UNDERSTANDING PICTURE-TEXT INSTRUCTIONS

M.D. Glock
George Bieger
Thomas Crandell
Marie Knowlton
Fran Schorr
David Stone

Final Report

Reproduction in whole or part is permitted for any purpose of the United States Government.

This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-80-C-0372, Contract Authority Identification Number NR157-452.

This report, No. 12, Series B, is issued by the Reading Research Group, Department of Education, New York State College of Agriculture and Life Sciences, a Statutory College of the State University, Cornell University, Ithaca, N.Y. 14853. It is supported in part by Hatch Funds Project #424 PRES. STRAT. IMP. COMP. PRINT TECH. MAT.

Approved for public release; distribution unlimited.
### Understanding Picture-Text Instructions

**Title**: Understanding Picture-Text Instructions  
**Authors**: Marvin D. Glock, Cornell University, and George R. Bieger, Bucknell University  
**Performing Organization**: Cornell University, Department of Education, NY State College of Agriculture and Life Sciences; A Statutory College of the State University  
**Contractor**: N04-80-C-0372  
**Type of Report**: Final Report  
**Period Covered**: 3/1/80-3/31/84  
**Performing Organization Report Number**: Technical Report B Series  
**Monitor Agency Name**: Personnel and Training Research Programs, Office of Naval Research (Code 458)  
**Report Date**: March 31, 1984  
**Number of Pages**: 21  
**Security Class.:** Unclassified  
**Distribution Statement**: Approved for Public Release; Distribution Unlimited  

This research was also supported by Hatch Funds Project #424, PRES, STRAT. IMP. COMP. PRINT TECH MAT, N.Y. State College of Agriculture and Life Sciences; A statutory college of the State University.  

The project investigated a variety of aspects of how procedural instructions comprised of pictures and text are comprehended and executed. The research focused on differences in comprehension and performance attributable to variations in the organization of the information, the metacognitive strategies employed by readers, and the interaction of readers with materials. The various studies confirmed that differences in comprehension and performance are related to these.
factors. Several categories of information were found to be very important if not essential for execution of procedural instructions, and it was further found that certain metacognitive strategies affected the speed and accuracy of performance. Additional investigations revealed that subjects rely on specific features of objects for purposes of identification and use those features to infer functional properties of the objects depicted. These findings are discussed in terms of a recursive model of the cognitive processing of picture-text information and implications for the design of instructional materials are discussed.
This project investigated a variety of aspects of how procedural instructions comprised of text and pictures are comprehended and executed. The research focused on differences in comprehension and performance attributable to variations in the organization of the information, the format of the information, the metacognitive strategies employed by readers, and the interaction of readers with materials. The various studies confirmed that differences in comprehension and performance are related to these factors. Several categories of information were found to be very important if not essential for execution of procedural instructions, and it was further found that certain metacognitive strategies affected the speed and accuracy of performance. Additional investigations revealed that subjects rely on specific features of objects for purposes of identification and use those features to infer functional properties of the objects depicted. These findings are discussed in terms of a recursive model of the cognitive processing of picture-text information and implications for the design of instructional materials are discussed.
UNDERSTANDING PICTURE-TEXT INSTRUCTIONS

The project discussed in this report had as its fundamental purpose the investigation of some of the cognitive processes involved in the acquisition, comprehension, and memory of information presented in picture-text combinations. The use of materials containing pictures and texts is very widespread, ranging from children's storybooks and beginning reading texts designed for young children, to complex technical manuals intended for use by engineering and scientific professionals. Because of the diversity in this domain of application and the realization that theoretical and methodological considerations dictated against addressing our research questions at a level that was too broad, we decided to narrow the scope of our investigations. Our research focused solely on issues related to the execution of technical procedural instructions comprised of text and pictures. We hoped to gain a greater understanding about how people comprehend a set of directions explaining how to do something, how to carry out some set of procedures, or how to achieve some goal. In particular, we intended to identify some of the variables and interactions among variables that affect performance on tasks involving the assembly of objects from a set of parts.

Questions about how procedural assembly instructions are comprehended and executed have important practical as well as theoretical significance (Pine & Bieger, 1980). Tasks involving the assembly of objects in accordance with instructions provide a kind of microcosm of human cognitive functioning. From the point of cognitive psychological theory such a microcosm can be a fertile domain for studying such processes as attention,
perception, comprehension, and memory and how those cognitive processes affect performance. Further, it is a domain capable of revealing some of the ways that those processes interact with a variety of task and individual variables. Therefore, investigations in this domain of procedural information can add to our understanding of how people acquire, process, and translate into action information in general.

From a practical perspective, issues related to the comprehension and memory of procedural assembly information have important implications for preparers of documents and instructions for use in the execution of all types of procedures. Additionally, the comprehensibility of procedural instructions has very practical implications which can be grasped most easily by examining the consequences of poorly understood instructions. These consequences can vary from trivial or merely inconvenient to catastrophic.

Many of us have had the experience of attempting to assemble an appliance or a child's toy using instructions that seemed to have been prepared by someone whose sole intent was to confuse. There is no accurate count of the number of sleepless Christmas eve's spent by parents disassembling and then reassembling bicycles in order to install a crucial washer that had been overlooked. Nor is there any tally of the number of back-yard barbeque grills that have their lid handles installed upside down because, having installed it incorrectly (due to misunderstanding the instructions), the owner found that the handle could not be removed without destroying it. At the other extreme, the potential consequences of misunderstood procedural information by a fire control technician or air
defense commander, or a nuclear power plant operator are of tragic and
catastrophic proportions.

Having examined many previous attempts to investigate the acquisition
of information from picture-text combinations, we realized that the general
lack of conclusive findings was due, at least in part, to the widespread
failure to consider relevant theoretical perspectives on this process. We
have adopted a particular perspective that is consistent with current
thinking in the area of cognitive science, that is that perception and
comprehension are processes that are cognitively mediated and not merely
direct responses to the world "as it is". What we perceive and comprehend
is not merely what we sense, but an alteration in some way, either
consciously or unconsciously, of what we sense. This is not to say that
the environment to be sensed is random and that meaning is solely
constructive, but rather that meaning is the product of the interaction
between a somewhat orderly arrangement of the environment (in which there
are constraints on the range of possible meanings) and the active cognitive
mediation of sense data by an individual. The quality and quantity of
that cognitive mediation is a function of an array of individual variables
including individual experiences.

The organized collection of past experiences, which governs the
interpretation of sensory input, has been termed a "schema". This term
became prominent through its use by Bartlett (1932) in accounting for
perceptual and memorial phenomena, and has been adopted by many modern
cognitive psychologists. The role of the schema in the interactive process
mentioned earlier has been described by Neisser (1976) in his formulation
of a model which takes into account the important variables involved in the perception - comprehension - action chain. This model, depicted in Figure 1, rejects the simple seriality of earlier information processing models, yet recognizes that the environment is not random at all, but that there is an elegant order in what we sense. According to this model, perception, and ultimately meaning, is a result of the interaction of the schemata of the perceiver and the structure of the information available to the senses.

In order to generate experimental hypotheses based on such a cyclical model, one must be able to operationalize each component of the model shown in Figure 1. That is, one must be able to describe the relevant characteristics of the perceiver, the relevant characteristics of the stimulus, and the ways in which the two interact. In the following sections of this report we shall describe series of studies in which these various components were operationally defined, measured, and systematically controlled with respect to the domain of procedural assembly instructions comprised of text and pictures. The use of several techniques for operationalizing the components of Neisser's (1976) cyclical model is illustrated in Figure 2. The stimulus characteristics are described using a method of discourse analysis and according to a taxonomy for classifying categories of information. The schema component is considered that part of the cycle which accounts for individual differences. It contains the aptitudes, abilities, experiences, knowledge, and other relevant attributes that a learner brings to the situation. Various methods for identifying
and measuring those factors have been attempted over the years, including intelligence and ability testing, descriptions according to personality types, and classification according to cultural or socio-economic variables. We have chosen to employ more recent techniques (Crandell, 1979; Crandell & Glock, 1981) such as descriptions based on educational cognitive style (Hill, 1976; Witkin, 1969) or according to the metacognitive strategies used by learners (Brown, 1975). The remaining component, that dealing with the manner in which the individual selectively explores the stimulus, has been quantified through the use of eye movement monitoring methods, in which the location, duration, and sequence of eye fixations were recorded.

Insert Figure 2 about here.

One of the important assertions of modern cognitive psychology is that, to a large degree, human behavior and experience is the product of mental processes which mediate sensory information in a variety of ways (Neisser, 1967). The result of this processing (i.e., attending, selecting, encoding, retrieving, etc.) is a unified subjective experience which enables the individual to obtain meaning from and act purposively on the environment. Implicit in these notions is the recognition that the range of possible meanings of an event, a text passage, or a procedural instruction is not unlimited, but is in fact considerably constrained (Keil, 1981). The constraints on meaning can be described in any number of ways such as physical or linguistic and are largely a function of an individual's experiences. In the study of instructions, we may consider
constraints on the ways that a particular instruction may be understood but we can also think of the instruction itself as a constraint. For example, consider the execution of a procedural task as a problem faced by a person. The steps taken to execute the procedures can be thought of as proceeding through a problem space. At each point in the process, the person is faced with a decision about what to do next, and is faced with a range of choices. This range of choices is constrained by the person's knowledge of the task at hand, the knowledge of the "world" in general, the choices that have already been made, etc. Instructions can then be considered as a means of placing added constraints on decisions in the problem space.

The research described in this report has dealt particularly with assembly tasks and has examined some of the ways that decisions about what to do next in an assembly are constrained, either by the assembly instructions, the experiences of the assembler, or the objects themselves. The remainder of the report summarizes some of these studies and their specific findings.

**Information Content of Picture-Text Instructions**

One of the major criticisms of past research on pictures and texts has been that the materials used in that research were rarely described in terms of their relevant characteristics (Stone, 1980; Stone & Glock, 1981). One possible remedy is the development and use of a "taxonomy" of categories of information to classify the content of such picture-text materials in a way that would permit generalizeability to other materials. Such a taxonomy was developed and tested in an experimental situation (Bieger & Glock, 1982a). The categories of information and their
definitions are shown in Table 1.

Insert Table 1 about here.

Three categories from this taxonomy were hypothesized to be very important for the successful execution of procedural instructions and when various combinations of the information were presented to 108 subjects the data indicated that when Inventory information was comprehended unequivocally Operational, Spatial, and Contextual information were virtually essential for a person to be successful in executing the instructions. Subjects who received instructions that contained all of these categories of information completed the assemblies in significantly less time and with significantly fewer errors than did subjects whose instructions were missing one or more of these categories of information.

These findings suggest that the categories of information described in the taxonomy may be a functional classification mechanism for describing the information content of procedural instructions. The identification of three categories of information that so dramatically affected performance has important implications for a cognitive theory of comprehension as well as for the preparation of procedural instructional materials. In a follow-up study Bieger and Glock (1982b) manipulated the location, in picture or text, of spatial, contextual, and operational information and measured the effect of location on speed and accuracy of performance in an assembly task. It was found that textual presentation of spatial information produced fewer errors, while pictorial presentation of spatial information reduced performance times dramatically. It was further found
that pictorial presentation of contextual information substantially reduced assembly times and slightly reduced the number of assembly errors. There were no differences between pictorial and textual depictions of operational information. Again, these findings have important implications both for preparers of instructional materials and for theories dealing with the acquisition and comprehension of information.

The Perception of Objects: Orientation and Function

In the preceding section we summarized some studies that investigated differences in performance as a function of changes to the information content of the instructions. In a sense, we were examining some of the constraints placed on assembly actions by the instructions given to the assembler. There are, of course, other factors which constrain the action decisions of an assembler. Some of these constraints are a function of the way in which objects, or depictions of objects are presented to and perceived by the viewer/assembler. Two such classes of constraints are those dealing with the identity of an object (Canonical information) and those pertaining to how an object functions or can be used (Functional information).

In a study that examined how people make decisions about the functional uses of objects (Knowlton, Keil, & Glock, 1982) it was hypothesized that subjects would attempt first to identify the object and from the information available from the object's identity determine the possible functions of the object. This hypothesis was tested by recording the eye movements of 43 subjects as they viewed pictures of objects and made decisions regarding their functional uses. The data from this study
supported this hypothesis for certain classes of objects. For objects having clear canonical orientations subjects attended to those features that enabled them to identify the objects. However, for objects not having a strong canonical orientation and for unfamiliar objects subjects attended those features that were related to the functional attributes of the object.

These findings form the basis of a cognitive model for determining the functional uses of objects (Figure 3) which can serve as the focus for further research on the relationship between functional and canonical information of objects and how this information is used in decision making. Such research may make significant contributions to theories of object perception, semantic theories of concept acquisition and comprehension, and general theories of cognitive processing.

Comprehension Monitoring Strategies

In order for a person to execute instructions accurately it is essential that they recognize when they have failed to understand a direction and know what to do when comprehension failures occur. The term metacognition has been used to refer to the knowledge that one has of his or her own cognitions and the regulation of those cognitions (Flavell, 1979). Examples of self-regulatory activities include checking the outcome of an action, planning a next action, monitoring effectiveness, and testing, revising, and evaluating one's strategies (Brown, 1978). Thus,
understanding is related to how well a person can monitor his or her ongoing comprehension of some material. This comprehension monitoring is an important part of metacognition (Winograd & Johnston, 1980).

The various comprehension monitoring techniques that readers use when trying to comprehend procedural instructions and execute those instructions were categorized in a taxonomy of comprehension monitoring strategies by Schorr (1982). This taxonomy is shown in Table 2.

In a study designed to examine how different comprehension monitoring strategies are related to performance Schorr and Glock (1983) observed the strategies employed by 68 subjects as they read various sets of procedural instructions that differed in mode of presentation (pictures, text, or pictures with text) and in the degree of explicitness of operational or "how to" information.

It was found that a number of comprehension monitoring strategies are related to comprehension. Surprisingly, planfulness was generally unrelated to comprehension. The other two taxonomic categories ("ways of following instructions" and "detection of errors") were significantly related to comprehension. For example, subjects displaying more care (either by checking their work or by showing a concern for detail) made fewer uncorrected errors but took longer to complete the task than did their less careful counterparts. Similarly, subjects who were quick in detecting their errors left fewer errors uncorrected than did those who were slower in noticing mistakes. There were no differences between these
groups in time to complete the task. Thus, being careful involves a trade-off in performance, increasing accuracy but decreasing speed, while efficient error detection involves no such trade-off.

The relationship between the remedial strategies, included in the fourth taxonomic category, and comprehension is less clear. The data indicate that the total number and variety of strategies employed were associated with longer completion times but not with greater accuracy.

The effect of explicit "how to" or operational information on comprehension and performance was also somewhat mixed. Explicit operational information improved the accuracy of performance, thus supporting Bieger and Glock's (1982a) contention that operational information is essential for procedural tasks, however explicit operational information did not have the same facilitative effect on performance time.

This study provides some data that are relevant to the design of procedural instructions and further provides insight into the relationship between some comprehension monitoring strategies and performance. These findings point to the need for additional research in these areas, yet they offer the hope of improving understanding and performance in procedural tasks.

**Reader-Material Interaction**

In several of the studies conducted as part of this project the interaction between readers and the stimulus materials was examined by recording the reader's eye movements during reading (Hirschfeld & Bieger, 1981). Another method that has been developed employs a computer-based training system that can both present information in textual and pictorial
media and can record the information needs of the learner as he/she reads the procedural instructions. This system (Stone & McMinn, 1982) has enabled researchers to present procedural information to a learner and to observe the pattern of exploration of the information presented. The computer-based delivery system enables the learner to obtain needed information from the computer on demand. The information demanded varies in mode of presentation (textual, computer-generated graphics, video pictures from computer-controlled videodisc displays, or any combination of these) and level of detail or generality.

One of the research applications of this technology has been a study involving the use of "hypertext" and a hierarchical arrangement of text and pictorial information which provides various degrees of embellishment and elaboration to the basic instructional information. Any concept, procedure, or instructional component can be elaborated or embellished by the reader touching the word or object depiction on the display screen. This action causes the computer to provide a more detailed depiction of the item selected. By recording the pattern of information requests, the experimenter can infer information needs and processing strategies.

Although the use of this technology is in its infancy the results of some preliminary studies are encouraging. For example, when the directions for a procedural assembly task were presented to people on paper (using text and graphics) about two-thirds of the people made assembly errors which were not recognized as mistakes. However, when people were taught the same task using the hypertext approach and the computer-based delivery system, they did not make mistakes. In fact, no subjects left any
uncorrected errors using the hypertext approach. The apparent utility of this approach as a learning delivery model is promising.

Similarly, it is apparent that the computer-based training system is a powerful research tool which may provide investigators with another much needed method for observing the way in which learners interact with instructional materials.

Although the research summarized in this report raises many new questions about how procedural information is perceived, understood, and acted upon, it is quite apparent that an important start has been made. A more complete understanding of the basic cognitive processes involved in acquiring and comprehending information and translating that information into action is a necessary part of any useful work toward the formulation of a general theory of cognition.
References


Table 1

**Categories of Information**

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inventory</strong></td>
<td>Information which specifies what objects or concepts are depicted.</td>
</tr>
<tr>
<td><strong>Descriptive</strong></td>
<td>Information which specifies the figurative detail of the objects or concepts depicted.</td>
</tr>
<tr>
<td><strong>Operational</strong></td>
<td>Information which directs an implied agent to engage in a specified action.</td>
</tr>
<tr>
<td><strong>Spatial</strong></td>
<td>Information which specifies the location, orientation, or composition of an object.</td>
</tr>
<tr>
<td>Location</td>
<td>Describes the position of an object in space in relation to another object or fixed point of reference.</td>
</tr>
<tr>
<td>Orientation</td>
<td>Describes the orientation in space of an object.</td>
</tr>
<tr>
<td>Composition</td>
<td>Specifies areas of filled or empty space and the density of filled space.</td>
</tr>
<tr>
<td><strong>Contextual</strong></td>
<td>Information which provides the theme or organization for other information which may precede or follow it.</td>
</tr>
<tr>
<td><strong>Covariant</strong></td>
<td>Information which specifies a relationship between two or more other pieces of information which vary together.</td>
</tr>
<tr>
<td><strong>Temporal</strong></td>
<td>Information about the time course of states or events.</td>
</tr>
<tr>
<td><strong>Qualifying</strong></td>
<td>Information which modifies other information by specifying the manner, attributes, or limits of that information.</td>
</tr>
<tr>
<td><strong>Emphatic</strong></td>
<td>Information which directs attention to other information.</td>
</tr>
</tbody>
</table>
Table 2

A Taxonomy of Comprehension Monitoring Strategies

A. Planfulness
   1. Initial approach to instructions
   2. Selection of parts

B. Ways of Following the Instructions
   1. Display concern with detail
   2. Check work after completion

C. Detection of Errors

D. Reactions to Mistakes or Problems (Remedial Strategies)
   1. Examine the construction vs. take it apart immediately
   2. Examine instructions subsequent to the problem
   3. Examine instructions preceding the problem
   4. Reexamine the same instruction continuously
   5. Experiment with the parts
   6. Hypothesize; reason
   7. Build another part of the assembly
   8. Examine the sheet depicting the parts
   9. Compare two instructions
  10. Replace one part with an identical part
  11. Rebuild the problem area in exactly the same way
Figure 1. Perceptual cycle (after Neisser, 1976)
STIMULUS
Picture-text reading materials
Semantic content of materials is described by propositional analysis of content. Lists concepts and relations among concepts depicted in material.
Propositions are classified according to the taxonomy of categories of information.
Modifies

Samples

SCHEMA
Abilities, aptitudes, knowledge, and experience of reader.
Indicated by various measures of cognitive style, or by metacognitive strategies used by reader.

EXPLORATION
Manner in which the reader interacts with the reading materials by selectively sampling.
Measured by monitoring a reader's eye movements and recording location, duration, and sequence of eye fixations.

Figure 2. Applicability of perceptual cycle to reading situations.
Figure 3. A model of the processes necessary for decision making regarding functional uses of common objects.
Navy Personnel & Training Research Group
Code 442PT
Office of Naval Research
Arlington, VA 22217

1 Dr. Robert Smillie
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Robert G. Smith
Office of Chief of Naval Operations
OP-987H
Washington, DC 20350

1 Dr. Richard Snow
Liaison Scientist
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

1 Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152

1 Dr. Frederick Steinheiser
CNO - OP115
Navy Annex
Arlington, VA 20370

Marine Corps

1 Dr. William Greene
Education Advisor (EO31)
Education Center, MCDEC
Quantico, VA 22134

1 Dr. A.L. Slapkowski
Scientific Advisor (Code RD-1)
HQ, U.S. Marine Corps
WASHINGTON, DC 20380

Army

1 Technical Director
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Harold F. O'Neill, Jr.
Director, Training Research Lab
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

1 Dr. Robert Sasser
U. S. Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333
Private Sector

1 ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

1 Dr. Anders Ericsson
Department of Psychology
University of Colorado
Boulder, CO 80309

1 Mr. Wallace Feurzeig
Department of Educational Technology
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02138

1 Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1 Dr. Dedre Gentner
Bolt Beranek & Newman
10 Moulton St.
Cambridge, MA 02138

1 Dr. Robert Glaser
Learning Research & Development Center
University of Pittsburgh
3939 O'Hara Street
PITTSBURGH, PA 15260

1 Dr. Joseph Goguen
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

1 Dr. Daniel Gopher
Faculty of Industrial Engineering
& Management
TECHNION
Haifa 32000
ISRAEL

1 Dr. James G. Greeno
LRC
UNIVERSITY OF PITTSBURGH
3939 O'HARA STREET
PITTSBURGH, PA 15213

1 Dr. David Kieras
Department of Psychology
University of Arizona
Tuscon, AZ 85721

Private Sector

1 Dr. Walter Kintsch
Department of Psychology
University of Colorado
Boulder, CO 80302

1 Dr. Stephen Kosslyn
1236 William James Hall
33 Kirkland St.
Cambridge, MA 02138

1 Dr. Jill Larkin
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213

1 Dr. Alan Leagold
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1 Dr. Jim Levin
University of California
at San Diego
Laboratory for Comparative Human Cognition - D003A
La Jolla, CA 92039

1 Dr. Don Lyon
AFHRL/OT (UDRI)
Williams AFB, AZ 85225

1 Dr. James R. Miller
Computer*Thought Corporation
1721 West Plano Highway
Plano, TX 75075

1 Dr. Tom Moran
Xerox PARC
3333 Coyote Hill Road
Palo Alto, CA 94304

1 Dr. Allen Munro
Behavioral Technology Laboratories
1845 Elena Ave., Fourth Floor
Redondo Beach, CA 90277

1 Dr. Donald A. Norman
Cognitive Science, C-015
Univ. of California, San Diego
La Jolla, CA 92039