAGINE THIS SHAPE

1) How easy is it for you to evoke this image?

1 2 3 4 5 6 7 8 9

VERY DIFFICULT — VERY EASY

2)How much detail is there in your image?

1 2 3 4 5 6 7 8 9

VERY LITTLE — A GREAT DEAL

3)How clear and sharp (in pictorial terms) is your image?

1 2 3 4 5 6 7 8 9

NOT CLEAR AT ALL — VERY SHARP AND CLEAR

4) How easily can you maintain this image now that you have evoked it?

1 2 3 4 5 6 7 8 9

NOT AT ALL — VERY EASILY

5) How much does a) the detail, and b) the clarity of your image change when you try to maintain your image?

a)the detail

1 2 3 4 5 6 7 8 9

NOT AT ALL — A GREAT DEAL

More or less detailed?..............................

b)the clarity

1 2 3 4 5 6 7 8 9

NOT AT ALL — A GREAT DEAL

More or less clear?................................

6)How much of the shape can you form an image of, at any one time?

1 2 3 4 5 6 7 8 9

1/10th — All

PLEASE MARK ACCURATELY ON THE SHAPE ABOVE THE PARTS YOU IMAGINE (circle the shape if you can imagine all).

7) How large is your image of the shape?

1 2 3 4 5 6 7 8 9

VERY SMALL (AS IF SEEING IT FROM A GREAT DISTANCE) — VERY LARGE (AS IF SEEING IT FROM VERY CLOSE UP)

8)How vivid is your image of the shape?

1 2 3 4 5 6 7 8 9

NOT AT ALL VIVID — VERY VIVID

9)IMAGINE THE SHAPE ROTATING IN THE PLANE OF THE PAPER; HOW EASILY CAN YOU PERFORM THE ROTATION? (NOTE we only wish you to rate how easy the rotation is, not how good the quality of your image is whilst you are doing it).

1 2 3 4 5 6 7 8 9

NOT AT ALL — VERY EASILY

10) Is the rotation you can imagine continuous or in discrete stages?

CONTINUOUS DISCRETE STAGES

11)Does the size of the imagined shape (or the apparent distance at which it is imagined) change when you rotate it? If so, in what way does your image of the shape change?

1	2	3	4	5	6	7	8	9
(AS IF) MUCH SMALLER								(AS IF) MUCH LARGER

12)Does rotating the shape affect a)the detail and b) the clarity of your image in any way?

a)the detail

1	2	3	4	5	6	7	8	9
NOT AT ALL								A GREAT DEAL

More or less detailed?.......................

b)the clarity

1	2	3	4	5	6	7	8	9
NOT AT ALL								A GREAT DEAL

More or less clear?.........................

13)How much of the shape do you form an image of, when rotating it.

1	2	3	4	5	6	7	8	9
1/10th								ALL

14)If you only imagine parts of the shape whilst rotating it, can you mark these on the shape below.

15)Are there any other ways in which your image of the shape changes when you rotate it? If there are can you say what these are, AND rate the amount of change on a 1-9 scale, where 1=VERY LITTLE.

....

....

....

16)Finally, how vivid is your image of the shape rotating?

1	2	3	4	5	6	7	8	9
NOT AT ALL VIVID								VERY VIVID

This Page Intentionally Left Blank

Part 7

Imagery, Neuropsychology, and Cognitive Science

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 23

A cognitive neuroscience of visual cognition: Further developments*

Stephen M. Kosslyn
Harvard University, Cambridge MA, USA

Visual mental imagery is one of the few cognitive abilities that can be easily related to brain function. It has been shown convincingly that visual mental imagery shares mechanisms with visual perception (e.g., for a review see Farah, 1988), and we know an enormous amount about the neural substrate of vision. In addition, imagery clearly relies on memory, and we also know a lot about the neural mechanisms underlying memory (e.g., Squire, 1987). One reason we know so much about vision and memory is that nonhuman primates have similar systems, and so animal models can be studied to understand these abilities. Animal models are not available for many other cognitive abilities, such as language. In this chapter I outline some ways in which findings about the neural substrates of vision and memory can inspire theories of human visual mental imagery.

In earlier papers, my colleagues and I (e.g., Kosslyn, 1987; Kosslyn, Flynn, Amsterdam, & Wang, 1990) have developed a theory of visual imagery. In this chapter I will indicate how this theory has developed more recently. As before, I use an approach that relies not only on results from neuroanatomy and neurophysiology, but also on computational analyses of how a machine with the structure of the brain could function in specific ways. Before beginning, then, I must briefly outline some key properties of imagery that must be explained.

*Preparation of this chapter was supported by AFOSR grant 88-0012, supplemented by funds from the ONR, in addition to NSF grant BNS-90-09619 and NIH grant N00014-90-J-1826. The chapter was greatly improved by the comments and suggestions of Christopher Chabris, Robert Logie, and Steven Pinker, although I am sure that there is still considerable disagreement between us, I thank them for forcing me to think more deeply about my claims and conjectures. Correspondence should be sent to S. M. Kosslyn, 1236 William James Hall, 33 Kirkland Street, Cambridge MA 02138.

Following this, I will consider the implications of facts about the primate visual system and memory system for how the brain might produce these behaviours.

Key Phenomena to be Explained

As is well illustrated in this volume, imagery is a complex phenomenon that has many distinct facets. I cannot hope to address all of the properties or uses of imagery in this chapter, and hence will restrict myself to a few, particularly central properties of imagery. I will focus on behaviours that reflect the nature, formation, and use of image representations, and will only allude to the role of imagery in learning.

Geometric representation

Visual imagery is used to help one recall information about previously perceived objects and events, to reason about visual and spatial properties of objects, and to learn new information (see Kosslyn, Segar, Hillger, & Pani, 1990). In all of these circumstances, the local geometry of surfaces of objects must be made explicit. Kosslyn (1980) argued that an array representation is an efficient way of serving this end. If images are patterns of points in a short-term memory structure that functions as an array, the spatial relations among portions of an object are depicted.

Generation

One of the most obvious facts about visual mental imagery is that we do not have images all of the time. Images come and go, depending on the situation. Patterns in the array are best viewed as short-term memory representations. Thus, there must be means of both storing visual representations in long-term memory, and activating the representations to form images in the array.

Part of our ability to activate images involves combining images of different objects into novel combinations. For example, one can imagine Margaret Thatcher riding a zebra, and determine whether she could see over the top of the zebra's head. Indeed, much of the power of imagery comes not only from the ability to image new combinations of objects, but also from the ability to generate new patterns; one can "mentally draw" in imagery, producing images of patterns never actually seen.

Inspection

Patterns in an array would be useless if they could not be interpreted. For example, if one is asked to image an upper case letter "a" and then to mentally rotate it 180°, most people can report the shape of the enclosed area (a triangle balanced on its apex). We must have some way of interpreting the patterns in images. Furthermore, we can "zoom in" on isolated parts of imaged patterns or scan across them (see Kosslyn, 1980, for reviews of experiments demonstrating these abilities).

Recoding

Not only can we interpret patterns in images, but we also can encode them into memory (cf. Paivio, 1971). After imagining objects in new combinations, or imaging new patterns altogether, we can remember them.

Maintenance

Many of our imagery abilities are limited by the fact that images require effort to maintain. The more perceptual units that are included in an image, the more difficult it is to maintain (see Kirby & Kosslyn, in press; Kosslyn, 1980).

Transformation

Finally, the ability to transform imaged patterns lies at the heart of the use of imagery in reasoning. For example, we can rotate patterns in images, including in the third dimension so that we "see" new portions as they come into view. We also can imagine objects growing or shrinking (Shepard & Cooper, 1982), and probably can perform many other types of transformations as well.

Any theory of imagery must provide accounts for these basic properties. The continued development of a theory of imagery in our laboratory is driven by this requirement. We have found numerous insights into these phenomena by considering facts about the brain, as is discussed in the following section.

A Cognitive Neuroscience of Imagery Processing

Kosslyn, Flynn, Amsterdam, and Wang (1990) described a theory of visual object identification. This theory posits a set of processing subsystems that work together to identify shapes and specify their locations. A processing subsystem corresponds to a neural network or set of related neural networks (i.e., which work together to perform part of an information processing task) and is defined by the type of input it accepts, the operation it performs on the input, and the type of output it produces (which in turn serves as input to other subsystems).

Kosslyn (1987) used an early version of the Kosslyn et al. theory to understand the relationship between visual mental imagery and visual perception. The present chapter continues the development of this effort, and in so doing modifies the more recent formulation of Kosslyn et al. (1990). My key assumption is that visual mental imagery shares processing subsystems with visual perception, which seems reasonable given the confluence of findings from numerous experiments using various methodologies (see Farah, 1988).

In this section I briefly describe each subsystem posited by the Kosslyn et al. theory, as well as how the subsystems are interconnected. In each case, I will describe the role of a subsystem in vision before turning to imagery, and will note the ways in which the previous theory has been modified. The architecture of the system underlying visual object recognition and identification is illustrated in Figure 1.

Input to the system

The input to high-level vision is a representational structure that stores the output from low-level visual processes in perception (i.e., those driven purely by stimulus input, which detect edges, colour, and so on); selected contents of this structure are then passed on for further processing.

Visual buffer. High-level visual processes take as input the patterns of activation in a series of topographically mapped areas of cortex. There are at least 15 such maps in the primate brain (for recent reviews, see Felleman & Van Essen, in press, and Van Essen, Felleman, DeYoe, Olavarria, & Knierim, 1990). I focus on the topographically mapped areas following V1 (and perhaps V2) in the processing stream (V1 apparently is dedicated to low-level visual processing), and conceive of these structures as forming a single functional structure that I call the *visual buffer.* The areas subsumed by this structure are localized in circumstriate cortex in the occipital lobe.

The visual buffer corresponds to the array in the theory of Kosslyn (1980). Kosslyn (1987) noted that the topographically mapped areas of cortex receive connections not only from the lower visual areas, but also from the higher ones. Thus, it is possible that a visual mental image is a pattern of activation in the visual buffer that is induced by stored information, as opposed to input from the eyes (which induces a pattern of activation during perception).

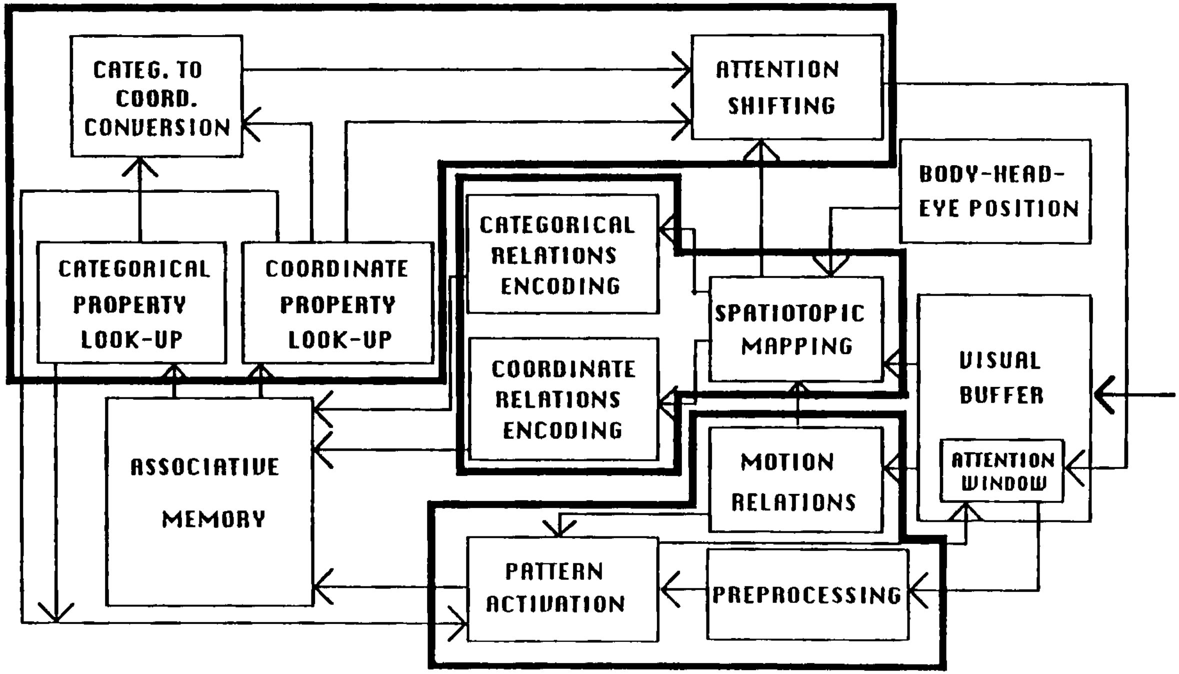

Figure 1. The subsystems of high-level vision

Kosslyn (1980) treated the visual buffer as a static structure, exactly analogous to an array in a computer. This seems overly simplistic. My present view is that the visual buffer itself performs much computation. I suspect that we do not store very complete information in long-term memory, and that when an image is generated the buffer itself must fill in many gaps in patterns. This filling-in process may rely on bottom-up processes that complete fragments that are colinear, fill in regions of the same colour, texture, and so forth. This sort of processing would allow stored fragments to engender a more complete pattern.

If some of the topographically mapped areas used in perception are also used in imagery, then at least some of the limits on our ability to maintain visual mental images make sense: In perception, one does not want smearing as one moves one's eyes from place to place. Thus the visual buffer does not retain patterns of activation long. This property is inherited in imagery, which uses the same structure - and so images fade quickly and require effort to maintain.

Furthermore, another property of the topographically mapped cortical areas allows us to understand why individual parts are hard to "see" when an object is imaged at a small size. "Spatial summation" is a neural averaging over variations within a given region, and is common within these visual areas. This property would also affect images, introducing a "grain" to the array; if objects are too small (i.e., cover too small a region of the visual buffer), details will not be represented.

Attention window. The visual buffer typically contains more information than can be processed during perception (there are more cells in these areas than there are projections to other visual areas; cf. Van Essen, 1985). Hence, some information must be given a high priority for further processing whereas other information must be placed in the background. The *attention window* selects a region within the visual buffer for detailed further processing. The size of the window in the visual buffer can be altered (cf. Larsen & Bundesen, 1978; Treisman & Gelade, 1980). Indeed, Larsen and Bundesen (1978) and Cave and Kosslyn (1989) showed that the time necessary to adjust the size of the attention window increases linearly with the amount of adjustment necessary.

In addition, the location of the attention window in the visual buffer can be shifted, independently of any overt attention shift. Kosslyn (1973) showed that people can scan visual mental images, even when their eyes are closed, and the farther they scan across the imaged object, the more time is required.

However, we do not "bump into the edge" of the visual buffer when we scan; rather, we can scan to portions of objects that initially were "off screen" (see Kosslyn, 1980, for evidence). This can be accomplished if new portions of an image are introduced on one side of the visual buffer and the pattern is slid towards the opposite side (rather like an image on a TV screen as the camera scans over a scene). Similarly, when we "zoom in" on an imaged object, further details of the object become apparent. Thus, there may be a means of fixing a portion of a pattern in the attention window, and adding more details to the pattern as the window is expanded.

Subsystems of the ventral system

A major anatomical pathway runs from the occipital lobes down to the inferior temporal lobes, which has been shown to be involved in the representation of object properties such as shape and colour (e.g., Maunsell & Newsome, 1987; Mishkin, Ungerleider, & Macko, 1983; and Ungerleider & Mishkin, 1982). This "ventral system" receives the information that is selected by the attention window. Kosslyn et al. (1990) decompose the ventral system into three subsystems.

Preprocessing. A vision system must be able to produce the same perceptual representation for an object when it is viewed in different locations in the visual field and from different points of view. Whenever a range of different inputs must be mapped to the same output, one seeks a set of common properties (or overlapping properties, exploiting Wittgensteinian "family resemblances"). Lowe (1987a, 1987b) calls these "nonaccidental properties" (see also Biederman, 1987). For example, properties such as parallel lines (usually indicating edges), line intersections, and symmetries are likely to remain invariant under translation, rotation, and scale changes. Some subsystem presumably computes these useful invariants for subsequent matching against stored information. Not all of the properties are likely to be preserved for all objects, but one cannot know that until the object has been identified; thus, the subsystem must operate in large part purely on the basis of the stimulus input. Kosslyn et al. (1990) hypothesize that such a preprocessing subsystem is implemented in the occipital-temporal area, which receives information from the lower visual areas in the occipital lobes, and sends information to higher visual areas in the temporal lobes.

Lowe's conception of nonaccidental properties is very powerful in certain domains, such as recognizing many manufactured objects. However, many natural objects are not easily described using such properties (e.g., trees, types of fruit, and so on). Indeed, such considerations led J. J. Gibson to emphasize the role of surfaces and texture fields in perception rather than the edge-based properties considered by Lowe. My own view is that the visual system is very opportunistic: Depending on the objects one has to distinguish, one encodes different kinds of information. A problem with this idea, however, is that one cannot know in advance what will be useful. To distinguish a tiger from a leopard, stripes are the key; but one does not thus look for stripes on every object one sees.

Such considerations have led me to revise my characterization of what the preprocessing subsystem does. I now suspect that it groups edges and regions using two kinds of principles. First, following classical Gestalt theory, the subsystem must use some bottom-up processes to group input, forming groups like those noted by Lowe but also grouping areas of similar colour and texture into regions. In the previous theory these functions were carried out in part by a "feature detection" subsystem; I no longer see principled reasons to assume that such a distinct subsystem exists. It is likely that different "channels" exist in the preprocessing subsystem (e.g., see Corbetta, Miezin, Dobmeyer, Shulman, & Petersen, 1990), but the information ultimately is used together to define perceptual units.

Second, I assume that the subsystem can be "tuned" via top-down "training" to organize material. That is, the preprocessing network receives feedback from higher areas so that it can more easily encode visual characteristics that have proven useful in the past. These characteristics can be anything, ranging from a peculiar coloured splotch, to a pattern of light intensity, to a configuration of bumps on a surface; an oddly shaped blotch on a cushion may be just the thing to distinguish one's chair from others of the same type.

Biederman and Shiffrar (1987) describe an unusual example of perceptual learning that seems to rely on this sort of opportunistic encoding. They found that subjects could learn to evaluate the sex of day-old chicks once they learned how to attend to the shape (convex versus concave or flat) of a particular cloacal structure. My view is that perceptual learning actually alters the way we organize perceptual input, changing processes in the preprocessing subsystem. Kosslyn (1987) sketches out an algorithm for such perceptual learning, a variant of which was implemented elegantly by Jacobs, Barto, and Jordan (in press).

The preprocessing subsystem would be used in imagery as part of "image inspection," particularly when imaged objects have been combined in novel ways. In this case, perceptual organizations produced by the subsystem would play a critical role in the matching processes that are carried out in a subsequent subsystem as well as in image retention (described below).

Motion relations. Kosslyn et al. (1990) did not consider an important source of information used to identify objects: characteristic patterns of movement. Such information is used in two ways. First, the visual system can infer "structure from motion." Fragments that move in the same way are grouped together. This organizational principle is very powerful (e.g., Ullman, 1979). Second, motion provides characteristic cues that can be used to identify objects. For example, Johansson (1950, 1975) noted that we can recognize a human form solely on the basis of the patterns of movements of its joints, and Cutting and Kozlowski (1977; see also Cutting and Proffitt, 1981) reported that people can recognize individuals solely on the basis of such information. In addition, it has long been known that neurons in some of the higher visual areas of the macaque respond selectively to different patterns of motion. For example, some neurons in the inferior temporal lobe respond selectively to different patterns of gait (e.g., Gross, Desimone, Albright, & Schwartz, 1984).

Because the computation of motion relations is distinct from the kinds of computations necessary to organize static perceptual units, I posit a distinct motion relations subsystem. Whereas the preprocessing subsystem organizes shapes into perceptual units, the motion relations subsystem extracts key aspects of motion fields, and sends this information to a visual memory (to be discussed in the following section) in which previously encountered motion patterns have been stored. This subsystem is used in imagery in the same way it is used in perception, allowing one to detect previously unnoticed patterns of movement in remembered or novel images.

Pattern activation. Visual memories must be stored somewhere in the system, or recognition could not take place; recognition, by definition, is the matching of input to stored information. Kosslyn et al. (1990) infer a pattern activation subsystem in which visual patterns are stored; these patterns correspond to shapes of objects or parts of objects. We hypothesized, based on results from nonhuman primates, that the pattern activation subsystem is implemented in the inferior temporal lobes.

Each visual memory is composed of a set of perceptual units (positioned in specific locations) and a set of motion relations. The pattern activation subsystem receives both sorts of information as inputs; perceptual units are organized by the preprocessing subsystem and motion relations are extracted by the motion relations subsystem. Both sorts of inputs are matched to the corresponding types of information stored in the visual memory. If both sorts of properties match those associated with a single stored pattern very well, this match is sufficient for object recognition.

Kosslyn et al. (1990) assumed that matching to stored information was performed using the *viewpoint consistency constraint* (Lowe, 1987b). According to this principle, the precise orientation or location of the perceptual units is irrelevant; all that is critical is that the configuration of perceptual units be consistent with seeing an object from a single point of view. This idea fails, however, to account for the wealth of data showing that pictures are more difficult to name in some orientations than others; indeed, the time to name a picture increases with the angular disparity from the upright, with a slight dip in this increase when it is upside down (for a review, see Jolicoeur, in press).

The fact that pictures require more time to name in various orientations suggested to Jolicoeur (in press), Tarr and Pinker (1989), and others that the representation is viewer-centered. Furthermore, they conjecture that viewer-centred input representations are matched directly against these stored representations. One of these two assumptions, about the stored representation or the matching process, must be incorrect, if only because memory for the left-right orientation of pictures is extraordinarily poor (e.g. Nickerson & Adams, 1976). Indeed, when people are asked to name previously seen pictures of objects, they identify mirror-reversed pictures as easily as the originals (Biederman, unpublished data). In fact, Kosslyn and Park (1990) showed that incidental memory for left-right orientation is at chance when previously memorized pictures are subsequently presented to the left visual field/right hemisphere, but are better than chance when they are presented to the right visual field/left hemisphere. This dissociation suggested to us that memory for left-right orientation is accessed separately from the representation of shape per se. (Indeed, if the left dorsal system is in fact better at specifying categorical spatial relations, as discussed below, the result is easily interpreted.)

The sensitivity to planar orientation can be reconciled with the insensitivity to left-right orientation if the stored representation is viewer-centred, but the matching process exploits the viewpoint consistency constraint. In this case, the viewpoint consistency constraint has only limited power, because it is used to match input to a restricted set of information in long-term memory (not a full three-dimensional model as suggested by Lowe). Although use of the viewpoint consistency constraint would match a pattern equally well to itself and its mirror reversal, the sensitivity to planar orientation would result because perceptual inputs are organized differently depending on how a stimulus is oriented. For example, Rock (1973) provided compelling demonstrations that forms are organized at least in part with reference to their gravitational upright. When an object is oriented oddly, at least some of its components may be organized differently in the preprocessing subsystem - and so will not match the information stored in the pattern activation subsystem. It is of interest that most of the effects of orientation are eliminated if a person is warned that an object may appear at an odd orientation - presumably because subjects override the default gravitational coordinate system and instead organize portions of the object relative to each other (for a similar idea, see Jolicoeur, in press).

In short, I am proposing that Lowe's viewpoint consistency constraint must be understood in the context of the effects of orientation on perceptual organization. Depending on the orientation, a pattern is organized into different units, and subsequent matching is between such units.

The claim that shapes are matched using the viewpoint-consistency constraint seems to contradict properties of neurons in the inferior temporal lobe. For example, Perrett et al. (1984) present good evidence that many neurons in the inferior temporal lobe not only are selectively tuned for faces, but also respond selectively to faces seen from particular points of view. Some neurons, for example, respond to the left profile of a face but not to the right, and others respond only if the eyes are pointed in a specific direction. My view is that Perrett et al. may be recording from an area that is used to direct action; this area is near the posterior portion of the superior temporal sulcus, which has rich interconnections to the parietal lobe. This portion of the parietal lobe has a role in directing action (Andersen, 1987; Harries & Perrett, in press, appear to adopt a similar perspective). Viewer-centered information clearly is necessary to guide reaching and other movements. There is no evidence, to my knowledge, that these cells are involved in recognition per se.

If an input does not match any representation very well or matches more than one stored representation to the same degree, additional processing is necessary. In this case, Lowe (1987a, b) found it useful to project back an image of the best-matching object, and then to compare this image, template-style, to the pattern in the input array (which corresponded to our visual buffer; see also Ullman, 1989). The image was rotated and its size adjusted until it matched the input as well as possible; this adjustment process may partially account for the increased time to name misoriented objects (Jolicoeur, in press). This operation is interesting in part because it suggests that imagery may have grown out of mechanisms that evolved to match stored representations to inputs during perception, and once it was available it was then used in other contexts.

Images of individual shapes, then, are formed by activating visual memories top-down, and this process in turn induces a pattern of activity in the visual buffer (Kosslyn, 1987). The areas that presumably are involved in storing visual memories are not topographically organized (Van Essen, 1985), and many of the geometric properties of stored shapes may be only implicit (not explicit) in the representation. By analogy, a list of coordinates does not make all information about colinearity explicit, but such information is implicit in the representation. In order to make local geometric relations explicit, it is necessary to use such stored information to produce a representation in an array format.

Furthermore, according to the present formulation, information about motion is implicit in the long-term memories stored in the pattern activation subsystem; to reinterpret motion, these representations must be unpacked in an image. Again, it is the geometric properties of the visual buffer that allow this information to be made explicit and hence subject to new interpretation; motion is registered by systematic shifts of points from location to location in the visual buffer.

The activation of a visual memory is but one component of visual image generation. As noted earlier, we can create composite images, which requires combining stored memories in novel ways. Furthermore, in some cases we mentally "draw" new patterns, "seeing" shapes that do not correspond to individually stored perceptual units. In order to understand these abilities, we need to consider additional components of the system.

Subsystems of the dorsal system

A second major cortical pathway projects dorsally from the occipital lobes, up to the parietal lobes. The usual description of this pathway is that this "dorsal system" is concerned with spatial properties, such as location, size, and orientation (see Maunsell & Newsome, 1987). Indeed, Ungerleider and Mishkin (1982) identify the ventral and dorsal systems as being concerned with "what" and "where," respectively. I infer that the dorsal system receives information from the attention window at the same time as the ventral system; hence, both systems are computing information about the contents of the same region of the visual buffer.

I have recently revised my thinking about the role of the dorsal system, in large part on the basis of findings in nonhuman primates. As Andersen (1987) and Hyvarinen (1982) point out, a pervasive property of neurons in the posterior parietal lobes is that they fire prior to the animal's initiating a movement or are sensitive to the consequences of a movement. The parietal lobe appears to be concerned in large part with controlling and monitoring movement, and spatial information must be encoded to serve these ends.

The idea that the parietal lobes are not simply concerned with encoding spatial properties, but rather with encoding information to guide action, may help to clarify a longstanding puzzle: In the experiments by Pohl (1972) and Ungerleider and Mishkin (1982), monkeys discriminated between patterns on food lids or between the locations of a small "landmark." When the animals' parietal lobes were removed, their performance on the landmark task was devastated, but they performed the pattern task well; this result is consistent with the idea that the parietal lobes are critically involved in encoding location. In contrast, when animal's temporal lobes were removed, their performance on the pattern discrimination task was devastated, but they performed the location task well; this result has been taken to show that the temporal lobes encode shape.

A problem with these interpretations is that spatial properties of the patterns in the shape task are often sufficient to discriminate among them. For example, a monkey may have had to discriminate between checks and stripes; in this case, there were fewer locations defined by the stripes than the checks, the patterns had different sizes, and they had different orientations (Holmes & Gross, 1984, showed that animals can discriminate orientation even when the temporal lobes are removed). Thus, all of the spatial properties of the patterns were sufficient to discriminate between the patterns. And yet monkeys without temporal lobes are severely impaired at the discrimination, even when the parietal lobes are intact.

I have puzzled over this apparent paradox for years, and only recently had a hint of a possible resolution from the behaviour of a patient studied by Kosslyn, Daly, McPeek, Alpert, and Caviness (1990). The patient had suffered damage to the left frontal lobe and had hypometabolism (revealed by PET scanning) in the occipital-temporal area on the left side. We asked this patient to discriminate between patterns that were formed by filling in cells of a 4 x 5 grid. He had some difficulty encoding patterns, and reported that he remembered the patterns in grids by *looking at each individual cell.* He apparently remembered the patterns as sets of filled locations in the grid. And in fact, the more segments the pattern had, the more time he required to encode them. When the grid lines were removed, so that cells were not clearly defined, he could not use this strategy and his response times changed accordingly; there now was no effect of the number of segments on the time to encode patterns. This difference in response times suggests that the patient was not making the same pattern of eye movements when viewing both types of displays.

One way to understand these results is to infer that the location information is normally encoded in a form suitable for directing action, and can only be used for recognition by making eye movements and recoding the location information into a different format. Think about how easy it is to toss an object into a wastepaper basket, compared to how difficult it is to estimate the distance of the basket from you. I have informally tested a series of people who enter my office, and found that some can throw better than estimate the distance and vice versa for others. The important claim is that there is a dissociation between the two kinds of information. This observation makes sense if the information about location is "encapsulated," and can only be directly used to guide action. McLeod, McLaughlin, and Nimmo-Smith (1985) provide good evidence for such a dissociation.

If so, then the monkeys without temporal lobes may have been unable to discriminate between patterns because they did not hit on the strategy of moving their eyes over the patterns, which would have allowed them to encode the spatial properties in a way useful for identification. It would be interesting to observe whether monkeys without temporal lobes could discriminate between checks and stripes if they had been trained to look at the dark regions of patterns prior to surgery.

Kosslyn et al. (1990) did not consider the idea that the parietal lobes encode spatial information in a format to be used to guide action. This idea leads me to modify Kosslyn et al.'s characterizations of the subsystems in the dorsal system.

Spatiotopic mapping. Location information is specified relative to the retina in the visual buffer (these maps are retinotopic; see Van Essen, 1985). Because a retinotopic representation changes whenever one moves one's eyes, it is not useful for object identification, navigation, or tracking. One needs a representation of an object's location relative to another object or part, not relative to the retina. Andersen, Essick, and Seigel (1985) found cells in area 7a (part of the parietal lobe) of the macaque that respond to location on the retina, as gated by eye position, and Zipser and Andersen (1988) showed that the outputs from sets of these neurons are sufficient to indicate location relative to the head.

I therefore infer a subsystem that receives as input a retinotopic position and the positions of the body, head, and eyes, and computes where an object or part is located relative to other objects or parts. During both vision and imagery, the output from the spatiotopic mapping subsystem is a set of spatiotopic coordinates that are tailored to guide action. Kosslyn et al. (1990) assumed that these coordinates were general purpose representations, but the present view is that they are dedicated for use in guiding actions. This idea has implications not only for how we form images, but for how we decode spatial information from images, as noted below.

Coordinate spatial relations encoding. We often want to store spatial information in memory. For example, to navigate efficiently in familiar rooms, it is useful to store the locations of furniture. This can even allow one to navigate in the dark. Thus, I hypothesize the existence of a subsystem that encodes the types of coordinates used to guide action. This subsystem does not encode motor programs, but rather coordinates that can be used to guide actions.

Fisk and Goodale (1988; see also Goodale, 1988) found that right-hemisphere damaged patients had difficulty in initiating a movement when asked to point at a dot. This result is consistent with the idea that the right hemisphere has a special role in encoding the coordinates that are used to guide actions. A key component of such computation is the precise specification of the location of an object, and hence it is of interest that Hellige and Michimata (1990), Koenig et al. (1990), Koenig, Reiss, and Kosslyn (1990), and Kosslyn, Koenig, Barrett, Cave, Tang, and Gabrieli (1989) provide evidence that the right hemisphere can encode metric spatial information more effectively than the left (see also De Renzi, 1982).

This subsystem can be used in imagery in at least two distinct ways: It can play a role both in image generation when multiple parts are assembled, as will be discussed shortly, and in image inspection, encoding spatial relations among parts of imaged objects.

Categorical spatial relations encoding. Different tasks require the use of different types of spatial relations. Consider the situation in which one is so close to an object that one only sees a small portion of it in a single fixation. In this case, the ventral system would identify parts, and the spatial relations would be encoded via the dorsal system. Many objects, such as a human form, can assume a wide range of positions as the parts move. In order to identify such objects, the spatial relations among the parts should be specified rather abstractly. The fact that the forearm is "connected to" the upper arm remains true no matter how the metric relations between them vary.

The *categorical* spatial relations encoding subsystem encodes relations such as "connected to," "left of," "under," or "above." These representations capture what is stable across instances that may differ in terms of precise metric relationships. As Kosslyn, Chabris, Marsolek and Koenig (in press) review, previous work provides evidence that this subsystem is relatively more effective in the left cerebral hemisphere. This finding is consistent with the long-standing reports that Gerstmann's syndrome, which includes left-right confusion as one component, occurs following damage to the left angular gyrus (e.g., see De Renzi, 1982).

A reinterpretation of the distinction. Sergent (in press) reports that the hemispheric dissociation between coordinate and categorical encoding only occurs when the stimuli are displayed at relatively low contrast. This result puts real pressure on the theory of Kosslyn (1987) and Kosslyn et al. (1990), and has caused me to reconceptualize the theory. The driving force behind the revised conception is a recent finding by Kosslyn, Hillger, Livingstone, and Hamilton (1990).

We asked subjects to view two short line segments presented in succession and to decide whether the lines had the same orientation. Both segments were presented in the same visual field while the subject stared at a central fixation point. The important variable was the distance between the locations of the lines in each pair; they were either relatively close (within 1° of visual angle) or far (up to 8° apart). When the segments were relatively close together, subjects were more accurate if the stimuli were presented initially to the left hemisphere; when they were relatively far, subjects were more accurate if the stimuli were presented initially to the right hemisphere.

One account of these findings hinges on the idea that neurons in the high-level visual areas in the two hemispheres have different sized receptive fields, perhaps because they receive input from different retinal ganglia. It is possible that some of the ganglia, such as the magnocellular neurons (see Livingstone & Hubel, 1987), have a special role in "preattentive" processing. The magnocellular ganglia encode motion and flicker very well, which is useful for guiding eye movements and subsequent "focal" attention (see Neisser, 1967). Furthermore, the magnocellular ganglia have relatively large receptive fields, which would help preattentive processing to monitor the entire visual field[1].

The finding that the right hemisphere encodes spatial location better than the left follows directly from the idea that the right hemisphere monitors larger, more overlapping receptive fields: Computer simulation modeling has shown that relatively large overlapping receptive fields are more effective at using "coarse coding" to register the location of a dot relative to a line than smaller, less overlapping receptive fields (Kosslyn, Chabris, Marsolek, & Koenig, in press). This notion appears to be

[1]I owe the idea that the magnocellular ganglia may project preferentially to the right hemisphere to Marge Livingstone.

consistent with Sergent's own interpretation of her results. In contrast, our computer simulations showed that smaller, less overlapping fields are more effective for dividing space into discrete bins, which correspond to some spatial relations categories (such as above/below or left/right). This idea, then, leads us to expect a left-hemisphere advantage only for some categorical spatial relations, namely those that allow space to be carved into discrete regions. Preliminary results in our laboratory suggest that this prediction is worth taking seriously.

The idea that the left hemisphere typically monitors smaller local regions than the right hemisphere is consistent with numerous findings. For example, the left hemisphere plays a critical role in encoding portions of objects, whereas the right hemisphere plays a critical role in encoding global patterns (e.g., Delis, Robertson, & Efron, 1986). Furthermore, people categorize parts of objects faster when the objects are shown initially to the left hemisphere, whereas they categorize overall shapes faster when they are presented initially to the right hemisphere (see Van Kleeck, 1989, for a review). Although large overlapping receptive fields are good for encoding location, they are not as good for encoding shape. For this purpose, smaller receptive fields provide greater resolution (because they average input over smaller areas).

Thus, the revised theory leads us to expect differences in the ventral systems in the two cerebral hemispheres. Kosslyn (1987) alluded to such possible differences, but did not provide detailed arguments for them. Specifically, the notion that the higher visual areas of the two hemispheres differ in the sizes of the receptive fields they monitor implies that the contents of the pattern activation subsystem may also differ: The left hemisphere may store better representations of separate portions of objects, whereas the right may store better representations of overall shapes.

The claim that the left hemisphere encodes portions of objects more effectively than the right might help to explain another of Fisk and Goodale's (1988) findings: Patients with left hemisphere damage could initiate a reaching movement normally, but had trouble controlling it (particularly in the deceleration phase). Reaching apparently has two phases, initiation (which is open-loop) and fine-tuning (which uses feedback). The right hemisphere may be critical in the first phase because it computes the location of the target better. And if the left hemisphere is more adept at encoding portions of objects, it may be critically involved in orchestrating the second phase of a reach; we typically reach for a portion of an object, such as the handle of a cup or the bottom segment of a pen.

Now let us return to Sergent's finding that the right-hemisphere advantage for encoding spatial coordinates depends on the level of contrast. Our computer simulations showed that if high contrast allows more input units to fire, the differences in the sizes of receptive fields no longer effect the ease of computing either metric distance or discrete bins. When very many units contribute, many of them have overlapping receptive fields and many do not. Thus, the networks can map both functions easily.

To summarize, the revised theory of categorical versus coordinate spatial relations encoding rests on the idea that the right hemisphere monitors larger receptive fields than the left, which is useful for detecting stimuli over the entire field. This information in turn is used to direct movement (such as head and eye movements towards a stimulus). These large fields overlap, conferring high resolution for specifying position via coarse coding. In contrast, by monitoring smaller receptive fields, the left hemisphere is better able to focus in on important characteristics of an object. These smaller receptive fields are also useful for carving space into bins, which may correspond to some types of categorical spatial relations. The differences between the hemispheres are a matter of degree, and when contrast is very high large amounts of all types of input are sent to both hemispheres, minimizing the differences.[2].

Like the coordinate spatial relations encoding subsystem, the categorical spatial relations encoding subsystem can be used in imagery in at least two distinct ways.

Associative memory. The simple fact that people can report from memory where furniture is placed in their living rooms indicates that the outputs from the dorsal and ventral systems are conjoined downstream. Kosslyn et al. infer an associative memory in which such conjunctions are stored. If an object is seen close up, so that it is examined over the course of multiple eye fixations, then associative memory will be used to build up a composite representation of the object and to identify it. During perception, the outputs from the ventral and dorsal systems are matched in parallel in associative memory to

---

[2]One can also ask why the right hemisphere monitors larger fields than the left rather than vice versa. A possible account rests on three ideas. First, the right hemisphere is more mature at birth (Taylor, 1969). Second, the infant, having little information in memory to guide attention, relies heavily on preattentive processes in vision. These processes are more effective if large receptive fields are monitored. Third, once the right hemisphere has been used heavily for this purpose, considerable neural reconfiguration would be required to allow it to be effective in controlling focal attention mechanisms. Hence, when the left hemisphere matures, it is able to accomplish these tasks easier than the right, and the specialization develops. (This idea was inspired by de Schonen & Mathivet, 1989; Hellige, 1989; Sergent, 1988.)

parts and relations of stored objects. The system converges on the identity of the object being viewed by finding the stored representation that is most consistent with the encoded parts and their spatial relations. When such evidence exceeds a threshold (which presumably can be varied, depending on context), identification occurs.

Goldman-Rakic (1987) summarizes evidence that one aspect of associative memory involves structures in the frontal lobes. In particular, she shows that area 46 in the dorsolateral prefrontal lobes is critically involved in storing memory for location. If this area is damaged in one hemisphere, an animal cannot retain in short-term memory the locations of stimuli in the contralateral field. The area is topographically organized; when different portions are damaged, memory is subsequently impaired for different regions of the visual field. Furthermore, Goldman-Rakic shows that areas of the parietal lobes that are involved in encoding spatial properties not only project to the frontal lobes, but also receive rich projections from them.

Associative memory plays a critical role in imagery for at least two reasons. First, this is where information is associated with an object's name. We often form images upon hearing the name of objects. Second, because associative memory integrates the outputs from the dorsal and ventral systems, it must contain representations of the structure of scenes and objects. To image an object that is composed of more than one part, we must access information about the structure of the object and use this information to activate the appropriate visual memories and the appropriate spatial relations representations. This process involves additional subsystems, as noted below.

Subsystems used in top-down hypothesis-testing

We see only about 2° of visual angle with high resolution. Thus, we often must move our eyes over an object or scene during recognition and identification. Logically, there are only three ways in which we can guide eye movements: randomly, on the basis of bottom-up information (e.g., motion), or using stored information. Yarbus (1967) provides ample evidence that knowledge is often used to guide one's sequence of attention fixations. Kosslyn et al. (1990) inferred a set of subsystems that are involved in accessing and using stored information to shift attention.

Coordinate property lookup. Often, the location of objects in a scene or the locations of parts on an object are important in identification. Thus, Kosslyn et al. (1990) postulate subsystems that can access stored information about the spatial arrangement of parts of objects and can use this information to shift attention to relevant locations. The present revision of the theory leads me to characterize the coordinate property lookup subsystem slightly differently from Kosslyn et al.; it accesses stored information that can be used to guide movements precisely. A subsystem that accesses such stored information appears to be implemented in the frontal lobes, near the frontal eye fields (area 8; cf. Luria, 1980).

The coordinate property lookup subsystem seems to be involved in many image generation tasks. For example, if asked to describe where the furniture is in their living rooms, most people move their eyes and report scanning to a location in an image and "seeing" the object. One interpretation of this finding is that the furniture is in fact not present until one scans to the appropriate location, and that such scanning involves activating motor-based coordinate representations of location. These representations are useful for guiding action, and in order to recover a representation of a specific location one must activate a motor program. One often may be able to inhibit the actual execution of the program, but perhaps not completely. Hence, one often moves one's eyes in the course of building up the image (cf. Hebb, 1949).

Categorical property lookup. Categorical representations group positions and treat them as equivalent; in contrast, coordinate representations specify the finest possible distinctions. Hence, the two representations are qualitatively distinct, and Kosslyn et al. (1990) argue that they logically require different operations to access. Thus, Kosslyn et al. (1990) infer a second lookup subsystem that accesses stored information about the categorical locations of objects in a scene or individual parts. This subsystem may also be implemented in the frontal lobes[3].

This idea implies that there are two distinct ways of adding parts to an image, one using coordinate spatial representations and one using categorical spatial representations to specify the parts' locations. If one images one's living room repeatedly, I have observed, one no longer moves one's eyes. It is possible that with repeated use, the motor-based coordinate representation is recoded into a categorical representation. Indeed, Koenig, Kosslyn, and Chabris, and Gabrieli (1990) found that the right-hemisphere superiority for metric judgments disappears after practice, which is consistent with this idea.

Attention shifting. Recent evidence suggests that the human visual system probably includes at least three subsystems that are used to shift

[3]Kosslyn et al. (1990) pointed out that because categorical spatial relations do not specify precise positions, additional processes are necessary to convert such representations to specific locations in a given image. They posited a separate subsystem to perform these conversions. I am no longer certain that this distinction is justified, and will be conservative by assuming for the moment that the categorical property lookup subsystem may perform the necessary conversion by itself.

attention: One that disengages attention from the current location (which appears to involve the parietal lobes); one that shifts attention to a new location in space (which appears to involve the superior colliculus); and one that engages attention at that new location (which appears to involve the thalamus; see Posner et al., 1987). Kosslyn et al. (1990) chose a coarser level of modeling in which all attentional control mechanisms were grouped into a single attention shifting subsystem.

The attention shifting subsystems guide the movement of the body, head and eyes, and also adjust the attention window in the visual buffer (both in perception and visual mental imagery). These mechanisms are important for several reasons. First, they guide image scanning and zooming. Second, they play a critical role in some forms of image generation. Consider, for example, a task developed by Podgorny and Shepard (1978). They showed people empty 5 x 5 grids, and asked them whether a dot or dots would be covered if a specific block letter were present in the grids (the subjects saw the block letters in advance, which were formed by selectively filling in cells in the grid). In this task, one selects specific cells to pay attention to; one does not activate stored visual memories.

The idea that images can be formed by allocating attention also allows us to consider "mental drawing." One can image a line simply by shifting attention over the visual buffer and activating each small region of the buffer in turn. This process will create a representation of a "path" in the visual buffer, which in turn can be processed just like any other pattern of activity (such as those arising during perception).

Thus, we are led to make another new distinction: Some forms of imagery involve activating stored visual memories, whereas others involve engaging attention in specific regions. This distinction leads to a simple prediction: There is no reason why the complexity of an object need affect the time to image it using the first method; if the object is stored as a single perceptual unit, the unit is simply activated. For example, a normal face might be easier to image than a face with scrambled features, even though both have the same number of features. The normal face has been seen so often that there may be perceptual grouping processes built into the preprocessing subsystem that produce a single representation of a face, which can imaged as such; in contrast, the scrambled display cannot be encoded as a single unit, and hence multiple units are encoded and must later be imaged individually. The other sort of imagery does not offer this possible difference; because the attention window can only pick out a regular region in the visual buffer, one will always need to shift it to attend to different regions, and so more time always will be required to image patterns that contain more component parts.

Thus, when imaging a letter in a grid, for example, one will need to attend to each segment in sequence. The more segments in the letter, the longer it should take to form the image. Kosslyn, Cave, Provost, and Von Gierke (1988) confirmed this prediction. In contrast, if one's eyes are closed and one is merely imaging what a previously seen letter looks like, there is no reason to expect that more segments should result in longer times; one simply activates the visual memory. Kosslyn, Hillger, Engel, Clegg, and Hamilton (1990) have confirmed this prediction.

Transformation. Lowe (1987a, b) proposed that when nonaccidental properties do not match the input very well during perception, an image is generated and matched to the input pattern. Lowe's computer vision system tried to maximize the match by rotating the generated image and adjusting its size scale. I have adopted his use of imagery in object recognition, which leads me to predict that there should be two distinct ways of imaging movement. First, if one has stored a visual memory of a moving object in the pattern activation subsystem, it can simply be reactivated. For example, imaging a horse running is simple if the visual memory itself contains information about its movement patterns. This information is purely visual.

Second, if the object was encoded without motion information, this information can be added by changing the spatial representations encoded in the dorsal system. In many cases, the only available representation of location, orientation, and size is encoded in a motor format. In these situations, one must execute motor operations on these representations to alter them. This idea predicts that people sometimes will perform implicit motor movements when transforming shape.

The two kinds of motion information may often be used together. One may not have encoded a pattern of movement over a length of time, but instead registered a succession of moving images. In this case, one will move one's eyes when replaying the image, with the eye movements indicating that the relative locations of the separately encoded images have been activated in the course of integrating them.

When one transforms an object that was not seen moving, one must actually alter the image in the visual buffer. When a three-dimensional object is rotated, new portions of the object will come into view and hence new visual memories must be activated. Thus, it is of interest that there are rich connections between area 7a in the parietal lobe and the regions of the inferior temporal lobe that presumably underlie visual memory (Harries & Perrett, in press). As one changes the spatial properties of the object, this in turn alters the aspects of the visual memory that are projected back to the visual buffer.

Summary and Critical Distinctions

The logic used to develop the theory of imagery hinges on the idea that perceptual mechanisms are used in imagery. Thus, I will summarize the way the system operates during perception proper before showing how it can provide accounts for the five key imagery phenomena reviewed at the outset.

Identifying objects

An object is identified by first positioning the attention window in the appropriate part of the visual buffer. Once the image of the object is enveloped by the attention window, it is sent simultaneously to the dorsal and ventral systems for further processing. The ventral system, which encodes object properties, attempts to organize perceptual units and match them to those of stored shapes. The dorsal system, which encodes spatial properties, converts retinal location to spatiotopic coordinates and encodes categorical spatial relations and motor coordinates. An object can be recognized at first glance if the match to a stored shape in the ventral system is very good. However, if the match does not definitively implicate a single object, then the identity of the closest matching object is treated as an hypothesis to be tested.

Hypothesis testing is done by accessing properties (such as parts or distinctive marking) and spatial relations between the properties of the candidate object stored in associative memory, and then positioning the attention window at the location of a sought property. The portion of the image at that location is then encoded via the ventral and dorsal systems. The subsequent output of these systems, which is sent to associative memory, may provide evidence in favour of the hypothesis or may lead to the formulation of a new hypothesis. The top-down hypothesis-testing cycle is repeated as many times as necessary until the stimulus has been identified (see Kosslyn et al., 1990, for details and computer simulations).

Imaging objects

The imagery phenomena considered earlier are explained in the following ways.

Geometric representation. The visual buffer functions to make explicit the local geometry of surfaces of objects. An image is a pattern of activation in topographically organized areas, and so portions of the representation correspond to portions of the object.

Generation. Images of single remembered shapes (that may or may not include colour, texture, or motion characteristics) are formed by activating stored visual memories in the pattern activation subsystem; this process results in a pattern of activation in the visual buffer, which is an image representation. In addition, we are led to posit four distinct types of image generation that are used when multiple parts are amalgamated or novel patterns are formed, defined by a two-by-two table: One either activates visual memories or allocates attention, and positions portions of the pattern using either categorical or coordinate representations of spatial relations. Consider first image generation when one activates visual memories of shapes, as occurs if one images a familiar scene. In this case, a description of the scene would be accessed in associative memory. This description would specify the objects and their spatial relations. Each object representation would in turn be used to activate a visual memory in the pattern activation subsystem, and the appropriate spatial relations representation would be used to position it correctly. If a coordinate spatial relations representation is encoded, a motor program is activated and its output is used to compute the location; I assume that the categorical spatial relations encoding subsystem then is used to encode the spatial relation into associative memory, where it is then used to position the image appropriately. The process of positioning the component object involves shifting the attention window to the appropriate region of the visual buffer, and forming the image at that location. (Recall that I assume, following Lowe, that the process of forming images can be adjusted to produce them in different locations in the visual buffer.) If a categorical spatial relations representation is stored, it can be used immediately to position the attention window and then form the image of the object or part in that location[4].

This sort of image generation may result in increased amounts of time for more complex objects. We expect such increases if objects are stored as separate visual memories of their constituent parts *and* each spatial relation specifies a part's location relative to a different part; hence the other part must be present before the new part's location can be computed. In principle, there is no reason why multiple parts cannot

[4]In either case, one cannot use the position information to adjust directly the location of an image in the visual buffer, without first moving the attention window; spatial relations are always specified relative to some part of an object or scene, and so the size and orientation of the object or scene will determine where the new part belongs. And the size and orientation of the object or scene is only explicit in the visual buffer, and may vary from instance to instance.

be imaged at the same time if their locations are specified relative to the body or another independent reference point. In each case, the imaged patterns may be static or moving.

The second sort of image generation is similar, except that no stored visual memories of patterns are activated. In this case, one simply picks out portions of the visual buffer to be activated. This process is done by guiding the attention window to different regions, which can be accomplished using either categorical or coordinate stored spatial relations representations.

Inspection. Objects in images are "inspected" using the exact same mechanisms as in perception. The pattern of activation in the visual buffer is surrounded by the attention window, and information is sent to the ventral and dorsal systems, as described above. These processes allow one to examine previously unconsidered shapes, colours, and textures as well as locations, orientations, and sizes. In addition, patterns of motion in the image can be encoded using the motion relations subsystem.

Recoding. The same processes are used in perception and imagery to store a new pattern in the pattern activation subsystem (i.e., enter a new visual memory) or in associative memory (i.e., enter a new structural description). I do not have a theory of how these processes operate, but the fact that the same subsystems and representations are used in the two types of processing implies that whatever mechanisms are responsible for learning in perception will also allow learning in imagery.

Maintenance. Image maintenance can be considered as a special case of image generation, with the generation mechanisms simply being used repeatedly to refresh an existing pattern of activation in the visual buffer. If a novel pattern is created, one must first encode the pattern into the pattern activation subsystem, and then activate this new representation to recreate the image. To the extent that one can "tune" the preprocessing subsystem to organize information into fewer units ("chunks") before these visual memories are created, one can hold more information in a single image.

The process of image maintenance plays a critical role in one form of "working memory" (Baddeley, 1986). In my view, there are three types of memory in the system: Short-term memory is the use of a perceptual buffer to represent information activated from long-term memory. The visual buffer is an example of such a short-term memory. Long-term memories may also be modality specific or may be amodal (i.e., in associative memory). The pattern activation subsystem is an example of a modality-specific long-term memory. Working memory is a) the combination of the information being held in the various short-term memory structures and the information that is activated in the various long-term memory structures, and b) the "control processes" that activate information in long-term memory and allow information to decay in short-term memory. That is, there is a dynamic relation between short-term and long-term memory. More information typically is activated in long-term memory than can be represented in short-term memory, and hence there often is a complex "swapping" process between the two types of structures at work, shuffling information in and out of short-term memory. Presumably the frontal lobes play a critical role in governing this swapping process, just as they do in selecting objects to be imaged. Note, however, that "loading up working memory" may consist of loading up the short-term buffers, which would not necessarily influence information stored in long-term memory.

Transformation. Finally, the revised theory leads us to expect that there are two distinct ways of transforming imaged patterns. First, if motion was an intrinsic part of a visual encoding, it can be recreated simply by activating the visual memory. Second, the spatial relations representations in the dorsal system can be altered, in part by running motor programs. This kind of operation is very flexible, and can be applied to a wide range of objects.

Conclusions

This chapter has briefly sketched some of the directions in which a theory of visual mental imagery may develop. The key innovations have been driven in part by recent discoveries about the neural underpinings of vision, and in part by developments in computer vision. At this stage in the history of the field, perhaps the most important thing about such theorizing is that it leads to innovative new empirical research. If it does, the theory is serving a crucial role. The present efforts seem to be fulfilling this requirement, and it will be of interest to see where we are led by the results of the many experiments suggested by the new ideas discussed here.

REFERENCES

Andersen, R.A. (1987). The role of the inferior parietal lobule in spatial perception and visual-motor integration. In F. Plum, V.B. Mountcastle, and S.R. Geiger (Eds.), *Handbook of Physiology: The Nervous System*

V, Higher Functions of the Brain Part 2. Bethesda, MD: American Physiological Society.

Andersen, R.A., Essick, G.K., & Siegel, R.M. (1985). Encoding of spatial location by posterior parietal neurons. *Science* **230**, 456-458.

Baddeley, A. (1986). *Working Memory.* Oxford: Oxford University Press.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review* **94**, 115-147.

Biederman, I., & Shiffrar, M.M. (1987). Sexing day-old chicks: a case study and expert systems analysis of a difficult perceptual-learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **13**, 640-645.

Cave, K.R., & Kosslyn, S.M. (1989). Varieties of size-specific visual selection. *Journal of Experimental Psychology: General* **118**, 148-164.

Cooper, L.A., & Shepard, R.N. (1973). Chronometric studies of the rotation of mental images. In W.G. Chase (Ed.), *Visual Information Processing.* New York: Academic Press.

Corbetta, M., Miezin, F.M., Dobmeyer, S., Shulman, G.L., & Petersen, S.E. (1990). Attentional modulation of neural processing of shape, colour, and velocity in humans. *Science* **248**, 1556-1559.

Cutting, J.E., & Kozlowski, L.T. (1977). Recognizing friends by their walk: Gait perception without familiarity. *Bulletin of the Psychonomic Society* **9**, 353-356.

Cutting, J.E., & Proffitt, D.R. (1981). Gait perception as an example of how we may perceive events. In R.D. Walk & H.L. Pick (Eds.), *Intersensory Perception and Sensory Integration.* New York: Plenum Press.

De Renzi, E. (1982). *Disorders of Space Exploration and Cognition.* New York: John Wiley & Sons.

de Schonen, S. & Mathivet, E. (1989). First come, first served: A scenario about the development of hemispheric specialization in face recognition during infancy. *European Bulletin of Cognitive Psychology* **1**, 2-44.

Delis, D.C., Robertson, L.C., & Efron, R. (1986). Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia* **24**, 205-214.

Farah, M.J. (1988). Is visual imagery really visual? Overlooked evidence from neuropsychology. *Psychological Review* **95**, *307-317.*

Felleman, D.J., & Van Essen, D.C. (in press). Distributed hierarchical processing in primate cerebral cortex. *Cerebral Cortex.*

Fisk, J.D., & Goodale, M.A. (1988). The effects of unilateral brain damage on visually guided reaching: hemispheric differences in the nature of the deficit. *Experimental Brain Research* **72**, 425-435.

Goldman-Rakic, P.S. (1987). Circuitry of primate prefrontal cortex and regulation of behavior by representational knowledge. In F. Plum and V. Mountcastle (Eds.), *Handbook of Physiology, Vol 5* pp 373-417. Washington, D.C.: American Physiological Society.

Goodale, M.A. (1988). Hemispheric differences in motor control. *Behavioural Brain Research* **30**, 203-214.

Gross, C.G., Desimone, R., Albright, T.D., & Schwartz, E.L. (1984). Inferior temporal cortex as a visual integration area. In F. Reinoso-Suarez & Ajmone-Marsan (Eds.), *Cortical Integration.* New York: Raven Press, pp 291-315.

Harries, M.H., & Perrett, D.I. (in press). Visual processing of faces in temporal cortex: physiological evidence for a modular organization and possible anatomical correlates. *Journal of Cognitive Neuroscience.*

Hebb, D.O. (1949). *The Organization of Behavior.* New York: Wiley.

Hellige, J.B. (1989). Endogenous and experimental determinants of cerebral laterality: What develops? *European Bulletin of Cognitive Psychology* **1**, 85-89.

Hellige, J.B. & Michimata, C. (1989). Categorization versus distance: Hemispheric differences for processing spatial information. *Memory & Cognition* **17**, 770-776.

Holmes, E.J., & Gross, C.G. (1984). Effects of temporal lesions on discrimination of stimuli differing in orientation. *Journal of Neuroscience* **4**, 3063-3068.

Hyvarinen, J. (1982). Posterior parietal lobe of the primate brain. *Physiological Reviews* **62**, 1060-1129.

Johansson, G. (1950). *Configurations in Event Perception.* Uppsala, Sweden: Almqvist & Wiksell.

Johansson, G. (1975). Visual motion perception. *Scientific American* **232**, 76-88.

Jolicoeur, P. (in press). On the role of mental rotation and feature extraction in the identification of disoriented objects: a dual-systems theory. *Mind and Language.*

Kirby, K.N., & Kosslyn, S.M. (in press). Thinking visually. *Mind and Language.*

Koenig, O., Kosslyn, S.M., & Chabris, C.F., and Gabrieli, J.D.E. (1990). Computational constraints on the acquisition of spatial knowledge in the cerebral hemispheres. University of Geneva manuscript.

Koenig, O., Reiss, L., & Kosslyn, S.M. (in press). The development of categorical and coordinate spatial relations representations. *Child Development.*

Kosslyn, S.M. (1980). *Image and Mind.* Cambridge, MA: Harvard University Press.

Kosslyn, S.M. (1987). Seeing and imagining in the cerebral hemispheres: a computational approach. *Psychological Review* **94**, 148-175.

Kosslyn, S.M., Cave, C.B., Provost, D., & Von Gierke, S. (1988). Sequential processes in image generation. *Cognitive Psychology* **20**, 319-343.

Kosslyn, S.M., Chabris, C.F., Marsolek, C.J., & Koenig, O. (in press). Categorical versus coordinate spatial representations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance.*

Kosslyn, S.M., Daly, P., McPeek, R., Alpert, N.A. & Caviness, V., (1990). Systems of dysfunction: A theory-driven analysis of a deficit in high-level visual processes. Harvard University manuscript.

Kosslyn, S.M., Flynn, R.A., Amsterdam, J.B., & Wang, G. (1990). Components of high-level vision: A cognitive neuroscience analysis and accounts of neurological syndromes. *Cognition* **34**, 203-277.

Kosslyn, S.M., Hillger, L.A., Engel, S.E., Clegg, B., & Hamilton, S.E. (1990). Attention-based and form-based imagery: Evidence from studies of transfer of practice. Harvard University manuscript.

Kosslyn, S.M., Hillger, L.A., Livingstone, M. S., & Hamilton, S.E. (1990). Does the right hemisphere monitor larger receptive fields than the left? Harvard University manuscript.

Kosslyn, S.M., Koenig, O., Barrett, A., Cave, C.B., Tang, J., & Gabrieli, J.D.E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance* **15**, 723-735.

Kosslyn, S.M. & Park, S. (1990). Incidental memory for left-right orientation in the cerebral hemispheres. Harvard University manuscript.

Kosslyn, S.M., Segar, C., Pani, J., & Hillger, L.A. (1990). When is imagery used? A diary study. *Journal of Mental Imagery* **14**, 131-152.

Larsen, A., & Bundesen, C. (1978). Size scaling in visual pattern recognition. *Journal of Experimental Psychology: Human Perception and Performance* **4**, 1 - 20.

Livingstone, M.S., & Hubel, D.H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement and depth. *Journal of Neuroscience* **7**, 3416-3468.

Lowe, D.G. (1987a). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence* **31**, 355-395.

Lowe, D.G. (1987b). The viewpoint consistency constraint. *International Journal of Computer Vision* **1**, 57-72.

Marr, D. (1982). *Vision.* San Francisco, CA: W. H. Freeman.

Maunsell, J.H.R., & Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience* **10**, 363-401.

McLeod, P., McLaughlin, C., & Nimmo-Smith, I. (1985). Information encapsulation and automaticity: evidence from the visual control of finely timed actions. In M.I. Posner & O.S.M. Marin (Eds.), *Attention and Performance XI.* Hillsdale, NJ: Erlbaum. pp 391-406

Mishkin, M., Ungerleider, L.G., & Macko, K.A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in NeuroSciences* **6**, 414-417.

Navon, D., & Norman, J. (1983). Does global precedence really depend on visual angle? *Journal of Experimental Psychology: Human Perception and Performance* **9**, 955-965.

Nickerson R.S., & Adams, M.J. (1976). Long term memory for a common object. *Cognitive Psychology* **11**, 287-307.

Paivio, A. (1971). *Imagery and Verbal Processes.* New York: Holt, Rinehart & Winston.

Perrett, D.I., Smith, P.A. J., Potter, D.D., Mistlin, A.J., Head, A.S., Milner, A.D., & Jeeves, M.A. (1985). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society London B* **223**, 293-317.

Podgorny, P., & Shepard, R.N. (1978). Functional representations common to visual perception and imagination. *Journal of Experimental Psychology: Human Perception and Performance* **4**, 21-35.

Pohl, W. (1973). Dissociation of spatial discrimination deficits following frontal and parietal lesions in monkeys. *Journal of Comparative and Physiological Psychology* **82**, 227-239.

Posner, M.I., Inhoff, A.W., Friedrich, F.J., & Cohen, A. (1987). Isolating attentional systems: A cognitive-anatomical analysis. *Psychobiology* **15**, 107 - 121.

Rock, I. (1973). *Orientation and Form.* New York: Academic Press.

Sergent, J. (1988). Face perception and the right hemisphere. *In* L. Weiskrantz (Ed.) *Thought without Language.* Oxford, UK: Oxford University Press, pp 108-131.

Sergent, J. (in press). Judgements of relative position and distance on representations of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance.*

Shepard, R.N., & Cooper, L.A. (1982). *Mental Images and their Transformations.* Cambridge, MA: MIT Press.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs* **74**, (11, Whole no. 498).

Squire, L.R. (1987). *Memory and Brain.* New York: Oxford University Press.

Tarr, M.J., & Pinker, S. (1989). Mental rotation and orientation dependence in shape recognition. *Cognitive Psychology* **21**, 233-282.

Taylor, D.C. (1969). Differential rates of cerebral maturation between sexes and between hemispheres. *Lancet* **2**, 140-148.

Treisman, A.M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology* **12**, 97-136.

Ullman, S. (1979). *The Interpretation of Visual Motion.* Cambridge, MA: MIT Press.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition* **32**, 193 - 254.

Ungerleider, L.G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, and R. J. W. Mansfield (Eds.), *Analysis of Visual Behavior.* Cambridge, MA: MIT Press.

Van Essen, D. (1985). Functional organization of primate visual cortex. In A. Peters and E. G. Jones (Eds.), *Cerebral Cortex (Vol. 3).* New York: Plenum Press.

Van Essen, D.C., Felleman, D.J., DeYoe, E.A., Olavarria, J., & Knierim, J. (1990). Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. *Quantitative Biology* **55**,

Van Kleeck, M.H. (1989). Hemispheric differences in global versus local processing of hierarchical visual stimuli by normal subjects: New data and a meta-analysis of previous studies. *Neuropsychologia* **27**, 1165-1178.

Yarbus (1967). *Eye Movements and Vision.* New York: Plenum Press.

Zipser, D. & Andersen, R.A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* **331**, 679-684.

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)

Chapter 24

Image generation and the territory of the left posterior cerebral artery*

Georg Goldenberg, Christa Artner and Ivo Podreka
Neurologische Universitätsklinik, Wien, Austria

INTRODUCTION

In her seminal paper on the neurological basis of mental imagery Farah (1984) has proposed that image generation depends on integrity of left posterior brain regions. Indeed, a considerable number of single case studies have documented that loss or impoverishment of mental visual imagery can follow lesions to the territory of the left posterior cerebral artery (PCA) (e.g., Stengel, 1948; Nielsen, 1955; Basso, Bisiach, & Luzatti, 1980; Farah, 1984; Levine, Warach, & Farah, 1985; Grossi, Orsini, & Modafferi, 1986; Deleval, DeMol, & Noterman, 1983; Pillon, Signoret, & Lhermitte, 1981; Wapner, Judd, & Gardner 1978; Davidoff & Wilson, 1985; Farah, Levine, & Calvanio, 1988a; Goldenberg, 1989), and recently, there have been several studies of event related brain potentials (Péronnet, Farah, & Gonon, 1987; Farah, Péronet, Gonon, & Giard, 1988b; Uhl et al., 1990) and of regional cerebral blood flow (Goldenberg, Podreka, Steiner, & Willmes, 1987; Goldenberg et al., 1989a; Goldenberg et al., 1989b; Goldenberg, Podreka, Steiner, Franzen, & Deecke, in press) which demonstrated an activation of the left occipital lobe in subjects performing imagery tasks. However, none of these studies unequivocally proves that it is specifically image generation which depends on the left PCA territory.

*The clinical study was supported by grant 7210M, and the SPECT study by grant 6085M from the Fonds zur Førderung der Wissenschaftlichen Forschung.

The expectation that image generation can be selectively disturbed by brain damage originates in cognitive theories of mental imagery which distinguish three main components of imagery (Kosslyn, 1983; Kosslyn, 1987; Farah, 1984): Knowledge about the visual appearance of the world stored in long-term memory; a visual buffer or working memory where that knowledge is presented and can be further processed during imagery; and an image generation process which transfers the needed information from long-term memory to the visual buffer and constructs the image. The visual long term knowledge used for imagery is also needed for recognition of perceived objects. The same visual buffer is used for retaining and processing visual images and visual percepts. Structural constraints of the visual buffer convey to the mental image "quasi-perceptual" properties, and scanning and manipulating of mental visual images follows the same rules as that of percepts. Image generation is the only component of the imagery process that does not participate in processing of perceived stimuli.

One manifestation of a selective image generation deficit should be that the patients are unable to imagine the appearance of objects which they are able to recognize from visual perception. Martha Farah concluded from her review of published cases that this dissociation was present in patients with left posterior lesions (Farah, 1984), but the comparison between imagery and recognition of objects is not a straightforward one. Loss of knowledge from long-term memory may affect only some visual attributes of objects, and partial knowledge of the visual appearance of an object can suffice for recognition. The difficulties raised by a comparison between visual imagery and recognition can be illustrated by the following observation of a patient's loss of imagery (Basso et al., 1980, p 436): The patient "was unable to describe his way home from the hospital, in spite of the fact that he could walk back and forth with no trouble, while admitting that all houses except his own looked unfamiliar". Concordance between imagery and recognition will be judged differently according to whether one assumes that the mental image to be described should depict the knowledge necessary for finding one's way or the knowledge necessary for recognising the familiarity of the route. The demonstration of a selective deficit of image generation affords a comparison between imagery and recognition of exactly the same visual attributes of objects (Farah et al., 1988a).

Another manifestation of the inability to generate mental visual images should be an inability to solve visuospatial problems with imagined stimuli which contrasts with preserved visuospatial processing of perceived stimuli. There are data which apparently conform to this dissociation, but they stem from patients with right hemisphere lesions. Morrow, Ratcliffe, & Johnston (1985) found that patients with right-brain damage performed very poorly on estimating the distances between major cities on an imagined map of the USA, although they were accurate in locating the cities on an empty map and in estimating perceived distances on an actual map. Goldenberg (1989) found that right-brain damaged patients were impaired in both, visuospatial tasks with imagined and with perceived stimuli, but concluded from an analysis of the correlations between test results that the visuospatial imagery tasks shared a component which distinguished them from the perceptual tasks. In both of these studies demands on image generation cannot be differentiated from demands on visual working memory. In the imagery tasks patients had to generate images and had to retain them in working memory for visuospatial processing, whereas in the control tasks the stimuli were present all of the time and there was no need to keep them in working memory at all.

Single photon emission computer tomography (SPECT) studies of rCBF showed left inferior occipital activation quite consistently in subjects who formed detailed images of objects, but not in subjects who imagined capital letters in order to count how many corners they have (Goldenberg et al., 1987; Goldenberg et al., 1989a; Goldenberg et al., 1990). Left inferior occipital CBF was higher in subjects who imagined faces that in subjects who imagined a single colour filling the whole visual field (Goldenberg et al., 1989b). Possibly, left inferior occipital activation depends on how rich on details the mental image is: A skeletal and schematic image of capital letters would be sufficient for the counting of their corners, and there is no specification of details in the image of a single colour filling the whole visual field. By contrast, images of faces or objects may start with a skeletal sketch but are then embellished by filling in many details. A dependence of regional activation on the number of details would be compatible with the assumption that the region is responsible for image generation, because the need to collect from long-term memory many details and to construct a mental image out of them should place high demands on the image generation process. The interpretation is not compelling, since the same dependence between left inferior occipital activation and richness on details could be expected if the region were itself the site of storage of long-term memory about details. In this case, activation would be linked to the retrieval or, respectively, activation of information from long-term memory but not to the additional step of image generation.

We addressed the question whether the left PCA territory is responsible for image generation by a clinical study of patients with

PCA lesions (Goldenberg & Artner, in press), and we supplemented the study by giving the same imagery tasks to healthy subjects and assessing rCBF. The studies employed the paradigm of low- and high-imagery sentences (Eddy & Glass, 1981). High-imagery sentences are sentences which make predicates about the visual appearance of objects. When asked to verify their factual correctness, subjects introspectively report that they generate a mental visual image and derive the answer from an inspection of that image. Patients with an image generation deficit should be unable to distinguish correct from incorrect high-imagery sentences because they cannot generate the mental images necessary for verification. By contrast they should be able to distinguish correct from incorrect pictorial presentations of the same alternative predicates, because picture verification does not call for image generation. If the deficit concerns the long-term memory information about the visual appearance of objects patients should fail with any version of the question.

Clinical study

25 low-imagery sentences, 25 high-imagery sentences concerning the shape and 25 concerning their colour were selected in a pilot study by obtaining judgements about the appearance and necessity of images for sentence verification from 23 volunteers. The sentences were in dual-choice format with one correct and one incorrect version. As far as possible, sentences of different classes concerned the same objects. For example, a low-imagery sentence was "Squirrels are mammals / lay eggs", a high-imagery sentence concerning shape "The eyes of the squirrel are slit / round", and one concerning colour "The brown of squirrels is rather yellowish / rather reddish". To all high-imagery sentences pictorial versions were prepared which showed two drawings of the same objects that differed in the detail asked for by the sentence, or, respectively, in their colour. To control for perceptual deficits which might lead to difficulties with picture verification two dual-choice perceptual discrimination tests were constructed. In one patients had to decide whether the curvatures of two triangles with convex boundaries were the same or not, and in the other whether two shades of a colour were the same or not.

Patients with isolated unilateral infarctions of either the left or the right PCA territory were selected from the records of our computer tomography department and invited to participate. Patients with history or signs of brain damage in other locations were excluded. There remained 19 patients with left and 15 with right PCA damage. In addition, 21 healthy controls were given the same tests. The groups did not differ in age, distribution of gender, and educational level.

The order of testing was as follows: Clinical history, visual acuity and visual fields were assessed. Reading of letters, of real and of nonsense words, and naming of basic colours was tested. Screening for colour blindness was performed with pseudo-isochromatic plates. In one deuteranopic patient no colour item was evaluated. Then the two perceptual discrimination tests were administered. The 75 sentences were presented in a fixed random order printed on cards with the two predicates below the common part. In addition, they were read aloud by the examiner. The patient was asked to select the correct predicate. Then the pictorial versions of the shape sentences were presented one after another. Patients were asked first to decide which of the two versions was the correct one, then to name the object and to designate the crucial difference. If they had not appreciated the crucial detail it was shown and they were asked whether a consideration of that detail would change their decision. For the pictorial versions of colour questions, patients were instructed that the only difference concerned the colour and they were only asked to name the object after their decision. Since all of the patients had successfully solved the pseudo-isochromatic plates, discrimination of the colours used in these questions did not present a problem.

A significant proportion of errors in reading and naming occurred only in the left PCA group. There were a few items of the visual versions which caused difficulties of identification in all groups and if this was noted, a semantic cue was given and answers were scored only after correct identification. An analysis of the naming errors committed by left-PCA patients did not show an enhanced number of misidentification errors.

With both perceptual discrimination tasks, the ANOVA showed significant group effects. On post-hoc testing right-PCA patient scored lower than controls whereas left-PCA patients did not significantly differ from any of the other groups.

There was no significant group effect in the ANOVA on low-imagery sentences. With shape and colour imagery sentences the ANOVA was significant, and on post-hoc testing left-PCA patients scored lower than controls. A two-factorial ANOVA, however, failed to prove the reliability of this apparent dissociation between low- and high-imagery sentences.

There was a highly significant ($P<0.00005$) group effect for the first choice of visual shape questions. On post-hoc left-PCA patients differed from controls, but right-PCA patients did not differ from any group. The number of crucial differences correctly perceived was also different across groups with both PCA-groups recognizing less crucial details than controls. When the second choices after the demonstration of crucial differences were considered, the significance of the ANOVA remained

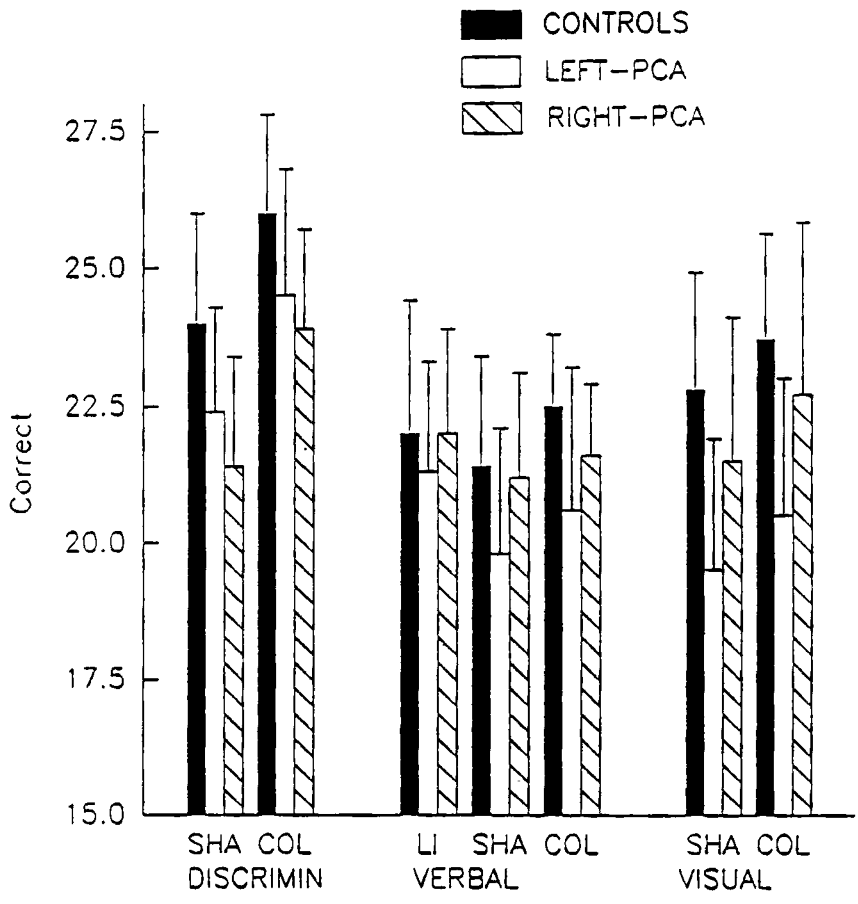

Figure 1. Scores on experimental tests (means and standard deviations) for Shape (SHA), Colour (COL), Low Imagery (LI), Discrimination (DISCRIMIN), Verbal questions (VERBAL), and Visual questions (VISUAL).

approximately the same as with the first choices, but now left-PCA patients differed from both right-PCA patients and controls which did not differ from each other. If not stated otherwise, all following analyses of visual shape questions will refer to the second choices. The results with visual colour questions were essentially the same: The ANOVA was highly significant, and left-PCA patients scored lower than both other groups which did not significantly differ from each other.

The hypothesis of a specific impairment of image generation in left-PCA patients would have predicted greater impairment on verbal than on visual questions and consequently a greater improvement from the verbal to the visual questions. For the group results, this was not the case. In fact, only controls had a significant improvement from verbal to visual versions of shape or colour questions. There were non-significant improvements in right-PCA patients and no improvement at all in left-PCA patients. We also looked whether there were single left-PCA patients with an outstanding difference in favour of the visual versions, but in no case did the difference exceed the range of controls.

There are three major cortical branches of the posterior cerebral artery, the temporal, temporo-occipital, and occipital. The PCA also supplies the thalamus, but isolated thalamic lesions had not been considered for the study. The temporal branch and the thalamus were lesioned in only one left-PCA patient in whom the temporo-occipital branch was affected as well. The CT-scans of all patients were classified as showing infarctions of the temporo-occipital or the occipital branches.

In right-PCA patients there were no significant variations of test scores related to this anatomical distinction. The anatomically defined subgroups of left-PCA patients did not significantly differ with respect to shape and colour discrimination, low imagery sentences and verbal and visual versions of shape questions. The low minimum scores on all tests of linguistic abilities stemmed from the one patient with the temporal and thalamic lesions, but there were no significant group differences on reading or naming. By contrast, there were significant group effects for both, the verbal and the visual version of colour questions, and on post-hoc testing patients with temporo-occipital lesions scored lower than those with occipital lesions on both of them. The scores of the patients with occipital lesions on visual and verbal colour questions were nearly identical to those of right-PCA patients and did not significantly differ from controls. A two-factorial ANOVA confirmed the reliability of the dissociation between visual shape and colour questions across the subgroups of left-PCA patients, but for the three kinds of verbal questions the group by task interaction failed to reach significance. There was a significant group by task interaction when the scores on low-imagery, shape and colour questions were compared between controls and patients with left temporo-occipital lesions. The patients with left temporo-occipital lesions were nearly as good as controls in responding to low imagery questions. The difference on shape imagery questions was marginally significant and that on colour imagery sentences highly significant.

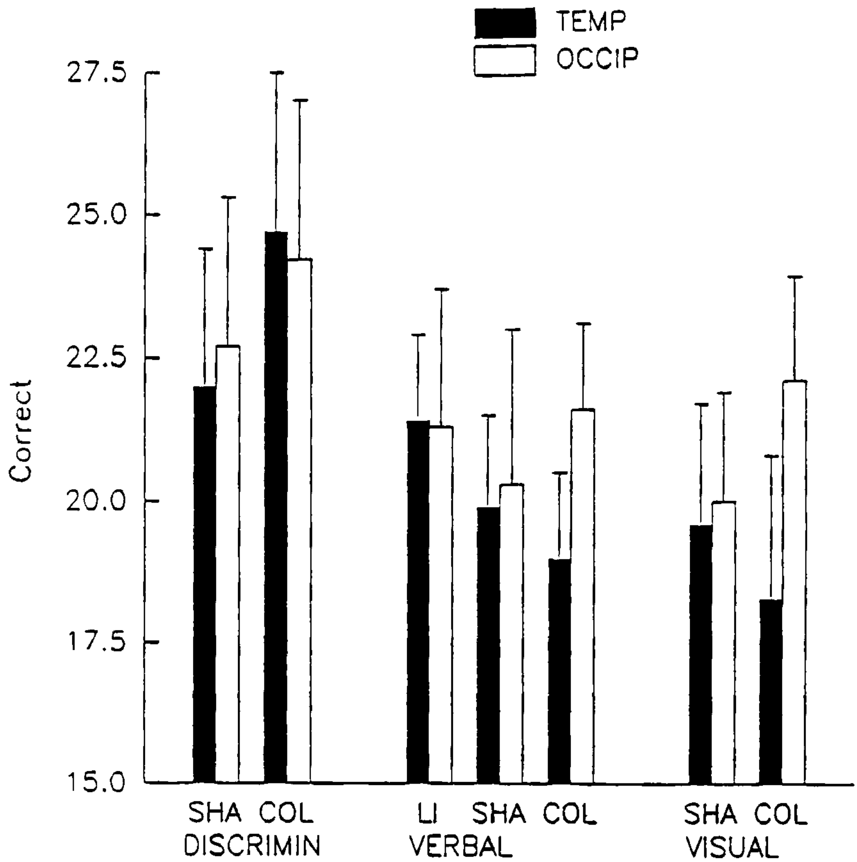

Figure 2. Scores on experimental tests (means and standard deviations) in relation to localization of left-PCA lesions. Temporo-occipital branch of PCA (TEMP) Occipital branch of PCA (OCCIP). Other abbreviations as in Figure 1.

A final analysis concerned the relationship of test results to the scores on naming and reading. They could be analyzed only in left-PCA patients because of the other groups' ceiling effect on naming and reading. There were significant correlations between these measures of linguistic abilities and errors on the visual shape and colour questions. There were eight left-PCA patients who had not committed a single naming error. Compared to controls these patients scored significantly more poorly on visual shape and colour questions. This makes it very unlikely that errors on visual shape and colour questions were a sequel of language impairment, or otherwise it had to be a language impairment which manifests itself in a non-verbal task stronger than in naming and responding to verbal questions. The association between knowledge about the visual appearance of objects and measures of language abilities need not necessarily betray any functional link at all, but may stem from anatomical contiguity between their respective cerebral correlates (Shallice, 1988).

SPECT study

The SPECT study examined rCBF in healthy volunteers who verified high and low imagery sentences. Subjects wore earphones connected to a tape recorder and a blindfold, and they were asked to close the eyes behind the blindfold. An intravenous line was laid to a cubital vein before the experiment, and in the other hand subjects held a small lamp which could be flashed by an easy touch. 25 sentences were read from the tape, 13 with a correct and 12 with an incorrect predicate, and subjects were asked to flash the lamp if the predicate was wrong. The isotope HMPAO was injected after the fifth sentence. This isotope gets trapped in brain tissue within about one minute and its distribution is proportional to rCBF during that time. There were two experimental groups and each was examined twice within one week. In the imagery condition, one group verified shape imagery sentences, the other colour imagery sentences, and in the control condition the same low-imagery sentences were given to both groups. The order of conditions was balanced within groups.

The shape- and colour imagery sentences were the same as in the clinical study, but the low-imagery sentences could not be used. Their criterion for validation was the necessity rather than the possibility to form visual images for sentence verification and they had deliberately been selected as referring to the same objects as the high-imagery sentences. Hence, it was possible to visually imagine the objects they referred to, and we found out in a pilot study that subjects in a SPECT study tended to do so, particularly when they had received high imagery sentences before. Therefore, 25 low-imagery sentences which referred to low-imagery referents were selected from a set used in a previous SPECT study (Goldenberg et al., 1989a).

The rate of errors did not differ between the three kinds of sentences. There were no significant differences at all between the two imagery conditions. This finding need not be of particular relevance. Presumably, subjects in the colour condition imagined not only the colour but also the shape of objects, and subjects in the shape condition may have had coloured mental images. Compared to low imagery sentences, high-imagery sentences led to significant decreases of regional flow in the right anterior frontal region and the left inferior temporal region and to a significant increase in the left inferior occipital region.

DISCUSSION

The SPECT study replicated the main result of a previous SPECT study using high- and low imagery sentences (Goldenberg et al., 1989a) and thus confirmed the reliability of left inferior occipital activation in the verification of high visual imagery sentences. It does not, however, contribute significantly to a componential analysis of the role played by that region in imagery. The clinical study yielded results which argue against the assumption that the left PCA territory is the site of image generation. This hypothesis would have predicted a more severe deficit with verbal than with visual versions of shape and colour questions, but the reverse was the case: Relative to controls and to right-PCA patients left-PCA patients were worse in selecting the correct pictorial versions of high-imagery sentences than in verifying the verbal sentences. Within the left-PCA territory there was an anatomically defined dissociation between performance on shape and colour questions. In order to accommodate this finding with a role of that region in image generation, one had to postulate that there are at least two anatomically distinct image generation modules, one for shape and one for colour. Of course, there is no formal obstacle against such an expansion of the theory. Once started, the division of image generation processes can easily be pushed farther, for example by postulating distinct generation processes for local patterns and for the spatial position of the components of an image (Kosslyn, 1983; Kosslyn, 1987). In the extreme, there can be as many image generation modules as there are kinds of knowledge about the visual appearance of objects. On a phenomenological level, a mental image appears as a unity which displays shape, size, position and colour simultaneously. It seems to us, that the main motive for the postulate of a distinct image generation process is to bring about this apparent unity from distinct pieces of knowledge. If the image generation process loses this unifying role, one must ask whether it is a necessary element of the functional architecture of imagery at all. The failure to find convincing cases of a pure image generation deficit or to find a unique location for image generation in the brain (Sergent, 1990; Goldenberg, 1989) may be due to the fact that the functional architecture of imagery does not contain an image generation process which is distinct from the recall of pieces of visual knowledge. This would imply that the mental image is not a unity but a collection of different pieces of information, and that visual working memory is not a structure (Logie 1989) but a limit on the capacity to simultaneously recall or, respectively, attend to several pieces of information. It further means that internal "perception" of mental images is equivalent to the simultaneous, and "scanning" to the successive recall of pieces of information about the visual appearance of objects. Several experiments on image scanning (Kosslyn, 1983; Denis, 1989; Finke, 1989) apparently contradict this implication, but may be amenable to new interpretations. The very notion of "perception" and "scanning" of mental images calls in the tricky question whether a mental image exists if one does not look at it.

Consider the possibility that imagery is equivalent to activation of visual knowledge within long-term memory and that lesions of left occipital regions erase parts of that knowledge. Verification of high-imagery sentences is an essentially verbal task and therefore requires an interaction between verbal processing, semantics and visual knowledge. This interaction is a source of errors even in normal controls, and if this source of errors is alleviated by a visual presentation of the questions, then the error rate is reduced. If visual knowledge is deficient, some of the sentences can still be correctly verified by drawing conclusions and inferences from other domains of knowledge but the visual presentation does not lead to any improvement and hence the difference from controls increases. This is the pattern of results observed in the patients with left PCA damage. The relationship of scores on shape and colour questions to parameters of linguistic abilities can be accounted for by the assumption that there is no sharp border between the anatomical locations of verbal and visual knowledge, but that their cerebral substrates correspond to adjacent positions in an extended neural network. Large lesions centred in the location of visual knowledge will encroach upon verbal knowledge and thus lead to problems with tasks taxing semantic aspects of language. Anatomical contiguity could thus be the source of the correlations between linguistic tasks and scores on visual shape and colour questions.

REFERENCES

Basso, A., Bisiach, E., & Luzatti, C. (1980). Loss of mental imagery: A case study. *Neuropsychologia* **18**, 435-442.

Davidoff, J., & Wilson, B. (1985). A case of visual agnosia showing a disorder of pre- semantic visual classification. *Cortex* **21**, 121-134.

Deleval, J., De Mol, J., & Noterman, J. (1983). La perte des images souvenirs. *Acta Neurologica Belgica* **83**, 61-79.

Denis, M. (1989). *Image et cognition.* Presses Universitaire de France, Paris, 284.

Eddy, J.K., & Glass, A.L. (1981). Reading and listening to high and low imagery sentences. *Journal of Verbal Learning and Verbal Behaviour* **20**, 333-345.

Farah, M.J. (1984). The neurological basis of mental imagery: A componential analysis. *Cognition* **18**, 245-272.

Farah, M.J., Levine, D.N., & Calvanio, R. (1988a). A case study of mental imagery deficit. *Brain and Cognition* **8**, 147-164.

Farah, M.J., Peronnet, F., Gonon, M.A., & Giard, M.H. (1988b). Electrophysiological evidence for a shared representational medium for visual images and percepts. *Journal of Experimental Psychology: General* **17**, 248-257.

Finke, R.A. (1989). *Principles of mental imagery.* MIT Press, Cambridge MA, London, 1-179.

Goldenberg, G. (1989). The ability of patients with brain damage to generate mental visual images. *Brain* **112**, 305-325.

Goldenberg, G., & Artner, C. (in press). Visual imagery and knowledge about the visual appearance of objects in patients with posterior cerebral artery lesions. *Brain and Cognition.*

Goldenberg, G., Podreka, I., Steiner, M., & Willmes, K. (1987). Patterns of regional cerebral blood flow related to meaningfulness and imaginability of words - An emission computer tomography study. *Neuropsychologia* **25**, 473-486.

Goldenberg, G., Podreka, I., Steiner, M., Willmes, K., Suess, E., & Deecke, L. (1989a). Regional cerebral blood flow patterns in visual imagery. *Neuropsychologia* **27**, 641-664.

Goldenberg, G., Podreka, I., Uhl, F., Steiner, M., Willmes, K., & Deecke, L. (1989b). Cerebral correlates of imagining colours, faces and a map. I. SPECT of regional cerebral blood flow. *Neuropsychologia* **27**, 1315-1328.

Goldenberg, G., Podreka, I., Steiner, M., Franzen, P., & Deecke, L. (in press). Contributions of occipital and temporal brain regions to visual and acoustic imagery. *Neuropsychologia.*

Grossi, D., Orsini, A., & Modafferi, A. (1986). Visuoimaginal constructional apraxia: On a case of selective deficit of imagery. *Brain and Cognition* **5**, 255-267.

Kosslyn, S.M. (1983). *Ghosts in the mind's machine: Creating and using images in the brain.* W.W.Norton and Company: New York, London.

Kosslyn, S.M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. *Psychological Review* **94**, 148-175.

Levine, D.N., Warach, J., & Farah, M. (1985). Two visual systems in mental imagery. Dissociation of "what" and "where" in imagery disorders due to bilateral posterior cerebral lesions. *Neurology* **35**, 1010-1018.

Logie, R.H. (1989). Characteristics of visual short-term memory. *European Journal of Cognitive Psychology* **1**, 275-284.

Morrow, L., Ratcliff, G., & Johnston, C.S. (1985). Externalising spatial knowledge in patients with right hemisphere lesions. *Cognitive Neuropsychology* **2**, 265-273.

Nielsen, J.M. (1955). Occipital lobes, dreams and psychosis. *Journal of Nervous and Mental Disease* **121**, 50-52.

Péeronnet, F., Farah, M.J., & Gonon, M.A. (1987). Evidence for shared structures between mental imagery and perception. *In* M. Denis, J. Engelkamp & J.T.E. Richardson (eds), *Cognitive and neuropsychological approaches to mental imagery,* pp 357-362. Martinus Nijhoff, Dordrecht.

Pillon, B., Signoret, J.L. & Lhermitte, F. (1981). Agnosie visuelle associative. Role de l'hémisphère gauche dans la perception visuelle. *Revue Neurologique* **137**, 831-842.

Sergent, J. (1990). The neuropsychology of visual image generation. *Brain and Cognition* **13**, 98-129.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge University Press, Cambridge, New York.

Stengel, E. (1948). The syndrome of visual alexia with color agnosia. *Journal of Mental Science* **94**, 46-58.

Uhl, F., Goldenberg, G., Lang, W., Lindinger, G., Steiner, M., & Deecke, L. (1990). Imagining faces, colours and a spatial map II. Negative cortical DC - potentials. *Neuropsychologia* **28**,81-93.

Wapner, W., Judd, T., & Gardner, H. (1978). Visual agnosia in an artist. *Cortex* **14**, 343-364.

This Page Intentionally Left Blank

Mental Images in Human Cognition
Robert H. Logie and Michel Denis (eds.)
Elsevier Science Publishers B.V., 1991

Chapter 25

Effortful and automatic activation of imagery: Evidence from right brain damage*

W.A. van Loon-Vervoorn, A. Elzinga-Plomp and J.H. Hennink
Rijksuniversiteit Utrecht, The Netherlands

INTRODUCTION

In this paper we explore the nature of the relationship between age of acquisition of words and word imageability, as it manifests itself in word association experiments. The focus will be on the effects of age of acquisition on differential imagery processes in meaning activation and word selection. We studied these issues both in normal subjects, and in neuropsychological patients with damage in the right hemisphere. In former presentations to this workshop it was argued that the adult lexicon consists of two major components, a basic lexicon and a higher order lexicon. It was also argued that this division is the result of differences in the ways words are learned in childhood (Elbers & Van Loon-Vervoorn, 1990; Van Loon-Vervoorn & Van der Ham-van Koppen, 1988; Van Loon-Vervoorn, 1988). Early acquired words generally refer to fundamental aspects of human experience, that is, they delineate real world correlational structures encountered in perceptual and motor interactions with the environment. For this reason, it can be assumed that the meaning representations of these words have a strong sensorimotor component. These words are particularly useful for communicating about the 'here and now'. However, since this basic lexicon must be fundamental for the extension of lexical knowledge at a later age, the words that belong to this lexicon may also be thought to provide the initial verbal contexts for the acquisition of other non-

*We wish to thank Iet van Bekkum, Loekie Elbers and Sjoerd Wiegersma for their critical discussion of earlier drafts of this paper.

sensorimotor meanings. For this reason we call them basic words. Non-sensorimotor meanings can be acquired through language use itself, i.e. without direct interaction with the external world. It can be assumed, that the meaning representations of these later acquired words have a strong verbal component.

In word association the distinction between a basic and a higher order lexicon is revealed, in that the relationship between stimulus and response is dependent on the age of acquisition of the stimulus word (Van Loon-Vervoorn, 1988; 1989). In early acquired words common responses specify a sensorimotor aspect of the meaning of the stimulus word, i.e the relationship between stimulus and response reflects how objects and events are perceived (horse-meadow; lemon-sour), how they function (bird-fly; stove-burn) or in what way they are acted upon (knife-cut; shoe-lace). We will call this kind of relationship feature specification. In contrast to this, late acquired stimulus words manifest relationships to response words that are verbal i.e. based on superordination (beetle-insect; willow-tree) or synonymy (robe-coat; doctor-physician). These relationships will be called definitions. In some cases however, in late acquired stimulus words the relationship between stimulus and response seems to be based on feature specification and so to resemble that of basic words, as in 'safe-money' and 'desert-sand'. These responses predominantly occur when no superordinate or synonym seems to be available for a late acquired stimulus word.

So, we may assume that in word association feature specification is the normal base for word selection in early acquired stimulus words, as is definition in late acquired words. When however, in late acquired words no definition becomes available from the associative network structure, feature specification occurs that resembles that of early acquired stimulus words. As was already said, in early acquired words common responses specify a sensorimotor aspect of the meaning of the stimulus word, whereas in late acquired words the relationship between a stimulus and its common response is verbal. Now the problem is to show whether feature specification in late acquired words is different from feature specification in early acquired words.

A first hint for this difference came from a study on word association with elderly subjects suffering from dementia (Van Loon-Vervoorn & Willemsen, 1989). In these subjects the response selection to early acquired stimulus words was relatively intact as was word selection to late acquired words with a valence for common responses of a 'defining' nature (robe-coat; willow-tree). However, the response selection to those late acquired stimulus words which, according to the word association norms, tend to evoke feature specifying common responses (safe-money; desert-sand) was disturbed. So, in dementia, feature specification in early acquired words seems to be intact, whereas in late acquired words it seems to be disturbed. How can this result be explained?

It is generally thought that in dementia effortful, attentive information processing is more disturbed than automatic forms of information processing. Therefore, we may suppose that word selection based on feature specification proceeds more automatically in early acquired words than in late acquired stimulus words. If this hypothesis is correct, then in normal subjects longer latencies between stimulus and response should occur in late acquired words with a valence for feature specification than in early acquired words. The testing of this hypothesis was the first aim of the research reported in this paper.

The next question is what kind of effort is required in the case of word selection in response to late acquired words with a valence for feature specification. To answer this question one should take a closer look at the process of word selection in word association. In early acquired stimulus words word selection mostly results in a feature specifying response (bath-water; bed-sleep) and in late acquired words in a defining response (dagger-knife; mango-fruit). Both kinds of responses, that are generally acquired early, becomes automatically available from the associative network structure of memory. When in late acquired words no such an early acquired defining response is available as a node in this network organisation, normal subjects tend to produce a feature specifying, also early acquired response. For this response a delay is predicted, if the presupposition is that word selection is attentive rather than automatic. Since the relationship between stimulus and response resembles that in early acquired words, it seems reasonable to assume that this selection is based on imagery processes. Because of the predicted delay of these responses these imagery processes are probably actively induced. If so, patients with right brain damage should encounter difficulties in just those stimulus words, for in these patients perceptual processes, and probably also imagery processes that rely on these, are disturbed[1]. On the other hand, in early acquired words these patients' word selection is expected to be normal for at least two reasons. First, the meanings of early acquired words are probably bilaterally represented. Second, if in these words imagery plays a role in

[1]Perceptual and imagery representations are to some extent based on isomorphic processes. For example the imagery values of the names of complete objects (bed) are higher then the names of parts of objects (beard) or substances (milk). In the latter cases the referential meaning incorporated a relationship, either between part and whole, as in beard-face) or between substance and container, as in milk-glass (Van Loon-Vervoorn, 1989).

meaning activation and word selection, these images will probably become automatically available.

EXPERIMENT 1. Response latencies in word association

Materials and Subjects

Twenty nouns were selected from Dutch word association norms (De Groot, 1980; De Groot & De Bil, 1987; Van Loon-Vervoorn & Van Bekkum, in press) in the following three word groups. Early acquired words with a feature specifying primary response[2], late acquired words with a defining and late acquired words with a feature specifying primary response. The primary response is the response word with the highest association frequency in the word association norms. The word groups were matched as to length and association frequency of the primary response.

In this experiment 30 subjects participated, 18 female and 12 male. All subjects were students of psychology at the University of Utrecht. The mean age of these subjects was 24.8 years with a standard deviation of 2.3.

Procedure

The experiment was run on an Apple-plus computer. The stimulus words were presented in uppercase letters, green on a gray monitor. Each presentation was preceded by an asterisk. The response times of the subject's oral responses were registered by a microphone that activated a voice-key. The responses were simultaneously recorded on tape. The subjects, who participated individually, were asked to pronounce, as rapidly as possible, the first word they thought of when reading the stimulus word. It was stressed that the response should consist of one word only. The stimulus word remained on the screen until the voice-key registered a sound. Response times were measured from the onset of the stimulus word to the onset of the response. The maximum presentation duration of a stimulus word was five seconds. Whenever this duration expired a blank was registered and a new trial was started. When the voice-key failed to respond to the subject's actual response or responded to an irrelevant noise the experimenter made note

[2]Early acquired words that evoke defining primary responses rarely occur in the consulted word association norms. These norms contain more than 2000 stimulus words.

of this. The response times of the latter were excluded from subsequent data analysis. They constituted less than 0.5% of all responses. Following 10 practice stimuli the 60 test items were presented in a random order (a new order for every subject).

Results

For each stimulus word the median response time of the primary response was calculated. The means of these medians for the three word groups are presented in Table 1.

Table 1 Mean response time of the primary response in early and late acquired words (D: defining, FS: feature specifying primary response).

	early	late (D)	late (FS)
Mean response time primary response	1147.7	1246.0	1472.8

In a one-way analysis of variance a significant effect of word group was found [$F(2,57)=17.09$, $p<.001$]. A post-hoc Scheffé analysis revealed only a significant difference between the two late acquired word groups, supporting our hypothesis that the selection of a feature specifying primary response to a late acquired stimulus word is more effortful than either the selection of a feature specifying response to an early acquired stimulus word or a defining primary response to a late acquired stimulus word. The difference of 226.8 ms between the two late acquired word groups is comparable to the difference of 189 ms reported in Van Loon-Vervoorn (1989) that was based on the word association norms with response times of De Groot and De Bil (1987).

EXPERIMENT 2. Word association in right brain damaged patients

Materials

Two word lists were used. List A contained early and late acquired words that were either high or low on imageability. The four word groups consisted of ten stimulus words (nouns) each. These 40 words, intermixed with 20 early acquired filler words, were also used in our study with elderly subjects with and without dementia (Van Loon-Vervoorn & Willemsen, 1989).

List A was used to study the general effects of age of acquisition and imageability on word association. List B consisted of 40 late acquired words, 20 high and 20 low on imageability. In each group ten words have, according to the word association norms, a defining primary response and 10 words a feature specifying response. These 40 words were intermixed with 20 early acquired words. The words of both lists were presented in a random order. This order was rotated for each subject.

Procedure and Subjects

Twenty patients (6 female and 14 male) with brain damage confined to the right hemisphere participated in this experiment. In all cases the damage was caused by a cerebral vascular accident, which took place at least six months before testing. These patients were recruited from Revalidatiecentrum De Hoogstraat in Utrecht. Their mean age was 59.6 years with a range of 41 to 75 years. The word association norms were used as control data. These norms are collected from groups of normal young subjects, usually psychology students. In an earlier study the response hierarchies of the word association norms and those of a control group that was matched as to age and social-educational background with aphasic, alcoholic (Korsakov) and epileptic experimental groups were compared (Van Loon-Vervoorn, 1989). The differences between the younger subjects (norms) and the older subjects (control group) did not exceed the normal variation that is found when word associations of two samples of students are compared. For example, 70% of the primary responses in the norms is also primary response in the age control group and of the remaining 30% primary responses 23% is secondary in this group. The mean association frequency of these primary responses in the norms is 31.6. In the age control group the association frequency of this words is 27.7.

The same procedure was used as in our former study with aphasia (Van Loon-Vervoorn & Van der Ham-van Koppen, 1988). The patients were individually tested in two sessions that were separated by at least one week. One half of the patients started with list A, the other half with list B.

Results

For the general list (A) the association frequency of the primary response is separately presented for the patients (RBD) and the word association norms for each word group in Table 2. The primary response is a measure for the commonality of a response.

Table 2 The mean association frequency of the primary response (percentages) for right brain damaged patients and word association norms (E-early, L-late, C-high imagery, A-low imagery).[3]

List A	EC	EA	LC	LA
association frequency (norms)	27.5	30.6	48.8	22.0
association frequency (RBD)	27.5	25.0	37.5	28.5

In normal subjects (norms) a two-way analysis of variance revealed that the association frequency of the primary response was dependent on imageability [$F(1,36)=5.30$, $p<.05$]. In this analysis also the interaction between age of acquisition and imageability was significant [$F(1,36)=8.83$, $p<.01$]. This pattern of results was also found in brain damaged epileptics with no focal problems in language or memory (Hennink, 1989). As can be seen in Table 2 the main effect of imageability can be accounted for by this interaction; a rather large difference occurred between LC and LA. In the right brain damaged group neither the main effect of imageability [$F(1,36)=2.11$] nor its interaction effect with age of acquisition was significant [$F(1,36)=0.41$]. Particularly the association frequency of the late acquired imagery words was reduced as compared to the norms. However not all words in this group were equally affected. In defining word pairs as (dagger-knife) the association frequency of the primary response equalled that of the norms (49%), whereas in feature specifying pairs (e.g. desert-sand) this frequency was quite low (26%). List A was not, as was list B, designed to explore the

[3]High imageability words usually are concrete and low imageability words abstract. For this reason the abbreviations C and A were used.

relationship between stimulus and response and its effect on word selection. Therefore, we now turn to the results of list B.

List B consisted of 40 late acquired stimulus words, ten high imagery and ten low imagery with a defining primary response (LCD and LAD); and ten high imagery and ten low imagery with a valence for feature specifying (LCFS and LAFS), as well as 20 early acquired, imagery words (EC). In the two defining word groups as well as in the two feature specifying groups the means of the association frequency of the primary responses were matched. In Table 3 the mean association frequency of the primary response of these four word groups is presented for right brain damaged patients and according to the word association norms.

Table 3 The mean association frequency of the primary response for the word association norms, the association frequency of this response in right brain damaged patients and the number of multiple word responses for these patients

List B	EC	LCD	LAD	LCFS	LAFS
association frequency (norms)	35.7	41.6	38.4	27.9	25.9
association frequency (RBD)	30.0	41.5	43.0	11.5	7.5
number of multiple words responses	2.5	3.4	2.9	4.9	4.7

As was expected, neither in the early acquired word group nor in the defining word groups was a difference found between the mean association frequencies of the patients and according to the norms (t-tests for dependent samples). However, as predicted, in the feature specifying groups the means of both groups differed significantly (LCFS: $t=5.80$, df 18 $p<.001$; LAFS: $t=4.96$, $df=18$, $p<.001$).

In the early acquired stimulus words as well as in the defining stimulus words 80% of the primary responses of the right brain damaged patients coincided with that of the norms. In the feature specifying stimulus words this percentage was only 15% (3 words). For the other words 80% (14 out of 17) of the primary responses were definitions of the stimulus word, and these have a relatively low mean association frequency (4.9) according to the norms. In these word groups almost twice as many multiple word responses occurred, compared to the other word groups (Table 3). In a one-way analysis of variance, comparing the three word groups (early, late defining and late feature specifying) on multiple word responses, the overall difference between the word groups was significant [$F(2,56)=9.86$, $p<.001$]. With a Scheffé analysis only the difference between the two late acquired word groups was significant [$F(2,57)=10.59$, $p<.001$]. Multiple word responses were in most cases attempts to define the stimulus word, as for example in "luxury-when you have a lot of money". Probably, when no lexical defining alternative, consisting of a single word, was available to the patient, as seems to have been the case in late acquired feature specifying words, definition was achieved by description.

From these data we may conclude that in right brain damaged patients a relatively normal pattern of primary responses appear in the case of early acquired words as well as in the case of late acquired words that have a valence for defining primary responses. However, in the case of late acquired words with a valence for feature specification the primary responses of the patients deviate from those of normal subjects in that many more definitions are given to the stimulus words as might be expected from the word association norms.

To explore whether this tendency for defining is a consequence of disturbed lexical access rather than of a response set, a third list was presented to the patients in a separate session, at least one week after the second presentation (list A or B). A response set is a tendency to associate according to a fixed strategy, for example predominantly antonyms or clang associations are produced. Such a strategy may be induced by the stimulus list. In this third list early and late acquired feature specifying stimulus words were presented in four blocks of ten words. The sequence was: early-late-early-late. Should a response set be responsible for the above mentioned results then a normal response pattern should be found in both early and late acquired blocks, since a set for feature specification would be induced in the first block of early acquired words. If access in late acquired lexical knowledge is disturbed then it would seem reasonable to expect that the early blocks elicit a normal response pattern. On the other hand the late acquired blocks are expected to deviate from normality in that these words elicit more defining responses as expected from the norms.

The mean association frequencies of the primary response according to the norms and the mean association frequency of this response in right brain damaged patients (percentages) are presented in Table 4.

Table 4. The mean association frequency of the primary response of the word association norms and the mean frequency of this response in right brain damaged patients for early and late acquired words of the block list. In parentheses the number of primary responses that are the same in both groups.

List C	early1	late1	early2	late2
association frequency (norms)	30.7	26.6	28.2	27.0
association frequency (RBD)	31.0 (7)	13.5 (2)	18.8 (3)	12.0 (1)

For the first block of early acquired words no difference in association frequency was present [t(18)=.06], and the number of common primary responses is 7 (out of 10). In the two late acquired groups the association frequency of the primary response of the norms is significantly reduced in the patients [late-1: t(18)=3.82, p<.01; late-2 t(18)=3.87, p<.01), as is the number of common primary responses (respectively 2 and 1). In the second early acquired group the difference between the patients and the norms is nearly significant at the 5% level [t(18)=2.09, p=.06) with only 3 common primary responses. The number of defining primary responses in the right brain damaged group is 0 in early-1, 7 in late-1, 4 in early-2 and 8 in late-2. These results suggest a combined effect of response set and disturbance of lexical access on response selection in word association (see discussion).

DISCUSSION

The main results of this study may be summarized as follows. First, in word association by normal subjects the selection of common responses that have a feature specifying relationship to the stimulus word was more effortful in late acquired than in early acquired stimulus words. Second, in patients with right brain damage the retrieval of lexical knowledge that was based on feature specification was disturbed in late acquired words but not in early acquired words. Third, the retrieval of defining lexical knowledge, that predominantly occurred in word association to late acquired words, is automatically achieved in normal subjects and is relatively intact in patients with right brain damage.

These results suggest that in the case of early acquired words, feature specifying relationships can be automatically derived from the associative network organisation of memory, as are the defining relationships to late acquired words. However, in the case of late acquired words feature specification, which predominantly occurs when defining knowledge based on verbalisation is absent in the associative network organisation, some extra operation is needed in the word selection process. This operation may consist of the active use of imagery, since in patients with right brain damage the retrieval of this lexical knowledge is disturbed. Since feature specification is less effortful in early acquired stimulus words and also relatively intact in right brain damaged patients, it seems reasonable to suppose that imagery processes for these words are more automatically available, probably because of their (bilaterally represented?) sensorimotor base.

If this hypothesis is correct then certain notions about lexical representation and word retrieval need revision. As already mentioned, semantic memory can be conceived of as an associative network structure with semantically related concepts directly linked within the network (Collins & Loftus, 1975; Chiarello, 1988). In word association the stimulus word activates the corresponding representation in the network. Some of this activation will spread, via links in the network, to semantically related words. Spreading of activation automatically makes available a set of alternatives for more controlled response selection in word association. According to Chiarello (1988) this automatic spreading of activation takes place in both hemispheres, but the more controlled response selection is dependent on the left hemisphere. Our data suggest that equipotentiality for semantic activation exists only for sensorimotor knowledge. When meaning activation presupposes verbal knowledge, as is the case in late acquired stimulus words, then in right brain damaged patients predominantly defining knowledge becomes available. This is also the case when a feature specifying response set is induced (block list). So, it may be hypothesized that in late acquired words spreading of activation in the left hemisphere makes use of defining links (superordination and synonymy), whereas spreading of activation in the right hemisphere is based on the active use of imagery. This latter kind of spreading of activation is more effortful than is the case in spreading of activation based on sensorimotor or defining knowledge. So, the general assumption that all spreading of activation is automatic needs revision.

REFERENCES

Chiarello, C. (1988). Semantic priming in the intact brain: Separate roles for the right and the left hemispheres? *In* Chiarello (ed), *Right hemisphere contributions to lexical semantics.* Berlin: Springer Verlag. pp 59-69.

Collins, A.M. & Loftus, E.F. (1975). A spreading-activation theory of semantic processing. *Psychological Review* **82**, 407-428.

Elbers, L.H. & Loon-Vervoorn, W.A. van (1990). Taal als resultaat van actief leren. *Logopedie en Foniatrie* **69**, 19-23.

Groot, A.M.B. de (1980). *Mondelinge woordassociatienormen.* Lisse: Swets & Zeitlinger.

Groot, A.M.B. de & Bil, J.M. de (1987). *Nederlandse woordassociatienormen met reactietijden.* Lisse: Swets & Zeitlinger.

Hennink, J.H. (1989). Een vergelijkend woordassociatie-onderzoek: epilepsie en normen. *Doctoraalonderzoek:* Vakgroep Psychonomie, Rijksuniversiteit Utrecht.

Loon-Vervoorn, W.A. van (1985). *Voorstelbaarheidswaarden van Nederlandse woorden.* Lisse: Swets & Zeitlinger.

Loon-Vervoorn, W.A. van (1988). Imageability and age of acquisition as determinants of word selection in word association. In: C. Cornoldi (ed), *Imagery and cognition,* Preproceedings of the Second Workshop on Imagery and Cognition, 101-110. Padova: Universita' di Padova.

Loon-Vervoorn, W.A. van (1989). *Eigenschappen van basiswoorden.* Lisse: Swets & Zeitlinger.

Loon-Vervoorn, W.A. van & Van Bekkum, I.J. van (in press). *Nederlands woordassociatielexicon.* Lisse: Swets & Zeitlinger.

Loon-Vervoorn, W.A. van & Ham-van Koppen, M. van der (1988). The importance of age of word acquisition for imageability in word processing. *In* M. Denis, J. Engelkamp & J.T.E. Richardson (eds.), *Cognitive and neuropsychological approaches to mental imagery.* Dordrecht: Martinus Nijhoff. pp 99-107

Loon-Vervoorn, W.A. van & Willemsen, I. (1989). Selectieve verstoring van lexicale kennis bij demente en normale ouderen. *Tijdschrift voor gerontologie en geriatrie 20,* 59-66.

Author Index

This Page Intentionally Left Blank

Subject Index

This Page Intentionally Left Blank