This paper reviews theories of transfer of learning to determine whether they offer guidelines for educational approaches. Two classes of theories are described. The first class of theories is based on the notion that the conditions for transfer are established when an original learning event and a transfer event share common stimulus properties. The second class of theories takes the position that facilitative transfer is a product of a successful memory search process. These two classes of theories are compared and their potential for providing guidelines for educational practice is examined. (Author/AA)
This article was written while I was a visiting research professor at the University of Illinois, Center for the Study of Reading, which is supported under Contract No. US-NIE-C-400-76-0116. I would like to thank John Elkins for his comments on an early version of this paper. Particular thanks go to Don Cunningham whose sunny disposition and cheerful words of encouragement contributed enormously to the quality of this paper.
Theories of Learning Transfer

There are few topics more central to the educative process than the transfer of learning. This is obvious when one considers the extent to which performance on a given educational task is influenced by prior learning history, or tries to think of any learning activity which is not influenced by something which was learned before. However, despite the importance of transfer, the topic has been neglected in recent years in the educational and psychological literature. There are several likely reasons for this neglect, including the association of transfer with "training" (a much more narrow and restrictive concept) rather than learning, and a reaction against the experimental tradition in which most basic research on transfer was conducted (e.g., paired-associate and serial learning tasks).

In discussing the boundaries of theories of transfer one can go astray in two ways. First, one can define the boundaries too narrowly. For example, if we were to equate transfer of learning with transfer of training, we would be restricted to talking about a very small segment of interesting transfer problems. But one could also err in the opposite direction. Since virtually all behavior is influenced by prior experience it is possible to equate theories of transfer with general theories of behavior. The current primitive state of psychological theorizing would not support such an ambitious approach.

I hope to steer a path in between these extremes. The domain for my discussion of theories of transfer will be defined by a subset of educationally important problems about which theories of transfer might have something to
In particular I will be considering two kinds of educational problems in this paper. The first is how to go about arranging instruction such that relevant previous learning can facilitate the acquisition of current materials; and the second is how to conduct instruction such that skills and knowledge acquired in schools can be used in solving and dealing with real world problems and events.

The purpose of this paper is to review a number of the theories of transfer which have been proposed, and to discuss the degree to which the theories offer approaches to the two educational problems posed above. In addition, the paper will examine the concept of transfer theory from the perspective of the recently emerging cognitive theories which have come to dominate much of the thinking about psychological and educational issues.

The paper is organized into four sections. The first section will be concerned with several distinctions which have appeared in the earlier literature. This section will consider the difference between lateral and vertical transfer, the difference between specific and nonspecific transfer, and the difference between literal and figural transfer. The second section will consider theories which emphasize the role of environmental events. The essential notion involved in these theories is that facilitative transfer (which will be the focus of this paper) can occur only when the learner recognizes that the transfer material and previously learned material share common features.

The third section of the paper will be concerned with theories which focus on internal cognitive events. This view, which has its origin in recent cognitive theory, takes the position that facilitative transfer
can be enhanced by increasing the likelihood that relevant prior knowledge will be retrieved in appropriate situations.

The final section of the paper contains a summary of the material presented in the previous three sections. This summary will focus on the limitations of environmental theories, the extension provided by cognitive theories, and the educational utility of both classes of theory.

**Basic Distinctions**

In this section of the paper I will review a number of distinctions made by previous writers and introduce a new distinction. Several of these distinctions seem timely and are relevant to the general theme of this paper. They are also of historical interest and are included for the purpose of giving one a better sense of how the concept of transfer of learning has been viewed in the past.

**Lateral and Vertical Transfer**

A number of years ago Gagné (1965) made a distinction between lateral and vertical transfer of learning. Vertical transfer occurs when a skill or bit of knowledge contributes directly to the acquisition of a superordinate skill or bit of knowledge. For example, a student who can multiply and subtract numbers will master the skill of long division more rapidly than a student who has not mastered multiplication and subtraction.

Gagné's (1970) definition of lateral transfer is less sharply focused than his definition of vertical transfer. He refers to lateral transfer as, "... a kind of generalization that spreads over a broad set of situations at roughly the same level of complexity" (p. 231). This is taken to
mean the sort of transfer which occurs when a child recognizes the fractions he is learning about in school are relevant to deciding how to divide a purloined pie into equal shares.

Historically, vertical transfer has received the bulk of attention from both psychologists and educators. Psychologists have been concerned with specifying the conditions (and occasionally, the underlying processes) under which vertical transfer occurs, and educators have been concerned with organizing and sequencing instruction so as to increase the occurrence of vertical transfer.

The relative neglect of lateral transfer is probably attributable to several reasons. One of these is that the historically dominant theoretical perspective for viewing transfer problems is ill suited for analyzing lateral transfer. This perspective, called the environmental perspective in this paper, focuses on an analysis of stimulus elements and is not a powerful tool for analyzing situations where the nature of the stimulus complex is impossible to control.

Another reason for the neglect of lateral transfer is that educators have not been overly concerned with determining whether school learned skills transfer to real world tasks. This situation is rapidly changing however as is testified to by recent interest in "minimal competencies" which presumably are directly related to out of school functioning.

Specific and Nonspecific Transfer

Specific transfer involves a situation where there is a clear similarity between stimulus elements in original learning and stimulus
elements in transfer learning. These stimulus elements can be clearly definable, as in the case of physical attributes such as the orthography or phonology of words and phrases, or the similarity can be less obvious, as in the case of a similarity of meaning between two instructional events. In either case, the notion is that the shared elements will be detected by the learner and this will lead to more rapid acquisition of the transfer task.

The classic examples of specific transfer are list learning experiments. For instance, a subject might learn an initial paired-associate list (an A-B list), and then learn a second list where the stimuli are the same and the responses are semantically similar to the original responses (an A-B' list). In this case it could be shown that the subject would learn the A-B' list faster than would a control subject who learned an initial list consisting of different stimuli and responses (e.g., a C-D list).

Even though the classic examples of specific transfer come from laboratory studies of verbal learning (cf., Ellis, 1965), it probably would not be difficult to find instances of specific transfer in highly structured educational approaches such as Individually Prescribed Instruction (e.g., Cooley & Glaser, 1969) or Project Plan (e.g., Weisgerber, 1971). These programs typically achieve specific transfer by carefully organizing and sequencing related instructional events.

Nonspecific transfer differs from specific transfer in that there are no obvious shared stimulus elements in the originally learned task.
Theories of Learning Transfer

and transfer task. The classic demonstrations of nonspecific transfer are the 'learning to learn' (e.g., Harlow, 1949; Postman, 1969) and 'warm up' (Ellis, 1965) effects frequently found in concept learning and list learning laboratory experiments. More recently, however, Royer and his associates (Royer & Cable, 1975; Royer & Cable, 1976; Royer & Perkins, 1977) have demonstrated nonspecific facilitative transfer with materials more similar to those used in classrooms. They have argued that the facilitation in these studies could be attributed to cognitive events occurring within the learner.

The distinction between specific and nonspecific transfer contains an important implication. The implication is that in the specific transfer situation there is a predictable set of dimensions along which an originally learned task and a transfer task could be similar, and in theory it should be possible to specify those dimensions in advance for any combination of original and transfer tasks. In contrast, the similarities between original and transfer tasks in a nonspecific transfer situation are likely to be impossible to specify on an a priori basis because instances of nonspecific transfer frequently do not share any obvious stimulus similarities. The line of reasoning initiated by this implication will be pursued at a later point in the paper.

Literal and Figural Transfer

At the risk of generating confusion I would like to introduce a further distinction relevant to a discussion of the transfer of learning. The distinction is between what I call literal transfer and figural transfer.
Literal transfer involves the transfer of an intact skill or bit of knowledge to a new learning task. So, for example, we can directly apply our knowledge about the workings of the electoral college to the problem posed by Samuel Tilden's loss to Rutherford Hayes in the 1876 presidential election (Hayes received fewer popular votes).

Most of the material in the past literature on learning transfer could be included under the concept of literal transfer. That is, specific and vertical transfer clearly involve the use of an intact skill or bit of knowledge in a new learning task. Further, many instances of lateral and nonspecific transfer could be considered to be instances of literal transfer.

Figural transfer does not involve the application of an intact skill or bit of knowledge. Rather, figural transfer involves the use of some segment of our world knowledge as a tool for thinking about, or learning about, a particular problem or issue. The clearest instances of figural transfer can be found in the use of figural language such as metaphor or simile. When we say things like "Encyclopedias are goldmines," or "Man is like a computer," we are asking the listener to use the world knowledge they have about the referent of the sentence as a tool for understanding or thinking about the subject of the sentence. One could hardly overestimate the importance of figural transfer in the thinking processes of human beings. Consider, for example, the degree to which much of the current work in psychology is dependent on the man as computer metaphor. Schon (1963) has argued quite persuasively that figural language, and in particular, metaphor, is the central mechanism in the development of new ideas and in the progress of science in general.
Theories of Learning Transfer

As important as figural transfer may be to the human thought process, there has been virtually no attention to the issue in discussions of learning transfer. As was the case with lateral transfer, the neglect of figural transfer can be traced to the environmental focused theory which has dominated the thinking about learning transfer for the past sixty years. The third section of this paper will devote some discussion to figural transfer from the perspective of recently emerging cognitive theory.

Comparison of Previous Distinctions

It is obvious that the distinctions introduced by previous writers (and the one introduced by this writer) are not mutually exclusive. In fact, one could argue that there is considerable overlap between the vertical-lateral distinction and the specific-nonspecific distinction. In general, it is probably the case that instances of vertical transfer could also be considered to be instances of specific transfer and many instances of lateral transfer are also instances of nonspecific transfer.

The distinction I have suggested between literal and figural transfer seems to clearly extend previous distinctions relating to transfer. Figural transfer involves a situation where an entire complex of ideas, concepts, and knowledge is juxtaposed against some new problem or situation. Teaching by analogy, for example, seems to embody this type of transfer. If I told a student that a pain signal from the toe to the brain travels in much the same way that a telephone signal is transmitted from one person to another, I would have activated an entire complex of knowledge. And hopefully, if the analogy is apt, the student would have
benefited from the juxtaposition of prior knowledge and new learning.

In the remainder of the paper, as a matter of convenience, I will use a terminology which is directed primarily towards the settings in which transfer occurs. I will use the term near transfer to refer to instances in which one classroom learned skill, or bit of knowledge, transfers to another classroom skill or bit of knowledge. I will use the term far transfer to refer to situations in which material learned in the classroom transfers to events or problems encountered outside of the classroom.

Theories Emphasizing Environmental Events

In this section of the paper I will review several variants of a theory which suggests that the way to approach a transfer problem is through a careful analysis of the stimulus properties of the learning events. At the heart of this theory is the notion that events which share stimulus properties will be recognized by the learner as being similar, and that the response learned to the first event can then be generalized to the second.

The Theory of Identical Elements

One of the first theories of transfer (ignoring the "formal discipline" theory which has largely been discounted) was proposed by Thorndike and Woodworth (1901). They suggested that transfer from one task to another would only occur when both tasks shared identical elements. Further, they proposed that the greater the number of shared elements, the greater the amount of transfer. When Thorndike and Woodworth talked about "elements,"
they were talking about shared features of the stimulus environment of the two tasks. Thus, two tasks which share some set of stimulus features are possible candidates for learning transfer.

I have labeled identical elements theory, and the theories which follow in this section, "environmental theories" because the critical step in the transfer process involves the recognition that one task (or problem situation) shares a set of stimulus features with another. If the recognition process does not occur, then the transfer of a previously learned response cannot occur.

Thorndike and Woodworth's (1901) theory of identical elements has heavily influenced many of the subsequent considerations of transfer theory. Osgood (1949), for example, formalized what was known about transfer at the time in his influential paper on the "transfer surface." In his paper Osgood indicated that facilitative and inhibitory transfer were functionally related to the similarity and difference relationships between stimuli and responses in an original and transfer task. The essence of these notions had been presented years before by Thorndike and Woodworth (1901); Osgood simply elaborated the theme. Likewise, Ellis' (1965) book on the transfer of learning simply updates the generalization contained in Osgood's (1949) paper, and follows the essential details of the theory of identical elements.

It seems clear that the theory of identical elements, and the subsequent elaborations of that theory, describes in good detail the boundary conditions of most situations which could be identified as instances of near transfer. Further, when applied to educational problems the theory
Theories of Learning Transfer

12

can provide useful guidelines for producing facilitative transfer in instructional settings.

There are, however, two possible problems with identical elements theory. The first problem is that the theory is not really a theory at all. And the second is that the theory does not account for a large segment of transfer situations.

Kintsch (1970) has suggested that the theory of identical elements, and subsequent developments of the theory, are not really theories at all. Rather, Kintsch argued, the theory was really a low-level empirical generalization based on the evidence available at the time. The argument is that the "theory" describes the conditions under which various kinds of transfer will be evident, but it does not specify the psychological processes which must be responsible for the transfer behavior. A theory of transfer, in the true sense of the word theory, would have to specify the psychological processes which support the observable behavior.

This criticism may not, in fact, be valid. One could argue, for example, that a statement of the underlying processes has been made. Hoffding (1892) presented one of the first (and still the most elegant) statements of what the underlying processes might be in his famous analysis of the problem of recall.

Hoffding suggested that recall could be conceptualized as consisting of four components: A-a-b-B. During learning the stimulus event A produces the internal sensory trace, a, which in turn becomes associated with the internal representation of the response event, b, and b gives rise to the overt response, B. After learning, successful recall is...
dependent on A (or a stimulus similar in some sense to A) again being connected to a, and a being connected to b. Recall can fail if A does not contact a, or if a does not reliably elicit b. The first steps in Hoffding's "function" are particularly relevant to an analysis of learning transfer. By definition, learning transfer is evidenced by the ability to apply a particular skill, or bit of knowledge, to situations differing from those encountered during original learning. In Hoffding's terms, in addition to wanting A to elicit B, we also want situations or problems which differ in some unspecified manner from A to also elicit B. The extent to which these other situations fail to elicit B is the extent to which we have failed to produce learning transfer. Hoffding's analysis would suggest that such failures occur most frequently because the learner fails to recognize that the new situation is similar to the one encountered previously. That is, A' (the new situation) does not give rise to a.

Hoffding's (1892) explanation of how similar events come to elicit the same response is essentially the same as the one suggested by more modern writers (cf., Ellis, 1965). The explanation, is, however, incomplete. The first problem is that we do not have an explanation of how A' (the new situation) comes to be connected to a (the old internal sensory trace). In all fairness to Hoffding, and all of the other identical element theorists who followed, it should be mentioned that no one else has come close to solving this problem either. Fame and fortune await the theorist who can explain how an event encountered in one's environment comes to be connected to a particular trace in memory.
Theories of Learning Transfer

14

The second way in which the theory of identical elements is incomplete is in terms of the breadth of the theory. That is, the theory provides an accounting of only those transfer situations which share obvious stimulus features. Upon consideration, the reasons for this become obvious. The theory says that transfer will occur between tasks in those cases where the two tasks share a set of common stimulus features. This means that the class of tasks to which a particular learned skill should transfer should be definable by a careful analysis of the conditions of original learning. One could not, for example, observe transfer between two tasks and then argue that the transfer was due to a set of shared features which were determined a posteriori. Such a situation would involve an obvious tautology.

The fact that identical elements theory can adequately account for transfer only in those situations where there are shared stimulus features which can be established a priori means that the theory has little to say about much behavior usually regarded as instances of transfer. As an example, assume that a child has learned to compute the area of a rectangle. After instruction, one might be able to predict with confidence that the child could successfully solve any problem involving the computation of the area of a rectangle. But now assume that the child is faced with the problem of determining the amount of carpet needed to cover a living room floor. Will he recognize that the mathematical skill learned in school is relevant to the solution of the real-world problem? Identical elements theory could only make a prediction in this situation after analyzing the stimulus features of the problem.
Theories of Learning Transfer

15

The above example points out a limitation of identical elements theory: the theory is only able to predict transfer in those situations where there is a clear and known relationship between an original and transfer task. But obviously transfer occurs in many situations besides these. And equally as obvious, educators are interested in developing instructional treatments which transfer to other stimulus features besides those encountered in classrooms.

Stimulus Generalization and Transfer Theory

The breadth problem in identical elements theory may not, however, be an inherent problem in the theory. One possible way to extend the theory would be to treat the problem of far transfer as a problem of stimulus generalization. Stimulus generalization occurs when a response learned in the presence of a particular stimulus is also elicited in the presence of a similar stimulus. So, for example, a dog conditioned to salivate to a 500 Hz tone will generally salivate (though in lesser quantity) to a 475 Hz tone.

The concept of stimulus generalization could conceivably be applied to the problem of far transfer in the following way. If it were possible to define a class of problems (school related and real world) which could be solved by using a particular skill or bit of knowledge, and if it were possible to identify the defining features of the problem class, then it should be possible to instruct learners such that the presence of the defining features in a given problem would reliably elicit the appropriate skill or bit of knowledge. This could be done by providing learners with
systematic instruction on the defining class of features, and with practice on recognizing instances and noninstances of the problem class.

The stimulus generalization approach to the problem of far transfer involves two critical assumptions: 1) that it is possible to define a class of problems to which a particular skill or bit of knowledge could be applied, and 2) that it is possible to isolate a set of defining features for the class of problems. At the present time the likelihood that there are many areas in which these two assumptions could be met appears remote.

The difficulty with both of the above assumptions is that they are hopelessly complex. Let us first take the problem of identifying the set of features which define a particular class of transfer situations. Consider, for example, the difficulty in identifying the defining features of a concept. What are the defining features of love or magnitude? Even concepts having concrete referents prove to be difficult. What, for example, are the defining features of vehicles or balls (e.g., footballs) or games? Anderson and Ortony (1975) and Wittgenstein (1963) have examined this problem, using examples such as those above, in some detail and have concluded that the possibility of identifying defining features for many concepts was virtually nil. At an even higher level of complexity is the problem of identifying the defining features of a class of tasks to which a learned skill or bit of knowledge might be transferable. Imagine, for example, trying to isolate the defining features of all of the situations (both real world and school based) where one's long division skills are called for.
The arguments above are directly relevant to the complexities involved in defining a class of problems to which a particular skill or bit of knowledge could be applied. In order to accomplish this analysis one would have to establish boundary conditions which would define the set of possible problems and situations to which a particular skill or bit of knowledge might be transferred. In doing this one would run into many of the same problems one encounters in attempting to identify the defining features of a concept.

Educational Applications of Identical Elements Theory

Educational applications of identical elements theory generally take the form of having students master skills and information and then instructing them such that a current problem will be recognized as calling for the use of the previously learned skill. This could be done in a variety of ways. One way would be to have the instructional materials organized and sequenced in such a way that tasks and activities which shared common elements would always be encountered in close temporal order. Another way would be to have the instructional agent (either teacher or learning materials) explicitly point out the relationship between prior and current learning. In this manner the label for the previous activity would become part of the stimulus complex for the current activity.

Both of these procedures are implicit (and sometimes explicit) guiding principles of transfer in several of the highly structured approaches to instruction (e.g., PLAN, IPI). Thus, the theory has provided the basis for some valuable approaches to the problem of near transfer. The theory
does not, however, provide guidelines for developing an approach to the problem of far transfer.

One could assure near transfer by simply instructing students to use a previously learned skill in a particular learning situation. However, it remains entirely possible that a learner could have mastered a particular skill, or bit of knowledge (as evidenced by performance on classroom activities) and still not be able to correctly apply the knowledge or skill to a task which differed from the original conditions of instruction (a far transfer situation). For example, a student could have mastered the skill of computing the area of a rectangle, and not recognize that the skill could be used to compute the square-footage of a rug needed to cover a living room floor. Given the fact that the original learning task and the transfer task do not share obvious stimulus features, identical elements theory is of little help in determining instructional procedures which will increase the likelihood of far transfer.

Environmental Theories in Perspective

The theories described in this section of the paper have dominated the thinking of psychologists and educators concerned with learning transfer since the turn of the century. The theories are associated with our behavioristic tradition, and being such, they are minimally concerned with events occurring inside the learner's head. Instead, the theories are primarily concerned with the observables in the learning situation. That is, stimulus events and response events.
This focus on stimulus and response events has lead to considerable progress in our ability to produce transfer of the specific and vertical varieties: the outcome studies from curriculums such as Distar, or from the behaviorally oriented Follow Through projects (cf., McDaniels, 1975) provide support for this assertion.

Similar progress has not been made, however, in developing procedures for promoting transfer from material learned in schools to problems encountered outside of classrooms. I would suggest that this lack of progress is due to the limitations of environmental theory; namely, the inability to theoretically handle transfer in many nonclassroom situations.

Theories Emphasizing Internal Events

The transfer theories discussed in the previous section were all based on the notion that the critical step in the transfer process was the recognition that one situation shared common elements with another. The theories to be presented in this section are based on the notion that the critical step in the transfer process is the retrieval of a relevant skill or bit of knowledge when a particular problem is encountered.

Early Precursors to Cognitive Theories

By stretching one's imagination a little one could trace the history of some of the ideas contained in this section back to Herbart (1896) and Huey (1908). Herbart's description of how experienced events become part of the "apperceptive mass" bears a more than passing resemblance to more modern descriptions of how incoming sensory experiences become integrated.
Theories of Learning Transfer

20

into existing knowledge structure. In a similar fashion, Huey's discussion of meaning as involving an interaction between an incoming sensory message and existing knowledge is similar to modern views on the topic.

One could hypothesize that these early cognitive views quickly become peripheral views in psychology and education because of the lack of critical conceptual tools. For example, early in this century there was no way to think in any sort of rigorous fashion about the nature of human knowledge. At best, one could talk about amorphous entities like the "aperceptive mass," but such discussions did not lead to any important empirical work.

A more recent precursor to the modern cognitive theories was Sir Frederick Bartlett's (1932) famous work on human memory. Bartlett talked of memory as consisting of schemata (singular-schema) which were dynamic storage structures which constantly changed as a function of the acquisition of new material. According to Bartlett, remembering involved a process whereby the schema reconstructed its previous state through a process of inference. That is, the schema inferred what its past state must have been on the basis of its current state.

Bartlett's (1932) theory ran into two kinds of problems. The first was that the evidence which Bartlett offered in support of his theory was disputed in subsequent research (e.g., Gauld & Stephenson, 1967; Gomulicki, 1956; Zangwill, 1972), and the second was that Bartlett's concept of knowledge schemas was no better defined than Herbart's apperceptive mass. The empirical research which derived from Bartlett's theory was primarily devoted to reproducing his results. None of it was devoted to testing the theoretical utility of his schema concept.
Theories of Learning Transfer

A Cognitive Theory of Transfer Based on a Relatively Static Concept of Knowledge Structure

The theory that will be discussed in this section emerges from information processing theories of human learning and memory which began to appear in the literature about a decade ago. Prior to describing the theory several assumptions need to be specified. First, the theory is based on the assumption that human memory is a highly structured storage system in which information is both stored and retrieved in a systematic manner. Thus, the theory to be described, contrary to the environmental theories discussed previously, makes strong assumptions about the nature of underlying memory representations. Second, the theory makes the assumption that the "richness" of knowledge structure is not uniformly constant, with richness referring to the number of interconnections between the "units" (e.g., nodes, propositions, etc.) in the structure. Thus, some parts of knowledge structure can be richly elaborated with a very large number of interconnections between the units, and some can be relatively impoverished with few interconnections between the units.

In addition to the above assumptions, the theory also has a fundamental premise: that comprehension is a necessary, but not sufficient, precursor to educationally relevant transfer of learning. We can, of course, learn information that has not been comprehended (by rote memorizing, for example). However, the conditions under which we can recall, or make use of, an uncomprehended message are very narrow. In fact, successful recall on use of an uncomprehended message (particularly after the passage
of time) probably depends to a large degree on the reinstitution of the same conditions under which the message was experienced. If this analysis is correct, then it follows that comprehension is a necessary first step in establishing the conditions for educationally relevant learning transfer.

The premise that comprehension must occur before transfer of learning can take place is probably so obvious, it may appear trivial. But the premise has an important implication. The implication is that if we understood the processes underlying comprehension, and if we understood the conditions which give rise to those processes, we will have moved a significant step toward specifying those conditions which give rise to learning which will transfer to a variety of situations.

My own view of comprehension has been heavily influenced by the writing of John Bransford, Nancy McCarrell, and Jeffery Franks (e.g., Bransford & McCarrell, 1974; Franks, 1974). These writers have argued that an adequate approach to linguistic comprehension must begin with a consideration of processes occurring within the comprehender as well as a consideration of the linguistic input. In brief, they argue that the process of comprehension entails the drawing of relations between the world knowledge (Franks uses "tacit knowledge," cf. Polanyi, 1966) possessed by the comprehender and the linguistic characteristics of the input message. In the event that relations cannot be drawn between world knowledge and linguistic input, comprehension will not occur. As a simple demonstration of this, consider the following sentences from Bransford and McCarrell (1974): "The haystack was important because the cloth ripped." "The trip was not delayed because the bottle broke." "The notes
were sour because the seam split." Each of these sentences is grammatical, and each refers to events or objects which are very familiar. However, we find these sentences difficult to comprehend because, by themselves, they are difficult to relate to our world knowledge. They become easy to comprehend, however, in the context of the words, parachute, ship launching, and bagpipes.

Bransford and McCarrell (1974) and Franks (1974) have assumed a relatively "weak" view of the comprehension process; weak in the sense that they do not specify the nature of the relational process they describe. A somewhat stronger view would be that comprehension entails a structural integration of the linguistic input into relevant existing knowledge structure. One implication of this view is that previous linguistic input can be accessed by making contact with the knowledge structure into which the input has been integrated. As will be seen shortly, this implication will be important in speculating about the conditions which give rise to certain kinds of learning transfer.

With this overview of how the comprehension process works, we can now consider the recall process. In its most general form (a number of more specific views exist) the process of recall is seen as entailing a search of the knowledge structure network until the relevant information is located. More specifically, search is initiated at a particular node or nodes (search could be parallel) in the memory network, and activation spreads from that node to connecting nodes (cf., Collins & Loftus, 1975). This spread of activation continues until the searched for memory node is encountered or until the search process is discontinued.
We now have a model of how information gets represented in human memory, and how we recall previously stored information. Let us now consider the implications of this model for learning transfer. From the perspective of the model I have described, the critical aspect of learning transfer does not revolve around the process of recognition as did the previous theories. Rather, the critical aspect of the present theory involves the process of retrieval. That is, the likelihood that learning transfer will occur is determined by the probability of retrieving the relevant prior learning during the search process. For example, the student learning long division, when told that one of the first steps involves multiplication, can enter memory at the node involving multiplication, and retrieve the appropriate information. In a similar fashion, the student who is told about the role of the electoral college in the 1876 presidential election can enter memory and retrieve the fact that plurality in the electoral college does not necessarily mean plurality in the popular votes, and thereby understand what otherwise would have been an anomaly.

The cognitive theory of transfer just described adds little to our ability to develop instructional procedures which promote near transfer. The environmental theories described previously seem more directly applicable to this type of transfer situation. What cognitive theory does do, however, is provide a powerful heuristic for thinking about transfer from school learning to real-world situations. In addition, the theory provides some guidelines for developing educational practices which could enhance the likelihood that far transfer will occur.
Educational Relevance of Cognitive Transfer Theory

Cognitive theory suggests that the likelihood of transfer is dependent upon the likelihood of encountering a relevant bit of information or skill during the memory search process. Given this framework the educational problem becomes one of increasing the probability that relevant material learned in the classroom will be retrieved when the individual is faced with a particular real-world problem. Since probability of retrieval is directly related to the number of interconnections between the school learned skill and the remainder of world knowledge structure, it follows that any educational procedure which increases the "richness" of this interconnecting network will also increase the likelihood of far transfer.

There are probably many ways to increase the richness of the interconnections in knowledge structure, but one way which seems to me to have particular promise is the use of application questions. As a hypothetical illustration of how this might work, consider the following example. Let's assume that two students are to be taught percentages. The first student is given repeated practice, consisting of numerical examples, until the teacher can say with confidence that the student has mastered the computation of percentages. At the end of instruction we might illustrate the representation of the newly acquired skill as I have in Figure 1.

-----------------------------
Insert Figure 1 about here
-----------------------------

Notice that the principle way of accessing the material contained in the percentage skill node is through the complex of skills and knowledge connected to the school learned math skills node.
Theories of Learning Transfer

26

Consider a second procedure for teaching the percentage skill. Imagine that our student again completed an instructional sequence which assured that the percentage computation skill had been mastered. Now, however, instead of proceeding on to a new instructional topic, we provide the student with a series of questions and problems which require the application of the learned skill. For example, we might ask the student to:

a) Compute baseball player's batting average; b) Determine the amount of ingredients needed in a cooking recipe if the dish is to serve six people rather than the four the recipe specifies; c) Compute the annual interest payment on a loan at a given dollar amount and a given interest rate; d) Compute the amount to be set aside out of a weekly pay check if 8% of one's income is to go into savings, etc. Additional guidelines for generating application questions are contained in Royer and Allan (1977).

Given the sorts of exercises mentioned above, we might expect the student's knowledge structure to appear as in Figure 2.

---------------------------------------------------------------------
Insert Figure 2 about here
---------------------------------------------------------------------

Notice that the primary difference between the representation in Figure 1 and Figure 2 is in the number of connections between the percentage skill and real-world knowledge. Given this situation it should be the case that a greater range of real-world problems dealing with percentages could be solved by the student with the richer connections between the skill and real-world knowledge.

There is a cautionary note, however, about using application questions to induce far transfer. The basic skills must be mastered before exposing
Theories of Learning Transfer

students to application questions. A number of both laboratory and school-based research projects (Andre, Smid, Groth, & Runge, Note 1; Gall, Ward, Berliner, Cohen, Crown, & Elashoff, 1975) have demonstrated that application questions at times inhibit rather than facilitate learning. A good guess would be that these negative effects of application questions are frequently associated with inadequate mastery of basic material.

An Evaluation of the Cognitive Theory of Transfer

It seems apparent that the cognitive transfer theory presented in the previous section extends our ability to think about, and to design educational approaches for, many kinds of transfer problems. More specifically, it provides a tool for considering far transfer problems as well as near transfer problems.

What the theory does not do, however, is to provide a vehicle for thinking about figural transfer. The cognitive theory I have presented provides a reasonable account of the comprehension process (incoming information is integrated into existing knowledge structure), the learning process (information is added to knowledge structures containing related information), and the retrieval process (search begins at an appropriate node and activation spreads from that node until the required information is located or until search is terminated). All of the above activities can be conceptualized as involving discrete units (e.g., ideas, propositions, etc.). Where the theory begins to have trouble is when we think of situations where an entire complex of previously learned information is activated simultaneously, and is used to facilitate the learning of new information.
Consider a series of experiments by Royer and his associates (e.g., Royer & Cable, 1975; Royer & Cable, 1976; Royer & Perkins, 1977) as examples of this sort of transfer. They presented subjects with two successive passages about either heat flow or electricity flow through metals. The subjects received as a first passage either an abstract passage (a passage as devoid of concrete referents as was possible), a concrete passage (a passage containing physical analogies for key parts of the passage), or an unrelated control passage, and either a concrete or an abstract passage as a second passage. The results of this series of studies indicated that subjects receiving an initial passage containing physical analogies learned more from the second abstract passage (there was no difference between groups when the second passage was concrete) than did the groups receiving the control or the abstract first passage. One way to interpret these results is to suggest that drawing an analogy between tinker toy models and the molecular structure of metals (one of Royer et al.'s analogies) activates an entire complex of previously learned information, much of which may be very abstract in nature. As an example, when we think of tinker toy models we typically think of regular units bonded together in some fashion with open spaces in between the units. These ideas, are not of course, ones we would regularly associate with a bar of iron. However, these are precisely the kind of ideas which must be grasped if one is to understand why metals are excellent conductors of both heat and electricity.

The cognitive theory I presented earlier is a poor vehicle for explaining the kind of transfer investigated by Royer and his associates.
The theory postulates a rather static memory structure in which activation spreads from a given node (or nodes in the case of parallel search) to other nodes. It is difficult to conceive of how this sort of theory would explain the immediate activation of an entire complex of information which could then be used as a tool for acquiring additional information.

A Schema Theory of Transfer

The theory I will present in this section is a distillation of work which has been going on in an area which has come to be called cognitive science. Since this area consists of a rather diverse set of disciplines (cognitive psychologists, linguists, computer scientists working in artificial intelligence) some of the terminology varies from writer to writer. I will be using the term schema to characterize the basic structure unit in the theory. This usage follows that of Adams and Collins (in press), Bobrow and Norman (1975), Norman (1975), and Rumelhart and Ortony (1977). Others have used the term, "frames," (Charniak, 1975; Minsky, 1975; Winograd, 1975) and "scripts and plans" (Schank & Abelson, 1975) to refer to conceptually similar structures.

Schemata, as I will use the term are of two kinds. The first is an abstract data structure consisting of generic entries for frequently experienced events or concepts. This data structure can be conceptualized as being hierarchical in nature with more specific schemata being embedded in a general schema. As an example, Schank and Abelson (1975) have talked of a "going to the restaurant" schema (they use script and plans instead of schema) which at the most general level would include such information.
as restaurants are places to eat, one pays for the food there, one does not have to cook or clean up at restaurants, etc. At levels below this, one could have embedded schemata which contain more specific information pertaining to ethnic restaurants (e.g., Greek, Italian, etc.), fast food restaurants (McDonald's, Chicken Delight, etc.).

The second kind of schemata contain procedural information. So, for example, we might activate a procedural schema when faced with learning a list of free recall words in a psychology experiment. The procedural schema would activate subschemas having to do with particular strategies, such as active rehearsal, category clustering, etc., for learning the words.

Schema theory suggests that during the learning process a particular schema (or schemata; one could have both a data structure and a procedural schema activated at the same time) is activated and serves as a structure for representing information and as a source of hypotheses about what kind of information to expect. One way to think of a schema is as a structure with a series of slots waiting to be filled by the incoming information. As long as the incoming information matches up to one of the slots in the schema, learning proceeds smoothly and easily. However, in the event that a bit of information is encountered which does not match up to a slot, or even worse, when information is encountered which does not match up to any of the available schema, learning becomes difficult and arduous.

In addition to the previously cited studies by Royer and his associates, there is considerable evidence from other sources which is consistent with
the schema view of learning. One could interpret, for example, the "theme" studies reported by Bransford and Johnson (1972) and by Dooling and Lachman (1971) as being consistent with expectancies from schema theory. One could also interpret the sentence learning studies reported by Begg and Paivio (1969) and by Pezdek and Royer (1974) as being instances of the nonavailability and the availability (respectively) of appropriate schema for interpreting abstract sentences.

In addition to the effects of schema on acquisition, there is also evidence that schemata can have effects on retrieval. Spiro (1977) and Anderson and his associates (Pichert & Anderson, 1977; Anderson & Pichert, in press) have reported studies which claim to show that the recall of previous stored information can be altered as a function of the nature of the schema in use at the time of recall.

When viewed from the perspective of schema theory, learning transfer involves the activation of a previously acquired schema upon encountering the new learning situation. Given that the activated schema is appropriate for the task, learning would occur much more rapidly than it would in the case where an appropriate schema was not available.

An Evaluation of the Schema Theory of Transfer

Schema theory, as it now exists, provides an account of transfer in situations where the previously discussed theories had difficulty. Figural transfer is the clearest example of this extension. Schema theory could easily account for figured transfer by proposing that a schema (consisting of a complex of generic information) acquired from prior experience would
be activated to interpret information from a heretofore unrelated problem. Thus, in the example provided by Royer and his associates the complex of information subjects had about tinker toy models was used to assist in the acquisition of information about the internal structure of metals.

Schema theory also provides a reasonably good account of transfer in those situations involving the utilization of a previously learned skill. If we conceptualize a skill (such as performing addition or long division) as a sequence of activities which could be performed in a variety of situations, we could then represent that skill as a set of procedurals for operating on a data base. This, of course, is one of the kinds of schemata which was mentioned near the beginning of this section of the paper. Vertical transfer then would involve the activation of a previously acquired procedural schema when the problem at hand signals for the use of that schema.

The weakness of schema theory at this point is that it is not really a well formalized theory. The many versions of schema theory, while similar at the conceptual level, are different, and perhaps even contradictory, at the specific level. In addition, the theories as they now exist are most frequently represented as working "models" subject to change. However, despite the relatively tentative stage of its development, schema theory has proven to be a highly useful heuristic for stimulating research and thinking in cognitive science.

Educational Relevance of Schema Theory

Schema theory has emerged so recently that at this point there is no strong evidence that the theory has educational utility. There are
many ways that it might have utility however, and several of these are worthy of empirical investigation.

One example of a place where schema theory might have a practical impact is in the formal analysis and development of teaching by analogy situations. Schema theory suggests that the reason analogies are a useful vehicle for promoting learning of new materials is that they provide a schema (i.e., a data structure) for interpreting the new material and for integrating it into existing knowledge structure. Conceptualizing the situation in this way suggests several ways for approaching the problem.

The first thing that is obvious about teaching by analogy is that in order to be effective, the analogies must be part of the learner's knowledge repertoire. One difficulty with conducting a formal analysis of the teaching by analogy situation has been that there was no way to determine if the knowledge necessary to make analogies effective was possessed by the learner. Conceptualizing the situation as a schema directed process does, however, suggest a way to approach this problem. If we think of schemata as abstract data structures consisting of generic entries for frequently experienced events or concepts, it is apparent that the prior knowledge which is important to assess in order to determine if a teaching analogy will work is not specific in the sense that we are interested in whether certain facts are known. Rather, the important prior knowledge must be context free in the sense that the critical knowledge must be applicable to all instances contained in the referent of the analogy, and in addition, must be applicable to the new material being taught:
The above analysis suggests that a procedure which assessed generic knowledge relevant to a class of entities would be useful as a procedure for determining if a particular analogy will work in a teaching situation. For example, assume that one wanted to use a telephone exchange as a teaching metaphor for how the human body responds to stepping on a hot coal. In order for the analogy to work one would need to know certain generic information about telephone exchanges. One would need to know that there is a source of initiated messages, that there is a mechanism for detecting the message, that there is a medium via which the message is transmitted, that there is a switching device which routes the message to the appropriate receiver, etc. If a list of such information were available one could assess the degree to which the relevant population possessed the information. One way this could be done is to ask learners to identify relevant generic information from a list which contained both relevant and irrelevant items. Based on this data one could then decide if a particular analogy is likely to work.

Another educational issue to which schema theory might be applied is the problem of determining whether students have sufficiently mastered a skill such that the skill can then be used as a tool for acquiring a superordinate skill. Normally this would be done by presenting a variety of problems requiring the use of the subskill and establishing the fact that the student could perform the skill under the variety of conditions calling for its use. Conceptualizing the skill as a schema, however, suggests another possible approach. In theory, it might be possible to develop an
assessment procedure which measured the degree to which a student possessed the abstract or procedural knowledge required to perform the skill given varying problem forms. This might be done in the same manner as was suggested for assessing whether the prior knowledge required to make an analogy work was available to the learner.

The ideas presented above are obviously speculative and empirical work would be necessary to establish their merit, or lack of merit. They are, however, suggestive of ways in which schema theory might seem as a source of ideas for new approaches to educational problems.

Summary

It seems clear that the two classes of theories discussed in this paper have both strengths and weaknesses. The cognitive theories seem to have the most breadth in the sense that they provide an explanation for transfer in a wider variety of situations. Cognitive theories suggest an accounting for transfer in both near and far transfer situations whereas the environmental theories are most applicable to near transfer situations.

The advantage for the cognitive theories in the breadth of the theory is counterbalanced if one considers predictive specificity as a criteria for evaluating theories. The problem with the cognitive theories from the perspective of the predictive specificity criteria is that they explain too much. That is, the theories are not well defined enough to produce theory testing predictions. As a result, one could interpret virtually any behavioral outcome in a manner which seemed consistent with the theory. Such interpretations, however, are post hoc, and do not take the place of predictions derived beforehand from the tenants of a theory.
The environmental theories receive better marks on the predictive specificity criteria. These theories specify situations in which facilitative transfer should occur, and they specify situations in which transfer should not occur (in fact, they also predict instances of inhibitory transfer). Thus, the scientific merit of the theory can be evaluated by comparing predicted with actual outcomes.

The comments above should be considered from the perspective that the theories being discussed are evolving. The cognitive theories described in this paper will surely become more specific. In fact, they already have. One need only compare Bartlett's (1932) description of a schema (he called it an "organized pattern") with recent computer programmable descriptions (e.g., Minsky, 1975; Winograd, 1975) in order to conclude that considerable evolutionary progress has already occurred. Since cognitive theories are very much in vogue there is every reason to believe that they will soon be specific enough to derive testable predictions.

In comparison to cognitive theories, environmental theories have remained curiously frozen. The identical elements theory suggested by Thorndike and Woodworth (1901) provided the essential details of modern versions of the theory. I would predict, however, that the next generation of transfer theories will resemble current environmental theories in the sense that heavy emphasis will be placed on an analysis of the stimulus event. As was mentioned earlier one of the most persistent theoretical puzzles in psychology is Hoffding's problem of how a sensory event comes to be connected to a particular representation in memory. I believe that current research on pattern recognition in visual and auditory perception may soon suggest
an answer to Hoffding's problem, and that this answer will serve as the basis for a new generation of transfer theories.

Cognitive and environmental theories also have strengths and weaknesses in terms of their educational utility. Environmental theories provide guidelines for developing and sequencing related instructional events in order to maximize facilitative transfer. Environmental theories do not, however, provide guidelines for achieving transfer from school learned material to real world events and problems.

Cognitive theories have really not developed to the point where they suggest specific guidelines for educational practice. This means that if one is faced with the problem of facilitating transfer from one school learned task to another it is probably best to look to environmental theories for guidance. Where cognitive theories can be useful, however, is in terms of suggesting procedures for facilitating transfer from school learned skills and knowledge to real world situations. In addition, cognitive theories suggest a framework for viewing the situation where real world knowledge facilitates the acquisition of school material (e.g., teaching by analogy).

If one were to examine educational psychology textbooks published a decade or more ago, one would almost always find at least several pages devoted to the topic of transfer of learning. An examination of recently published textbooks reveals that the topic is covered only cursorily at best. This observation corresponds with another: The vast bulk of educational psychological research in recent years has been concerned with the learning of isolated tasks. It is the belief of this author
that some of this research effort could have been more productively spent on considering the issue of how the learning of one task influences the acquisition of another. Hopefully, this paper will suggest some frameworks from which these types of questions might be asked.
Reference Note

References


Theories of Learning Transfer


Thorndike, E. L., & Woodworth, R. S. The influence of improvement in one mental function upon the efficiency of other functions. Psychological Review, 1901, 8, 247-261.


Footnote

What is not universally shared is Huey's view of the importance of imagery in deriving meaning. While some psychologists share Huey's view on this, many do not.
Theories of Learning Transfer

46

Figure Captions

Figure 1. Hypothetical knowledge representation after learning without application questions.

Figure 2. Hypothetical knowledge representation after learning with application questions.

No. 2: Asher, S. R. Sex Differences in Reading Achievement, October 1977.


No. 4: Jenkins, J. R., & Pany, D. Teaching Reading Comprehension in the Middle Grades, January 1978.
* Available only through ERIC

**No. 1:** Halff, H. M. Graphical Evaluation of Hierarchical Clustering Schemes, October 1975. (ERIC Document Reproduction Service No. ED 134 926, 11p., HC-$1.67, MF-$0.83)

**No. 2:** Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse, October 1975. (ERIC Document Reproduction Service No. ED 136 187, 81p., HC-$4.67, MF-$0.83)

**No. 3:** Goetz, E. T. Sentences in Lists and in Connected Discourse, November 1975. (ERIC Document Reproduction Service No. ED 134 927, 75p., HC-$3.50, MF-$0.83)

**No. 4:** Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management, November 1975. (ERIC Document Reproduction Service No. ED 134 928, 21p., HC-$1.67, MF-$0.83)

**No. 5:** Schallert, D. L. Improving Memory for Prose: The Relationship Between Depth of Processing and Context, November 1975. (ERIC Document Reproduction Service No. ED 134 929, 37p., HC-$2.06, MF-$0.83)

**No. 6:** Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. Two Faces of the Conceptual Peg Hypothesis, January 1976. (ERIC Document Reproduction Service No. ED 134 930, 29p., HC-$2.06, MF-$0.83)

**No. 7:** Ortony, A. Names, Descriptions, and Pragmatics, February 1976. (ERIC Document Reproduction Service No. ED 134 931, 25p., HC-$1.67, MF-$0.83)

**No. 8:** Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976. (Journal of Educational Psychology, 1977, 69, 288-297)

**No. 9:** Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education, April 1976. (ERIC Document Reproduction Service No. ED 134 932, 42p., HC-$2.06, MF-$0.83)


**No. 11:** Armbruster, B. B. Learning Principles from Prose: A Cognitive Approach Based on Schema Theory, July 1976. (ERIC Document Reproduction Service No. ED 134 934, 48p., HC-$2.06, MF-$0.83)


No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976. (ERIC Document Reproduction Service No. ED 134 938, 24p., HC-$1.67, MF-$0.83)


No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977. (ERIC Document Reproduction Service No. ED 134 941, 76p., HC-$4.67, MF-$0.83)


No. 28: Ortony, A. Remembering and Understanding Jabberwocky and Small-Talk, March 1977. (ERIC Document Reproduction Service No. ED 137 753, 36 p., HC-$2.06, MF-$0.83)


No. 34: Bruce, B. C. Plans and Social Actions, April 1977.


No. 54: Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and Decontextualized Practice Conditions on Word Recognition, July 1977.


No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers’ Developing Conceptions of Print, September 1977.


No. 77: Nash-Webber, B. L. Inference in an Approach to Discourse Anaphora, January 1978.

