The Effect of Context on Training: Is Learning Situated?*

Lynne Reder ‡ and Roberta L. Klatzky ‡

13 September 1994
CMU-CS-94-187

School of Computer Science
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213-3890
(‡Department of Psychology)

Also appears as:
Human-Computer Interaction Institute Technical Report CMU-HCII-94-108

Abstract

This paper addresses issues associated with transfer of training; in particular, the importance of mimicking the contexts of application during training: When will a skill transfer from one situation of application to another, and what aspects of training affect a learner's ability to transfer the learning to new situations? Although there exist many examples of failure to show good transfer of training, the failures are more surprising and tend to mask the successes. With appropriate training there is generalization to a variety of contexts. Such training should include variable contexts during instruction, concrete examples, and abstract rules and procedures. Claims made by proponents of Situated Learning Theory suggest that training must be done in the context of the situation of application. We agree that learning is unlikely to transfer if taught in only one context with no examples from other contexts; however, because it is difficult to anticipate all future contexts of application, trainers are advised not to so constrain instruction. Although training and transfer should have identical elements, there is evidence that part-whole training is often the appropriate method of instruction. This argues against apprenticeship learning during early stages of acquisition for many skills. Further, too much fidelity in simulation may lack cost effectiveness and may even be detrimental to the early acquisition process.

* This report is a slightly modified version of chapter 3: Transfer: Training for Performance, in D. Druckman and R.A. Bjork (Eds.), *Learning, Remembering, Believing: Enhancing Human Performance*, pp. 25-56. Washington, D.C.: National Academy Press.

Research on this chapter was supported by a grant from the U. S. Army Research Institute to the National Research Council and by Grant BNS-8908030 from the National Science Foundation to Lynne M. Reder.

Keywords: transfer, training, Situated, Learning, context, Apprenticeship, identical elements, abstractions, instruction, performance, part-whole training.

Introduction

Probably the most critical issue in any type of learning is how well the learning transfers from one situation to another. Although there is a broad consensus that transfer is an important aspect of learning, training, and performance, it is not always clear what is meant by transfer or how to achieve it. In this report, we focus particularly on situations where there is some period of training prior to on-the-job execution of a task. We also focus on the individual within the training context. That context may or may not be a group setting.

The ability to transfer between the training and application contexts is the crux of the frequently made distinction between learning and performance. One may learn to perform a task quite well during training, according to some criterion, but later find that the acquired knowledge is not sufficient to perform in the day-to-day task environment.

The distinction between learning and performance is made because, more often than not in everyday training situations, the training and task contexts differ in some way. The differences may involve situational characteristics external to the task per se, such as social interaction patterns, stress or ambient noise; they may involve explicit or implicit rules that directly govern task performance, the range of variation in the stimulus environment, the nature of available responses, performance schedules, or characteristics of the performer such as levels of motivation, fatigue, or stress. Indeed, training cannot generally anticipate the full range of circumstances that will be encountered subsequently, and even an anticipated circumstance may be impossible to fully simulate in training. A training program should ideally produce the ability to accommodate some degree of variability-- within the task environment and between the training and task environments--as well as establish basic skills required for the task itself.

The principal question addressed in this report is: How much does the training context have to incorporate the performance context in order to produce effective transfer? This concern applies to training on math problems at school, flying a plane, training for combat, controlling airplane traffic at the airport, playing a championship tennis match, or operating a nuclear power plant. It is a fundamental issue in the design of simulator-based training devices, where one asks how close in fidelity a device must be to the performance situation, in order to be effective. Simulating unnecessary aspects of the performance context means that resources have been wasted, but failing to simulate necessary ones means that training will be inadequate. The fundamental problem underlying the issue of specificity of training, or fidelity of simulation in training, is determining the factors that produce transfer. Only when those factors are understood can we balance resources against outcomes to design optimal training programs.

An issue related to training fidelity is the relative importance of training individuals on abstract generalizations, as opposed to specific, contextualized examples. Abstractions may or may not be important, and even if they are, it is arguable whether they are best taught directly (e.g., by the instructor's articulating rules) or whether they are best inferred from examples. This report will discuss a strong position in regard to this issue, which has been called the Theory of Situated Learning.

Implicit in our comparisons between situated learning and other approaches to learning are certain basic constructs concerning transfer. Positive transfer refers to the facilitation, in learning or performance, of a new task based on what has been learned during a previous one. Negative transfer refers to any decline in learning or performance of a second task, due to learning a previous one. These types of transfer are often measured as a percent savings or loss, respectively: How fast (or accurately) is the target task acquired after learning a similar (transfer) task, as compared with learning without a prior task? Sometimes transfer occurs from a component of one task to a more complex task that encompasses that component; this is called part-to-whole or vertical transfer. In contrast, horizontal transfer can occur between tasks similar in complexity that do not have an inclusion relation.

1. Transfer by Identical Elements

As we have said, a major issue in developing training programs is the required similarity between the training and transfer contexts. At the limit, transfer should be perfect if the training context exactly simulates the performance context, but this is relatively rare in applications. Theories of transfer have stressed the requirement of learning/performance similarities. In this section, we review such theories.

Thorndike and Woodworth's theory of "identical elements," published in 1901, stated that the determinant of transfer was the extent to which two tasks contain identical elements: The more shared elements, the more similar the two tasks and the more transfer there would be. This position was in stark contrast to the long-standing view that the condition of one's mental faculties accounts for transfer. Thorndike rejected the view that the mind is a muscle that must be strengthened with good exercises such as the study of topics like Latin and geometry, and that with such rigorous studies, transfer between any two fields would be straightforward. The problem with Thorndike's approach is that it is unclear exactly what defined the identical elements. There was some suggestion that he meant mental elements, although his theory was typically interpreted to mean stimulus-response connections.

Identifying the elements that should be identical, in order to produce transfer, is critically important. If one first observes transfer and subsequently infers what the identical elements must have been to produce it, then one's reasoning is obviously circular. But advance specification of identical elements is more difficult than it might appear. Some situations that seem to have substantial identical elements produce little or no transfer, and some that do not seem to have similarities produce a substantial amount. Here we give some examples.

> Case 1: Singley and Anderson (1989) taught people three different computer text editors, varying the order of acquisition of the three editors across subjects. They developed a set of 107 rules capable of simulating editing in these three editors. Some of the rules were shared by all three editors, some by two, and some were unique to a specific editor. A given editing task using a particular editor might or might not share many rules with the comparable task using a different editor. Their expectation was that there would be savings in performance for the second editor to the degree that the rules had already been acquired. The data were in strong support of their predictions.
>
> Case 2: MacKay (1982) reported data from a group of English/German bilinguals who were asked to produce the same sentence in the same language 12 times, with a 20-sec pause between utterances. They were to produce as quickly as possible. The production time declined regularly over the 12 repetitions, to an asymptote of about 2 seconds. In the last 20-sec pause interval, each subject was asked to produce the same sentence, but in the alternate language. The speed of production for this transfer sentence was found to be identical to that of the trial before. That is, the subjects remained at asymptote and the new sentence functioned as if it were identical to those previous, despite a complete change in the motor movements necessary for output.
>
> Case 3: Logan and Klapp (1992) had subjects solve alphabetic arithmetic problems (e.g., if A = 1, B = 2, etc., does A + 2 = D?) with one set of 10 letter addends and the digits 2-5 for 12 sessions of nearly 500 trials each. Initially, their response times increased markedly with the digit addend (e.g., there was a longer time for B + 5 than for B + 2 problems), as if they were moving forward through the alphabet from the given letter for the required number of digits (e.g., for B + 2: B, C, D). The slope of this increasing function was 486 ms per count in the first session. By the 12th session, however, the slope was only 45 ms, suggesting development of a new strategy. The subjects were then transferred on session 13 to a new set of 10 letters. Although the task remained unchanged, the slope of the function relating response time to digit addend increased dramatically -- to nearly the value it had held in the initial sessions. Thus there was only a small amount of transfer.

The situations we have described represent very different degrees of transfer: Partial transfer, depending on shared rules; virtually perfect transfer despite apparently substantial differences between the training and transfer contexts; and virtually no transfer from one problem to another. How can we understand these situations? The answer lies in what elements are and are not identical. Let us consider the theories behind the experiments.

1.1 Identical Elements at the Level of Cognitive Abstractions

Singley and Anderson (1989) and Bovair, Kieras and Polson (1990) have developed formal theories that specify what are the "elements" of cognitive tasks. They then show that holding these elements identical across learning situations predicts positive transfer. At the heart of these theories is the idea that transfer is produced when cognitive abstractions that are formed in one context--rules or knowledge chunks--can be used in another. The models of Polson and Kieras (1985), Bovair et al, and Singley and Anderson are similar to Anderson's (1983) ACT model, so ACT will be described as an illustration. (See Gray and Orasanu, 1987, for an excellent review of work on skill transfer explained within this framework.)

The theory assumes that there are two types of memory, a declarative or fact memory and a procedural or skill memory. Many of the tasks that we perform are originally encoded in memory in our declarative memory as verbalizable rules. As they become practiced and strengthened, these declarative facts are often compiled into procedural rules that are executed automatically and are not open to inspection. So, for example, when one first learns to drive a stick shift in an automobile, one says implicitly to oneself: "Push in the clutch with the left foot, lift up the right foot to release the gas pedal and use the right hand to move the stick shift to another position." As this skill becomes compiled into an automatic procedure, one might not be aware of exactly what one's feet are doing while shifting (Anderson, 1976).

Declarative facts are stored in a type of semantic network of nodes or chunks that are connected together via associations acquired though experience. These nodes or chunks may vary in strength as will the associations among them, depending on the amount of exposure (practice) they have received. Performance will vary as a function of the amount of activation that any particular memory structure receives. Activation of a structure depends on the number of competing associations linked to a node and the relative strength of these associations and nodes.

Procedural memories are condition-action rules. An action is specified by the right hand side of a production and it is executed when all of the elements of the condition side (lefthand side) are met, i.e., match the contents of working memory. When the condition side of a production is matched, and thus the action side is executed (the production "fires"), the production is also strengthened. When more than one production could fire, because the condition side of multiple productions could match to the situation, the selection of a single production to fire (or conflict resolution as it is called) is determined by the specificity and the strength of the competing productions. Specificity refers to the number of condition elements that must match on the left hand side. More specific (or complex) productions are preferred to more general productions, all else being equal.

The action of a production can also vary in complexity. For example, it can set up new goals to be achieved, place new elements in working memory, or initiate a simple response. The action of setting up a new goal may enable the conditions of another production to be met.

According to theories such as ACT, transfer between tasks is achieved when the tasks share elements, either chunks in declarative memory, or productions in procedural memory. Acquiring and strengthening declarative memory chunks and productions for one task will facilitate performance of the other task to the extent that these memory chunks and productions are shared. Anderson (in press) suggests that there are other means by which one experience can influence another, such as analogy--extending knowledge from one situation to a new situation that is only similar.

MacKay (1982) has also emphasized transfer based on identical elements at abstract levels, focusing particularly on the domain of motor performance. His model makes hierarchical distinctions between high-level structures representing abstract rules for a task and low-level structures representing commands to peripheral mechanisms that will directly produce performance. In the speech domain, for example, the hierarchy comprises distinct structures for meaning, sound, and control of the speech musculature.

Transfer occurs in this model between tasks that share mental structures. Practice on one task will then strengthen linkages that are necessary for the other. The model actually predicts perfect transfer for one situation: when the unshared structures are so extensively practiced already that their linkage strengths are at asymptotic levels and could not benefit further. This is the case for transfer between a sentence in German and its translation in English. The idea is that the shared structures between these tasks are conceptual ones that have not been previously practiced (since the phrase is novel), and the unshared structures are phonological and muscle-movement ones that are well practiced in both languages and hence would not benefit from further practice. Transfer at the conceptual level is all, then, that is needed. Given that the two languages share a single conceptual system, perfect transfer will occur, because practicing the phrase in one language is equivalent at the conceptual level to practicing it in the other.

Transfer will not be complete, however, if the unshared structures between two tasks are not well practiced. This is the case, for example, when one attempts to write with the nondominant hand. Only the abstract level of structure pertaining to the written words is shared in this case. The muscle-movement structures for the non-dominant hand are not well practiced, and being unshared, they do not benefit from practice with the dominant ones. One would expect minimal transfer from writing one's own name with the dominant hand to writing with the non-dominant, because in this case the shared mental structure--the name--is already well practiced and cannot benefit further from the commonalities between learning and performance.

1.2 Identical Elements as Exemplars

A theory of Logan (1988) makes the fundamental assumption that people performing in a task store instances of past performance in memory. Further, each instance is assumed to be stored as an independent copy or "exemplar." On their first encounter with the task, having no stored instance, they will use whatever strategic, rule-based tools they have available; this constitutes a task "algorithm." Subsequently, however, they will have available not only the algorithm, but memory for the past instance of performance--potentially many instances, according to the number of times the task has been performed. When the task recurs, performance is based on the first solution that is retrieved from memory, which will be the algorithm or a retrieved instance. The time to retrieve each past instance is assumed to vary stochastically, so that the algorithm competes for retrieval with that instance having the fastest current retrieval time, that is, the lowest value drawn from a set with similar distributions, one for each instance that has been stored. With enough stored instances, the algorithm will tend not to be retrieved faster than the fastest instance, even if its mean retrieval time is the fastest. Responses will come to rely virtually exclusively on past instances. Automaticity, according to the instance model, corresponds to this shift from algorithmic to instance-based retrieval.

A principal assumption of the instance model is that if a previously solved task is presented and attended to, past instances of that same task will be retrieved from memory. The emphasis is on "that same task." Under the assumption that the subject comes to rely on specific stored instances, the task here constitutes not only the general type of problem that must be solved but the specific parameters that are provided. Learning is item-specific, and retrieval is of the same items that were previously used in the task. Transfer between distinct items within the same task is assumed not to occur, because presentation of a novel item will not lead to retrieval of an item used previously in training. This explains why so little transfer was obtained in the Logan and Klapp experiment described above, when the specific letter addends changed. What little did occur was attributed to better learning of an algorithmic solution (e.g., with practice, subjects became able to count faster from the letter addend).

It should be noted that Logan emphasizes tasks in which conceptual structures are minimally important and instance retrieval provides an effective solution. Without identical elements at the instance level, there is then little transfer. Singley and Anderson, and MacKay as well, emphasize tasks in which conceptual structures are very important and responses are either extremely well practiced or minimally taxing parts of the task. Hence non-identical elements at the output level will matter little. We will return to the issue of how transfer may be governed by different principles in different tasks in a later section.

Logan's experiment clearly supports Thorndike's view that the mind is not like a muscle. Subjects had ample "exercise" with the mental arithmetic task, but without identical problems, there was virtually no transfer in the task.

2. Situated Learning

Recently there has been a somewhat different approach to the idea of identical elements, one that emphasizes the context in which learning and performing occur. This general approach has come to be called Situated Learning (e.g., Lave, 1988 ; Lave & Wenger, 1991), although closely associated terms are situated cognition and situated action. The situated approach is a somewhat loose alliance of researchers from several fields, including psychology, anthropology, and philosophy. Its fundamental tenets include an emphasis on contextual determinants of performance, particularly on social interactions in the task environment, and on the importance of situating the learner in the context of application, as in apprenticeship learning, called "Legitimate Peripheral Participation" by Lave and Wenger (1991). Part of this view is that learning is fundamentally a social activity.

Although proponents of situated learning might differ in their views to an extent, certain general principles can be abstracted that seem to characterize the approach, as follows:

> (1) Action is grounded in the concrete situation in which it occurs. The potentialities for action cannot fully be described independently of the specific situation, and a person's task-relevant knowledge is specific to the situation in which the task has been performed.
>
> (2) Psychological models of the performer in terms of abstract information structures and processes, are inadequate, if not inappropriate, to describe performance. A task is not accomplished by the rule-based manipulation of mediating symbolic representations. Certain task-governing elements are present only in situations, not representations.
>
> (3) Training by abstraction is of little use; learning occurs by doing. Because current performance will be facilitated to the degree that the context more closely matches prior experience, the most effective training is to act in an apprenticeship relation to others in the performance situation.
>
> (4) Performance environments tend to be social in nature. To understand performance, it is necessary to understand the social situation in which it occurs, including the way in which social interactions enter into performance.

Situated learning has become a major theoretical framework that is hotly debated by those concerned with education and training. There is little disagreement among cognitive scientists that there exist contextual effects in learning, transfer and retention. Indeed it is not news that performance is affected by context, as has been amply demonstrated in many domains. In the area of verbal learning, for example, the principle of encoding specificity holds that retrieval of learned information directly depends on similarities between the retrieval and learning contexts (Thomson & Tulving, 1971). In applied research based on this principle, Godden and Baddeley (1975) demonstrated remarkably strong encoding-specificity effects when deep-sea divers learned material on land or underwater and were tested in matching or non-matching situations.

What makes the situated-learning approach relatively unique is the extreme degree to which learning is claimed to be context specific and the implications of this claim for education and training. For example, Greeno, Smith and Moore (1993) state (p. 99) that "knowledge--perhaps better called knowing--is not an invariant property of an individual, something that he or she has in any situation. Instead, knowing is a property that is relative to situations...(just as) motion is not a property of an object". An important implication is the idea that school is just another context and therefore that what one has learned in school can only be used there. Lave and others reject the premise that school constitutes a neutral ground apart from the real world in which things are learned that are later applied in the real world. Brown, Collins and Duguid (1988) argue that success with schooling has little bearing on performance elsewhere. So, for example, rather than being taught mathematics as an abstract skill, a person should learn the mathematical techniques relevant for his or her trade right in the situations where they will be needed. Moreover, some theorists argue, people tend not to remember the skills nor are they able to apply them if they are taught in an abstract manner. Only when they are learned "on the job," embedded in the appropriate situation, can the skills be used in those situations.

Another important aspect of the situated learning approach is the view that cognition cannot be represented symbolically, and that instead of representing the environment, people perceive it directly and use it to support thought. Stucky (1990) argues that rather than representing relevant aspects of the situation, people use contextual cues directly to calculate their actions. Following the general approach of ecological psychologists, Greeno et al. (1993) similarly suggest that performers perceive physical properties of situations that make certain activities possible. For example, they find it unlikely that children have symbolic representations of physical properties of objects, such as the flatness of a surface; rather, they propose, children directly perceive flatness and its affordance of operations like stacking. Transfer from one environment to another then depends on common affordances that produce invariance of interaction.

3. General Principles of Transfer: An Evaluation

A general principle of transfer seems to be that identical elements are necessary. One specification of those elements that has proven successful is in terms of production rules. But how much identity is necessary? Must learning be situated in the transfer context in order to be effective? At this point we move to a discussion of general proposals related to the broader question. Each of these proposals is listed as a question below. Some of these proposals appear to be correct generally, others hold only for certain contexts, and still others appear to be generally incorrect.

3.1 Is Transfer of Training Difficult to Achieve?

The cases of successful transfer discussed in Section 2, along with much of the literature review to follow, make clear that transfer is not invariably difficult to achieve. However, there is also a surprising amount of data testifying to the difficulties of achieving transfer. In such cases, performance often seems to be over-embedded in the training context, so that the identity of elements across contexts fails to take effect.

The failure of learners to recognize and capitalize on identical elements between training and task contexts has been well documented (see Patrick, 1992, and as reviewed by Chipman, Segal & Glaser, 1985, and Segal, Chipman & Glaser, 1985). Hayes and Simon (1977), for example, showed little transfer between two structurally or formally identical problems that differed only in superficial characteristics. Gick and Holyoak (1983) used a task in which a memorized story should help, by analogy, to determine the solution of a subsequent problem. Analogical transfer was rare, however, unless a hint was given.

Charney and Reder (1987) proposed that an important component of cognitive skill acquisition, such as learning to use a personal computer operating system or an electronic spreadsheet, is the ability to recognize the situations or conditions under which a particular procedure should be used. If training always makes obvious which procedure should be practiced, then an important element of the skill--the ability to recognize when to use the procedure--is not trained.

Theories based on production rules describe lack of transfer as being caused by production rules that are too specific in the condition elements contained in the IF part of an IF-THEN rule. That is, people do not realize that an old rule can be applied in a new context. Anderson's (1983) model of skill acquisition allowed for generalization to occur by reinforcing the more general (less specific) productions when there was variable practice. Unfortunately, there was little evidence that people actually generalized from variable practice to novel situations, if these new situations were very different from the example or training problems. It seems that generalizations have to be explicitly encoded, either through conscious discovery by the learner who encounters multiple scenarios and notices the similarities or from explicit instruction.

Transfer failure also occurs not because the context or situation is very different, but the task has dramatically changed. For example, Knerr, Morrison, Mumaw, Stein, Sticha, Hoffman, Buede, and Holding (1987) reported that a group trained to recognize correct flight paths was superior in recognizing correct flight paths, but not in producing them. McKendree and Anderson (1987) and Kessler (1988) found that being trained to evaluate LISP functions facilitated performance on evaluation of other LISP functions (compared to a control), but not on generating LISP functions, and vice versa.

In theory, when productions are formed initially they are very condition and action specific. The specificity of these rules would go unnoticed in everyday situations because typically one would not practice only half of a skill, such as learning to evaluate LISP functions but not to generate them. In order to achieve transfer the learner must make a conceptual generalization of actions. For example, if a learner is used to using a command-based computer and then moves to one that is menu-driven, successful transfer requires that the learner be able to map old actions (such as typing in a string of letters in order to rename a file) into new ones (such as calling up a menu dealing with file manipulation and clicking on a "rename" option). This type of successful transfer occurs frequently; however, when people accomplish the mapping (the transfer), the success is taken for granted. Only when the attempted transfer fails is it considered remarkable.

3.2 Is Transfer Optimal if Training is Situated in the Context of Application?

It cannot be contested that one of the most important goals of training and education is improving teaching techniques in such a way that the application of a skill in new situations is easily achieved. An implication that might be taken from the failures of transfer just described is that training must be situated in the target context in order to be effective. This is of course a fundamental premise of the situated learning approach. And in support of the premise, there exist many illustrations of superior performance with more specific, relevant practice. Because it is difficult to anticipate all of the features of the target context that will be important to learn during training, it would seem appropriate to train in the target context to insure that the relevant features are trained.

However, a strong position that basic training is a waste of time, because learning must be situated, is certainly false. There are many domains in which fundamental skills are critical to acquire before more specific training can occur. Further, there are certainly basic skills acquired in school that transfer easily and are heavily used in situations outside the classroom. Reading and writing are obvious examples.

Conceivably, there are individual differences among learners for their need for concrete, motivating "real-life" contexts. Consider mathematics learning, for example. The view that learning of mathematics will necessarily be better if it is learned in real-world contexts is of course debatable. We suspect that at the upper end of the distribution of math ability, the insistence on embedding math in real-life contexts takes away valuable time from the actual acquisition and practice of the underlying procedures.

The proposal that training in the target context is optimal led the advocates of situated learning to suggest that the best form of learning involves an apprenticeship in the real-world context where the training is to be applied. However, in evaluating apprenticeship training, it is important to consider costs external to training per se. To the extent that the presence of an apprentice disrupts performance in the work context, the output of the system as a whole may be diminished. Indeed, there are cases where it would be inconceivable to place a novice directly in the work context. Playing in an orchestra, co-piloting an airplane, or serving in a tank crew are just a few of the examples that come to mind.

It is also unclear how one graduates from being an apprentice in a real world task to an actual participant who performs the task competently. If learning is so contextualized that one must serve as an apprentice to learn, how does one actually acquire the skills of the master if one has only been the apprentice? Current theories that emphasize situated learning do not adequately deal with the transition from apprenticeship to mastery.

Yet another problem with training in the target context is the potential for variability within that context. One cannot always anticipate the future contexts that will be required of the learner. Therefore the instructional goal should be to devise a training procedure that will optimize performance in various contexts that may or may not be anticipated.

Training in the context of application may be unfeasible for economic reasons as well. The target context may limit the number of trainees that can be handled at one time, or it may involve costly equipment or supplies.

For a variety of reasons, then, a training environment that simulates the relevant features of the task might be a more feasible and appropriate technique for many tasks than fully situated learning. Since 1982, soldiers have been trained in simulated battles at the National Training Center at Ft. Irwin (see Wiering, 1992). Over the course of a 4-week rotation at the center, trainees become demonstrably better at fighting simulated battles there. There is reason to believe that this training, using real armoured vehicles, negotiating in real deserts, improved performance in the transfer task of fighting real desert battles in the Middle East. Simulation will be discussed further below.

One point that can be well taken from situated-learning theorists is the stress on evaluating the context of performance-particularly the social context--when designing a training regimen. The commissioned report on the Los Angeles Police Department following the 1991 riots mentions the common statement to police officers newly placed on the beat: "Forget everything you learned at the [Police] Academy" (Independent Commission on the Los Angeles Police Department, 1991, p. 125). The implication is that the social milieu of the working officer has little to do with the training environment. Although this is an extreme example, there are social-interaction characteristics of most workplaces, and training may often ignore them. A workplace analysis may reveal these characteristics and motivate their being dealt with during training, to the extent that is possible.

Several questions arise from consideration of the claim that situated learning is necessary to produce transfer. Would an abstract form of instruction facilitate, to any degree, transfer from one task context to another? Does learning in one context become so situated that it cannot be generalized to other contexts? Below we address the issue of the relative merits of abstract instruction that generalizes across contexts, specific instruction in one context, or varied instruction in several contexts with no abstract instruction presented.

3.3 Does Traditional Training on Abstract Concepts and Rules Facilitate Transfer?

A position of some advocates of situated learning is that skills to be learned should not be taught in an abstract fashion. One of the influences of this approach was Whitehead's (1929) inert knowledge problem, which points to knowledge that can be recalled when specifically asked for, in something like a school setting, but is not used spontaneously when needed in actual problem solving, where the knowledge could be applied.

Lave (1988; Lave & Wenger, 1991) is one of the most outspoken proponents of this position. She has made a distinction between "indoor" research and "outdoor" or real-world research and has claimed that school-learned (indoor) algorithms are not the procedures used in the real world. For example, mathematics training in school is said to be irrelevant to mathematical performance on the job in later life and in other real-world situations. In this view, the educational system provides minimal preparation for real-world problems.

Greeno (1989) has also argued that symbolically represented knowledge does not translate well into useful skills. Moore and Greeno (1991) proposed that symbolic knowledge domains such as physics and math should be taught using physical models rather than using symbolic formulas and algorithms. The strong version of this position would hold that mathematics should not be taught using formulas and algorithms because they are abstractions and not tied to physical models of the world. As stated in an earlier section, proponents of situated learning tend to emphasize the importance of knowledge embedded in the situation and to denigrate the role of manipulating internal symbols in accomplishing a task. One problem with such a strong claim is that it fails to explain cases where the abstract concept is developed before the real-world model. For a classic example, Riemann developed his non-Euclidean geometry as an abstract theory much before Einstein used it to provide an explanation of the universe based on General Relativity.

Greeno, Smith and Moore (1993) propose a more general view that transfer can occur to the extent that the situations share affordances for interaction, and that affordances can be perceived without there being a symbolic representation of the properties that specify the affordances. In the example given above, children can directly perceive flatness of a surface without having a symbolic representation of that property. Vera and Simon (1993, pp. 22-23) respond by noting that "Greeno argues that physical models having component objects that correspond closely with those found in real situations are better pedagogical tools than symbolic formulas and algorithms. Does this argument imply that symbolic knowledge does not underlie the central processes of ordinary everyday cognition? We think not."

There are numerous examples where abstract instruction has been shown to be superior to concrete instruction when the transfer task is not very similar to the original training situation. Singley and Anderson (1989) showed this type of result in the context of learning to solve algebra word problems. They presented students with either concrete or abstract tabular representations of "mixture" word problems, e.g., a coin problem in which pennies and nickels are mixed together to yield some total amount. By appropriately labeling a tabular representation, this problem could be treated as specific to coins (using labels of "penny," "nickel," or "total value") or as involving, more abstractly, the combination of parts into wholes (using labels of "part", "part," and "whole."). The abstract labels could be applied to a wide range of mixture problems, involving mixing two solutions together, selling different kinds of sandwiches at a picnic, gambling with a certain amount of money won or lost with each bet, and so on. Singley and Anderson found that, with the same amount of training (number of problems solved), those students who got concrete labels did better on problems of the same type but did worse on mixture problems of different types.

This near transfer/far transfer interaction is reminiscent of much earlier Gestalt research that showed that it is sometimes easier to learn a rote procedure, but the procedure only applies in very limited contexts. In a classic study, for example, Katona (1940) showed that in problems requiring that matchstick shapes be altered to make new ones, subjects who memorized the required moves outperformed subjects who tried to discover general principles -- but only on the originally learned tasks. On transfer to new problems, the discovery group excelled.

Further evidence for the usefulness of teaching abstract concepts to facilitate transfer has been shown in a number of subsequent studies. Mayer and Greeno (1972) found that subjects who were taught to solve binomial probability problems with a formula outperformed those taught to induce the formula from principles and examples, but only on solving problems similar to those encountered during training. The induction group showed better comprehension of the formula and better ability to recognize problems that could not be solved. Similarly, Singley (1986) explicitly taught subjects the abstract goal structure for solving related-rates calculus word problems and found that this led to faster learning and more transfer to new types of problems. Klahr and Carver (1988) found that when children learned how to debug programs in the Logo computer language, they performed better when the high-level goal structure for debugging was explicitly presented in its abstract form.

A somewhat different example of transfer that was facilitated by abstract instruction concerns learning to throw darts to hit a target under water (Scholckow & Judd, as described by Judd, 1908). If one were to aim directly at the target under water, the dart would go beyond the actual location of the target, due to the misleading cues from the refraction of light. One group of elementary (5th and 6th graders) students received an explanation of light refraction, while the other group did not. On the initial task, where the depth of water was 12 inches, both groups performed about equally well; however, when the water level was lowered to 4 inches, the difference between groups became striking. Those without the abstract instruction were confused and made large and persistent errors, whereas those who received the abstract instruction corrected their aim quickly. When the depth of water was changed once more to 8 inches, the abstract-instruction group did better once again. A conceptual replication by Hendrickson and Schroeder (1941) used two levels of abstract explanation as well as a control group and found more transfer with more abstract instruction.

A seemingly visual skill, pattern identification, can also be aided by abstract instruction. Biederman and Shiffrar (1987) found that novice subjects could be taught to determine the sex of pictured chickens at the level of experts, merely by receiving instructions as to the location and qualitative shape (convex, concave, or flat) of a critical feature distinguishing males from females. This skill had previously been thought to require many hours of visual training. (Note, however, that the experimental task involved sex-typing photographs of chickens, while an additional part of the expert's skill involves knowing how to pick up a chicken to sex it.)

Despite the studies mentioned above that provide support for the benefits of abstract instruction, it is important to emphasize that abstract instruction in the absence of concrete examples is almost always a bad idea. It is well understood among educational and cognitive psychologists that concrete examples are important to facilitate appropriate use of acquired knowledge (Simon, 1980). Sandberg and Wielinga (1992) point out that the real problem has always been how to design teaching methods that teach both the declarative subject matter and its use. Anchored Instruction (Cognition and Technology Group at Vanderbilt, 1990) provides an approach to education that creates learning experiences in a school setting that have some of the properties thought to be important in apprenticeship training.

3.4 Must Task-irrelevant Elements be Duplicated for Transfer to Occur?

This question has been addressed primarily in the domain of verbal learning, where the issue is whether changes in the environmental context from learning to remembering affect the amount that is remembered. A prototypical experiment would have subjects learn in a room with walls of one color, then recall in a room with walls of the same or a different color. If recall is superior in the same-color environment, a positive effect would be said to occur. Environmental context effects have also been addressed by asking whether experiencing multiple environments during learning helps people to remember (a paradigm that is similar to that used to study variable practice effects; see below).

It is important to realize (Bjork & Richardson-Klavehn, 1989) that this issue refers to the effects of incidental context. That is defined as features of the learning environment that are not part of the to-be-learned material itself and that should not directly affect how people deal with that material. For example, the presence of posters citing cancer statistics in a room where people receive an anti-smoking lecture might seem task-irrelevant but is likely to be influential context rather than incidental.

The literature on environmental context effects in verbal learning has a somewhat checkered history. Sometimes effects have been found, and sometimes they have not. There have also been failures to replicate seemingly robust cases of environmental influences (e.g., Fernandez & Glenberg, 1985). We can better understand this inconsistency in outcomes if we consider two factors that have been suggested to modulate the degree of environmental influence. They are (a) the availability of memory cues other than the environment, and (b) the extent to which the subject tries to recollect the original context at the time of the memory test.

According to the "outshining hypothesis" (Smith, 1988; Smith & Vela, 1986), environmental influences will be reduced (outshone) when there are strong retrieval cues present at the time of the memory test. This can occur, for example, because the test itself provides strong cues (as does a recognition situation) or because the original learning situation promoted self cueing (as would an instruction at the time of learning that induced subjects to think about items conceptually and relate them to one another, e.g., forming words into a story). Bjork and Richardson-Klavehn suggested that this hypothesis should be augmented by a "reinstatement hypothesis," which states that physically reinstating the context of learning will be useful only when subjects cannot mentally do so for themselves. For example, if I studied in a pink room but am tested in a green one, imagining the green walls to be pink may compensate for the change in environment (e.g., Smith, 1979). According to these hypotheses, then, variations in the nature of the retrieval and learning situations may underlie the inconsistent effects of environmental context that have been observed.

In a meta-analysis of the literature on environmental context, Vela (1989) found support for the outshining hypothesis. Analyzing over 50 studies and calculating the size of the context effect (defined as the performance difference between no-context-change and context-change groups, normalized by the standard deviation), Vela found overall a moderate, statistically significant advantage for preserving the context at test. This was modulated by both the nature of the retrieval test and the conditions of study, as predicted by the outshining hypotheses.

Extrapolating from these studies to the more general topic of transfer, we could infer that changes in seemingly irrelevant aspects of task context between training and performance can be detrimental. The effects are not always robust and can apparently be reduced by providing strong cues to performance in the transfer context and by motivating strong cueing from relevant aspects of the task (as opposed to incidental environmental features) during the training procedure. It would be useful to verify these conclusions in a broader range of performance domains.

3.5 Should Training Mimic the Experiences that are Anticipated?

It is interesting that the traditional literature on transfer gives many examples where a situation that does not closely mirror the target task is considered optimal. Two types of data are particularly relevant, dealing with superiority of transfer after variable practice relative to fixed practice, and with transfer to the whole after training on parts.

Variable Practice. Few tasks that people learn to perform are totally repetitive from one occurrence to the next, regardless of whether we consider tasks like flying a fighter plane in combat or seemingly invariant tasks like performing the broad jump. Schmidt (1988) has called these open and closed tasks, respectively. Not only are almost all tasks variable in some aspects of their context, but this variability is generally not predictable. This means that it is not possible to train an individual on every task variation that will be encountered. The learner must aspire to be flexible enough to handle the variation that will be encountered.

A potential way to induce flexibility in transfer is by introducing variability in training. Generally, the effects of variability in training are positive (reviews may be found in Cormier, 1984; Shapiro & Schmidt, 1982; Johnson, 1984; Newell, 1985; Schmidt, 1988; and Shea & Zimny, 1983). It is important to note that variable practice can facilitate subsequent performance not only in open tasks, where intrinsic variation is high, but also in closed tasks, where variations in the transfer context are relatively minimal. Consider for example the study by Kerr and Booth (1978). They had children throw miniature beanbags at a target on the floor. A screen prevented them from actually seeing the target or their hand while they were making any actual throw, but they could look at the target before a throw and see the outcome. Half the 8-yr. olds practiced at a 3-ft distance, whereas the other half got a mixture of 2-ft and 4-ft trials. Performance was superior for the variable practice group despite the fact that the fixed practice group was training on the target distance.

The previous volume in this series (Druckman & Bjork, 1991) discussed two effects related to the issue of variable practice: contextual interference and use of variable examples during training. It has been suggested that these are part of a common phenomenon (Schmidt & Young, 1987; Schmidt, 1988), and we will treat them together here. Contextual interference refers to the addition of task demands during training, which tends to inhibit initial acquisition of task competency but ultimately to facilitate transfer. Demands may be introduced, for example, by randomly varying conditions from trial to trial rather than blocking (Reder, Charney & Morgan, 1986; Shea & Morgan, 1979) or by adding a second task requirement (Battig, 1956). Variability of examples refers to training with varied task content but with the same general task requirement, for example, throwing bean bags of several different weights in practice before transferring to a novel weight (Carson & Wiegand, 1979). Again, such variation has been found to facilitate transfer, although not without exception (van Rossum, 1990).

Several theories have been proposed to explain practice variability and contextual interference effects. Battig (1979) proposed that in an effort to overcome contextual interference, learners undertake elaborative and variable, rather than rote and repetitive, encoding of task information. This strategy has two consequences: First, the information is more retrievable, so that previously practiced instances will be performed better at transfer. Second, task-relevant information becomes better distinguished from irrelevant context, so that elements central to the task achieve higher strength and adaptation to new instances is promoted. Similar consequences could also arise because variability in task instances leads to distributed rather than massed practice. That is, when training is varied, repetitions of the same instance tend to occur at longer intervals, which again would promote richer and more varied encoding and would decontextualize the encoded information.

Elaborations of these strengthening and de-contextualizing mechanisms have been offered in subsequent theories. Shea and Zimny (1983) proposed that variable presentations increase the likelihood that different elements of information will reside concurrently in working memory and hence will be associatively related at encoding. Anderson (1983) has suggested that variable practice leads to more general productions with wider applicability. This happens because each specific production (with more contextual elements) is only strengthened when it is practiced, while the more general form of the production is practiced each time one of the variations is practiced. Thus one develops a stronger version of the general form of the production.

A number of additional accounts of variable-practice effects have been offered. Charney and Reder (1987) have suggested that variable practice facilitates a component of skill acquisition in which the learner becomes able to recognize the appropriate procedure to use in a given context. Variable practice of exercises forces the learner to practice the procedure selection component of the task (figuring out which procedure to use for this situation). This makes the task more difficult initially, but target performance is better because the learner now is able to select the procedure that is appropriate in each context. In keeping with this view, Reder, Charney & Morgan (1986) found that when exercises were not blocked by type and subjects had to figure out which procedure to use to solve the problem, initial performance was worse, but final performance was better.

Variable-practice effects have also been interpreted in the context of theories that stress the importance of schematic knowledge representations as the elements that must be identical for transfer to occur. Schmidt's (1975) schema theory for motor learning is such an account. It assumes that performance is guided by schemata that represent movement parameters and movement outcomes, on the basis of past experience. The recall schema represents an abstract motor program, and the recognition schema represents expected feedback consequences of an action -- in essence, what it should feel like. The two are used together to plan and generate actions. For example, there might be a generalized schema for throwing a ball, with a force parameter used to adjust the throw so that balls of different weights could be thrown equivalent distances. Formation of these schemata is assumed to be promoted by presentation of varying instances, because it provides more data about the underlying rules that relate movement parameters and sensory consequences (including observed outcomes). Transfer to new instances will then be facilitated because the existing schema can be applied, allowing the performer to generate appropriate parameters and expected consequences. One can see why this would help in open tasks, where instance variability is high and new instances are frequently introduced. Under the assumption that a schema is more stable or retrievable in memory than isolated instances, this approach might also explain why performance in a closed task benefits from variations in training. (See also Shapiro & Schmidt, 1982.)

The idea of schema abstraction has been similarly proposed in theories of analogical transfer, where solution of an initial "source" problem (e.g., an algebra "word problem") is intended to facilitate later solution of a "target" problem that is structurally similar but different in surface description. It has been suggested that successful analogical transfer leads to the induction of a general schema for the solved problems that can then be applied to subsequent instantiations (Holyoak, 1984; Novick & Holyoak, 1991; Ross, 1989). If a schema is used in this way, one would expect to find that practice with a greater number of instances facilitates analogical problem solving. Consistent with this idea, analogical transfer has been found to be facilitated by the provision of multiple analogous source problems, along with instructions to compare them (Catrambone & Holyoak, 1989; Gick & Holyoak, 1983).

Yet another possibility is that the set of component processes that is trained under a variable practice regimen is more inclusive than one trained under specific practice. This is suggested by studies using a variant on a contextual-interference paradigm (Carnahan & Lee, 1989; Langley & Zelaznik, 1984; after Shea & Morgan, 1979). Subjects were trained to knock down three wooden barriers placed an equal distance apart. Some subjects were trained to produce a target total time (duration training), whereas others were trained on different component times for each barrier that added to the same total time (phase training). When the subjects transferred to a task requiring a new total duration, duration-trained subjects performed no better than those given prior phase training, but the phase training group was superior on transfer to a new phase-control task. Langley and Zelaznik (1984) suggested that this might occur because control of phasing is a higher-order skill that incorporates control of duration as well. Thus phase-trained subjects had learned skills applicable to both tasks, but duration-trained subjects had not. An alternative hypothesis is that of contextual interference, since the phase-trained group had three different movement components to learn whereas the durationtrained group had only one. However, evidence against this hypothesis was provided by Carnahan and Lee (1989), who contrasted phase-trained groups that were trained either on three distinct intervals, one for each submovement, or on the same duration for each submovement. Although the variable phasing group performed with more error during learning (as would be expected if they had higher contextual interference), the two phase-trained groups performed equally at transfer, and both outperformed a duration-trained group. Thus the higher-level skill requirements of the phasing task, rather than interference during training, appeared to be responsible for more effective transfer.

It might be useful to consider the issue of practice variability in the context of a distinction between the content of a task and the task requirements for processing that content (Smith, 1990). Initial training may lead to both content specific learning and mastery of the process. When variable instances are introduced in training, but the task itself remains the same, learners are faced with varying content but a constant set of task processes. The positive effects of variable training on transfer to new instances can then be viewed in two ways: First, variability could have a greater strengthening effect on task processes than constant practice (e.g., by eliminating strategies that work only for limited content). Second, exposure to varying previous content could improve accessibility to new content (e.g., by associative priming or by precluding perseveration in retrieval routes). It seems likely that the extent of these effects would vary with the task. For example, if process learning is a small part of a task, and if there is little interaction between old instances and new ones, there should be little advantage for variable training. This might be the case with alphabet arithmetic, where as we have seen, transfer is minimal after extended training with multiple instances (Logan & Klapp, 1992). In contrast, Smith, Branscombe and Bormann (1988) had people make judgments of the form, Is behavior X trait Y? (e.g.., Is hitting friendly?). When subjects were given 200 trials with one trait, but using different behaviors, there was substantial transfer to a new trait. (This study did not address the effects of practice variability, since all subjects were given varied instances in training).

Part-task training. We next consider part-to-whole (or vertical) transfer, often called part-task training in applied contexts. There are many everyday situations where we learn a task in parts, then transfer to a whole situation. Learning to drive a car is an example; novices are typically trained separately on shifting gears and steering. There has been interest in the efficacy of this approach since the inception of formal studies of learning (see Knerr, Morrison, Mumaw, Stein, Sticha, Hoffman, Buede, & Holding, 1987; McGeoch & Irion, 1952, Naylor, 1962, and Wightman & Lintern, 1985 for reviews). A general issue is whether part- task training is superior to training that uses the whole task from the outset, which would violate the general claim that training should simulate the transfer context as closely as possible. Even lacking an overall advantage, one can still ask to what extent part-task training produces positive transfer to the task as a whole.

Knerr et al. reviewed part-task training in the context of airplane flight skills and pointed to practical reasons for developing such training methods. One is the level of complexity of flight tasks. When tasks are complex enough to comprise multiple distinct components, part-task training seems both natural and imperative. There is also substantial potential for saving money by training in parts, especially when some task components can be trained via simulation. Training transitions, such as updating techniques in response to new equipment, may also be facilitated by part-task retraining. There is also the fact that many applied tasks involve working in teams, and training of individuals outside of the team constitutes a form of part-task training (Salas, Dickinson, Converse & Tannenbaum, in press). These practical considerations could justify a part task training program as long as it produced a reasonable degree of positive transfer, even if there was no overall advantage over whole-task training.

Does part-task training produce positive transfer, and if so, is it as effective as or even more effective than whole-task training? Like many issues related to transfer of training, the answer to this question is complex. The efficacy of part-task training depends in large part on the nature of the task that is to be learned. Two task variables that have received particular emphasis are the difficulty or complexity of the task and the degree to which it is structurally integrated or organized. Naylor (1962) hypothesized that these variables interact: When a task is highly organized, the usefulness of whole-task training will increase with task complexity. When a task is easily decomposed into parts, the usefulness of part-task training will increase with task complexity. Naylor's hypothesis has received substantial support, but there have also been violations (see Knerr et al., 1987).

In the domain of motor performance, Schmidt and Young (1987) suggested that when a task constitutes a sequence of distinct components or "programs" and has a relatively long duration (e.g., > 10 s), practicing individual subcomponents will produce positive transfer. One could potentially identify boundaries between such subcomponents by looking for points where temporal aspects of the task vary the most from one performance to another. However, when the task is continuous and its subcomponents do not form a clear sequence, there is often little transfer from part-task training. This is likely to be the case for rapid, ballistic tasks (Schmidt, 1991). In this case negative transfer from part-task training may actually occur, because the neuromuscular structure of a component may be fundamentally different when it is practiced in isolation from the structure in the whole-task context.

Tasks may be divided into parts not only sequentially, but on the basis of the cognitive/perceptual processes involved. A successful part-task training program based on task division into cognitively distinct components was devised by Mane, Adams, and Donchin (1989). Patrick (1992) suggested that task components might be distinguished on the basis of required processes such as perceptual detection, concept learning, problem solving, motor coordination, and rule following. Christina and Corcos (1988) pointed to another cognitive factor, the attention span of the learner, which may limit the ability of an individual to deal with the whole-task situation.

The success of part-task training also depends on training procedures, such as how the task is broken down during training and how it is reconstructed at transfer. Three general methods for decomposing a task are simplifying, fractionating, and segmenting. Simplifying is done by modifying or eliminating task demands, for example, reducing dimensions of control or eliminating time constraints. Fractionating refers to separate practice on components of tasks that in the whole-task form would overlap in time; an example is dividing control of a plane into separate pitch, roll, and yaw components. Thus each practice condition is simplified, but no task component is eliminated from consideration. Segmenting refers to division of a task into temporal or spatial components.

The effectiveness of these various methods depends in part on how they are recombined. Methods for recombination include pure part (practice each part in isolation, then combine all) progressive part (incrementally add parts to a combination, practicing each separately before adding it), and repetitive part (incrementally add parts to a combination, but without separate part practice). Incremental additions may be further varied by whether parts are added in the order of whole-task execution or the reverse order. There are also variations in how much the combined group of subtasks is practiced between new additions.

The aviation training literature suggests that segmentation is highly effective when tasks are recombined by a reverse repetitive-part technique, that is, successively adding task components in the order from those performed latest to those performed earliest. This technique, called "backward chaining," was used successfully by Bailey, Hughes, and Jones (1980), for example, in training a dive-bomb maneuver as a four-segment event. Pilots trained with segmentation and backward chaining tended to learn faster and when transferred to the whole sequence, they outperformed those given whole-task training.

The case for fractionation and simplification is somewhat mixed. By its very nature, fractionation eliminates training on how to integrate and share components of the task over time. Not surprisingly, then, it appears to be particularly deficient, relative to whole-task training, when tasks have interdependent, time-shared components. On the other hand, fractionation does generally produce some positive transfer and may become more effective at higher levels of practice (Knerr et al., 1987).

With simplification, there is some danger of negative transfer, since the more complex version of the task may call for new responses to the same stimulus conditions used previously (Lintern & Roscoe, 1980). Wightman (1983) suggested that simplification will be most effective if the simplified task focuses on those components that potentially produce the greatest error. In landing on an aircraft carrier, one such task is controlling the glideslope to the landing. He found that a manipulation designed to simplify this component -- shortening the lag between the throttle input and the visual glideslope indicator -- had little effect on training or transfer. Again, the problem appeared to be negative transfer: The progressive lengthening of lag over training called for new responses while there was no corresponding change in the stimulus display.

A general message from the part-task training literature is that careful analysis of the task is called for. There is a clear need at this point for a task taxonomy that will (a) identify variables predicting the potential success of a part task training program and (b) for a given task, suggest techniques for decomposition and recombination that will optimize part training. It seems clear that at least some tasks are aided by part-task training, and given the practical considerations mentioned above, there is strong motivation for the development of these training techniques.

3.6 Does Longer Training Lead to Better Transfer?

The relationship between level of initial learning and transfer is complex. While it is commonly the case that higher levels of initial learning lead to longer retention, that does not necessarily mean that greater initial learning will produce greater benefits in a transfer situation. Put simply, this reflects the fact that differences between the training and transfer contexts may be more difficult to accommodate, the greater the initial learning level.

Early studies of transfer in a verbal learning context attempted to experimentally isolate the basic underlying stimulus-response processes. They typically used paired associate paradigms, in which subjects learned a list of A-B pairs (e.g., A terms are words; B terms are digits) and were then transferred to new lists. Performance of these experimental subjects could then be compared to subjects who learned the second list without prior training on another. Learning in this task can be subdivided into (at least) four components: learning the stimulus terms, learning the response terms, learning stimulus-response associations, and learning response-stimulus associations. The transfer context can retain any of these component processes while placing demands on others. For example, transfer from pairs like sweet-7 to sour-7 and happy-8 to sad-8 should take advantage of learning on all prior components (assuming mediation between antonyms like sweet and sour), whereas transfer from pairs like sweet-7 and happy-8 to sweet-8 and happy-7 should benefit from initial learning of the individual stimuli and responses, but suffer from interference due to remapping of associations. Not surprisingly, these procedures produce positive and negative transfer, respectively.

The effect of degree of prior learning will depend in part on the rate of learning of the various components. For example, in a transfer situation where old stimulus and response items are retained but are associated in new ways, there will be positive effects from prior item learning but negative effects from prior associative learning. If the items (positive effects) are learned faster than the associations (negative effects), it would be better to transfer early in learning, when positive benefits have been attained without too much cost on the negative side. But we cannot be sure that items are learned faster than associations; indeed, this depends greatly on the nature of the paired associates themselves. It is not surprising that in reviewing the verbal transfer literature, Kausler (1974, p. 232) termed degree of learning to be an "enigmatic variable."

Similar principles can be applied outside of the relatively simple context of paired verbal associates. Mastery of a task requires learning the pool of relevant stimulus cues, learning to perform the response repertoire, and learning the relationship between the cue context and responses. Longer training periods should produce greater learning of all of these components; however, some of that learning may not be productive in the transfer context.

The picture is further complicated when we consider that the nature of processing can change qualitatively over the course of learning, in this way affecting transfer. Changes that have been suggested include modification of the stimulus cues that control responses, for example, changes from visual control to proprioceptive control in tracking tasks (see Cormier, 1984). Stimuli may also be redefined; for example, they may become responded to as category members rather than instances (Cheng, 1985). Another idea is that people learn to extract the "invariants" (Gibson, 1979) that are the predictors of response requirements and to ignore uncorrelated cues that accompany them (Lintern, 1991). Still another assumption is that attentional demands of tasks are reduced, that is, that individuals become automatic responders over the course of skill acquisition (e.g., Ackerman, 1988; Schneider & Shiffrin, 1977). All of these changes over the course of learning should have effects on transfer that will interact with the relationship between the training and transfer contexts.

Automatized processing, in particular, has been suggested to produce potentially negative effects because of the specificity of learning that results (see Cormier, 1984). It is difficult to change automatized responses given a change in task conditions. Fisk, Lee and Rogers (1991) demonstrated these effects in a category search task, in which people indicated the position of an exemplar of some target taxonomic category among a vertical array of three items (e.g., looking for a bird name in an array: robin, arm, hand). Under consistent mapping conditions, a given taxonomic category was always used for targets and never for distractors; for example, a subject might search for a bird name among body parts but never search for a building name among birds. Under variable mapping conditions, a given taxonomic category provided targets on some trials and distractors on others. Consistent mapping from stimulus to response has been shown to be critical for developing the characteristics of automatic processing: low response times and errors and null effects of increasing workload (e.g., increasing the number of targets to be sought on any one trial).

After 10 training sessions, subjects showed positive transfer from consistently mapped tasks to new tasks using the same targets with novel distractors or the same distractors with novel targets, relative to performance with entirely new sets of targets and distractors. In contrast, pronounced negative transfer resulted from reversing responses to consistently mapped targets or distractors, so that previous targets became distractors or vice versa. This occurred even when a subset of the original consistently mapped items was maintained unchanged, that is, when old targets were sought among items that had previously been consistent targets, or when previously consistent distractors became new targets that were sought among old distractors. The cost of remapping targets was greater than that for remapping distractors.

These results were interpreted in terms of a strength theory that is reminiscent of earlier theories of transfer in pairedassociate learning tasks. Automaticity is assumed to result in consistently trained targets' having a high strength in memory and consistent distractors' having a low strength. Negative transfer when target/distractor roles are reversed results because the low-strength items have relatively weak signals and form poor targets, or because the high-strength items have relatively strong signals and cannot be rejected easily. These changes in strength resulting from consistent training, particularly of target items, are not easily overcome.

An important conclusion from this study is that once automatic responses are attained to a stimulus set, changes in even part of that set will substantially disrupt performance. If lengthy training leads to automatic responses to task-irrelevant stimuli, which are not similarly linked to those responses in the transfer context, transfer performance will be impaired. Against the potentially negative effects of automaticity, with concomitant over-specificity of the stimulus representation, must be balanced the positive effects of extended learning, in the form of mastering the repertoire of responses, integrating and strengthening the representation of relevant stimuli, and learning the rules that associate them.

3.7 Should Feedback be Given During Training? If So, When and How Much?

A surprising generalization that comes out of the research exploring the importance of feedback to ultimate performance is that such feedback rarely helps and sometimes actually hurts long-term learning. This is true whether the task to be learned is motoric or cognitive in nature. It is important to distinguish between intrinsic and extrinsic feedback, concurrent vs. terminal feedback, immediate vs. delayed, and between separate and accumulated feedback. Intrinsic feedback refers to the knowledge one gets immediately and automatically simply by attempting to perform the task. It is obvious when one hits the tennis ball out of the court, and it is obvious when one misses when swinging a baseball bat at a ball. Extrinsic feedback, however, comes from outside the learner, provided either by an individual or a training device. If the feedback is received while the activity is being performed, such as while riding a bicycle, we would say that the feedback is concurrent; however, the success of a tennis stroke might be thought of as terminal, when one sees where the ball lands. Extrinsic feedback that is provided after the execution of the task (terminal) can be further classified by whether the feedback is immediate or delayed, and if not delayed whether it is given for each attempt (separate) or is accumulated across trials, and only a summary of performance is given. A final distinction that should be made is between knowledge of results (KR) and knowledge of performance (KP). The former refers to how well one did at executing a task in terms of the outcome of the action; the latter refers to the action or movement pattern involved in the skill (e.g., "your elbow was bent when you tried to hit the ball"). Most of the research has looked at the effectiveness of knowledge of results.

There are numerous reviews on the effects of feedback (e.g., Salmoni, Schmidt & Walter, 1984; Wheaton, Rose, Fingerman, Karotkin, & Holding, 1976). Schmidt (1988) provides a thorough review of much of the literature pertaining to motor learning. An interesting finding from his laboratory is that if the feedback (KR) is provided after each trial, immediate performance is better than when only summary performance is given every 15 trials; however, long-term or ultimate learning is facilitated by giving only summary feedback intermittently such as after 15 trials.

Anderson, Conrad and Corbett (1989) examined feedback in the context of learning the programming language LISP. They taught the language using a computer based tutor and varied whether or not they provided feedback and when it was provided. Like Schmidt, they also found an immediate benefit of feedback without finding any long-term benefit from it. In contrast to some of the results that Schmidt reports, Anderson et al. did not find a long term deficit of feedback.

An explanation for the lack of long-term benefit from error feedback was suggested by Miller (1953), who realized that offering feedback concurrent with doing a task might serve as a crutch. When the feedback is removed during later performance, those trained with the feedback are at a disadvantage. Goldstein and Rittenhouse (1954) provided some of the first evidence that concurrent feedback as opposed to summary information at a delay produces short term gains, but no long term benefits. One should not conclude that it is optimal to deprive the learner of any knowledge of results; rather, feedback should not provide too much information too soon.

This conclusion may seem at odds with a literature from the tradition of verbal learning and concept formation, in which subjects learn to assign multi-attribute items to classes according to some experimenter-designated rule (e.g., Restle, 1962; Bower & Trabasso, 1964). Usually in these tasks, there is an arbitrary stimulus-to-response mapping, and no intrinsic feedback about the correctness of a response is available. Feedback then apparently enables subjects to prune the set of possible hypotheses about the correct rule (Levine, 1966; Trabasso & Bower, 1966), and its delay will only impose the load of holding previous stimulus-response pairings in memory. In contrast, this section has emphasized task situations in which intrinsic feedback is typically available and the stimulus-response mapping is non-arbitrary.

3.8 How Can One Predict When Transfer Will Occur?

Transfer is a problem of high dimensionality. Much of the controversy over transfer may be seen to result from different theorists and experimenters working within different areas of a high-dimensional transfer "space," in which transfer is predicted by a host of variables, some of which have yet to be considered.

The notion of a transfer space was suggested for the learning of lists of verbal stimulus-response associates by Osgood, in 1949. He constructed a theoretical three-dimensional transfer surface, where two of the dimensions were the similarity between the original and transfer lists with respect to the stimulus and response terms, and the third dimension was the amount (and direction) of transfer. For example, having learned the pair dog-basket, transfer to canine-basket would involve a semantically similar stimulus term, while transfer to dog-bucket would involve a semantically similar response term. Transfer situations with semantically similar stimuli and responses were predicted to produce positive transfer; in contrast, considerable negative transfer was predicted when the same stimulus items were trained with one set of response terms and then paired with new ones at transfer.

However, predictions of the Osgood model were not always accurate, particularly with respect to negative transfer (Bugelski & Cadwallader, 1956). Negative transfer was found to be anomalously strong when old stimuli were paired with responses that were new but similar to the old ones (an effect sometimes called the Skaggs-Robinson paradox; Robinson, 1927). In addition, particularly in the domain of motor skills learning, negative transfer effects were often small and vanished with practice (Bilodeau and Bilodeau, 1961).

Holding (1976) proposed a modification of Osgood's transfer surface that, while admittedly imperfect, was consistent with general findings regarding similarity effects in motor learning. These findings include a decrease of positive transfer as the stimulus similarity between training and transfer tasks decreases, no transfer when entirely new stimuli and responses are used in the transfer task, large positive transfer when the previously trained responses are used with different stimuli, and negative transfer when the same stimuli are used at transfer but with responses that are different -- but similar to the originals. Holding suggested that most practical transfer is positive, and that negative transfer is most likely when there is a failure to discriminate between distinct training and transfer stimuli that are intended to elicit distinct responses, or when the responses themselves cannot be well discriminated.

At the same time, Holding (1976, p. 8) noted that "there are, in fact, good reasons for viewing all transfer surfaces with mistrust." A major problem is their low dimensionality; many important influences on transfer have been ignored by any existing formalism. We have been stressing (as do transfer surfaces) the similarity of the training and task environment. We have also discussed influences from the nature of training, such as training schedules and the availability of feedback. Situated learning theorists stress social variables; one could also discuss motivation and learner intelligence.

Efforts to understand transfer would clearly be helped by greater efforts to characterize the nature of relevant variables. A valuable addition would be a task taxonomy, stressing distinctions such as those between cognitive and motor tasks, between tasks performed in isolation and those performed in groups, or between tasks limited by perceptual vs. memory vs. response factors. All of these (and more) are likely to modulate the effects of other variables on the nature and degree of transfer. Efforts to make relevant distinctions can be seen, for example, in the literature on part-whole training, which focuses on the intrinsic integration (or lack thereof) of task components as a determinant of training outcomes.

3.9 How Important is Fidelity to Simulation?

The fundamental tenet of situated learning, that learning should occur in the same context as eventual execution, has been evaluated particularly carefully in the context of simulator devices. Training devices that simulate the actual situations in which learners will ultimately find themselves vary in terms of the fidelity or realism of the device, in relation to the actual performance environment. The size and complexity of the task that is simulated can also vary, e.g., from a hand grenade that has its explosive charge replaced with sand to teach the skill of activating and throwing a grenade, to a full-scale flight simulator. The latter can employ computer-generated imagery of the landing field, a full motion platform, complete cockpit instrumentation, etc. The common feature to all simulations is that they provide practice on aspects of the task environment.

The reasons for simulating the actual task situation as opposed to training in it are several: The real environment may be too dangerous, too costly, too time-consuming, or too rare to find. For example, the National Training Center at Ft. Irwin provides soldiers with tough, realistic training, using air-land battle exercises in the Mojave Desert involving "infantry, armor, artillery, aviation, chemical, logistics, air defense, engineering, military police, electronic warfare, and special operations units" (Wiering, 1992, p. 18). There is a very high priority placed on realism: If a commander wants a ditch dug, he must send his soldiers out and they must dig that ditch. If a soldier is injured in a battle, he must be evacuated, and if he dies a replacement soldier must be requisitioned. The simulated enemy forces are actually American troops trained to fight according to Soviet military doctrine. There are key differences between this simulation and real warfare that make NTC training an effective teaching device. These include the use of observer controllers who monitor performance during each simulated battle. An important learning vehicle is the intense after-action review, where strategic mistakes are discussed and soldiers are induced to discover their own correct actions and mistakes. Of course, soldiers killed (tagged by a laser weapon) in a simulated battle have a chance to learn from their mistakes. Historically, it is the first major battle of any war that creates the most casualties. With experience, soldiers make fewer mistakes.

Another obvious place where simulators are important is for training airplane pilots. Not only is it considerably safer to use training devices, it is often cheaper than flying a real plane that requires gas and maintenance. Errors can be corrected during training via a simulation that might cost a life in a real airplane. Further, specific types of errors can be identified and enable the learner to practice those subcomponents of the task that need most attention. Another advantage of simulations is that difficult or emergency situations, e.g., failure of an indicator, can be simulated at will to see whether or not the operator can cope.

There are others who think that high levels of fidelity are not cost effective, and may even be detrimental (e.g., Andrews, 1988; Patrick, 1992). Andrews (1988) notes that too much fidelity can sometimes be worse than a simpler representation of the environment. She states, "Trainees who are new to a particular piece of equipment such as an aircraft, a sonar system, or a nuclear control panel often have a difficult time in learning the proper operation or maintenance techniques. This is so because the cues of the real equipment to which the trainees must respond are too subtle, fast, transitory or complex for the novice to fully comprehend" (pg. 49).

Patrick (1992) suggested, based on ideas of Miller (1954) and Gagne (1954), that different types of simulation be used for different levels of proficiency, when acquiring the same skill. During the initial stages of skill acquisition, learners would do better to be familiarized with very simple representations of the equipment, if the task and equipment are quite complex: "In the first stage of skill acquisition...simple simulations (pictures, diagrams, mock-ups) could be used to familiarize the trainee with the nomenclature and location of the displays, and with the controls involved in the perceptual-motor tasks...In the second stage of skill acquisition... the trainee should be able to practice coordinating movements and also making anticipations in the same manner as that required by the operational task. In the final stage ...simulation should support high levels of practice of the task at high speeds and under heavy workloads or in low signal-to-noise ratios" (p. 503).

Although developers of simulation devices are often concerned with how many real-world phenomena can be mimicked, the important question is: Which perceptual cues should be reproduced in the training cycle? Miller (1954) has made an important distinction between engineering fidelity and psychological fidelity. The former is the degree to which the training device duplicates the physical characteristics of equipment or duplicates the surrounding environment in which the learner will ultimately find him or herself. Miller argued that there is a point of diminishing returns in terms of cost-savings of training per degree of increased engineering fidelity. That is, increasing the fidelity of the simulation device will yield only small improvements in performance over a simpler device beyond certain levels. Patrick (1992) has suggested that engineering fidelity is not the issue at all; that the critical determinant of transfer is the psychological fidelity of the simulation device. He notes that when training cognitive tasks and procedures, high transfer can be achieved with simulations of low physical fidelity, and so the theoretical tradeoff of increasing savings in training with increased fidelity does not apply. The key issue is, what are the necessary factors that produce psychological fidelity in a simulation? Research on transfer of training is critical to addressing this issue.

There have been a number of studies that have shown no advantage for real equipment or realistic simulators as compared with very inexpensive cardboard mock-ups or simple drawings when teaching various types of procedural tasks (e.g., Caro, Corley, Spears, & Blaiwes, 1984; Cox, Wood, Boren & Thorne, 1965; Crawford & Crawford, 1978; Grimsley, 1969; Johnson, 1981; Prophet & Boyd, 1970; Trollip, 1979; see Valverde, 1973, for a review of early studies). For example, Cox et al. (1965) compared training using real equipment (costing $11,000), with training using a realistic simulator (costing $1000) and cheap cardboard mock-ups and photographs. There were no observed differences in training time or in long-term performance or retention. Prophet and Boyd compared acquisition of the startup and shut-down procedures for a Mohawk aircraft when trained with the real thing or a cheap mock-up, and here too there was no difference.

The degree to which simulations must incorporate specific features of the real-world environment can be thought of in terms of a given feature's "cueing potential" (Cormier, 1987). A person needs to be able to recognize the cue in the target task as the same cue that was used for training to enable the learner to execute the appropriate actions. This view is similar to the notion of identical elements, but emphasizes that the recognition of a cue may fail if the surrounding context is different. Instead of merely including the surrounding context as other elements that must match, one can speak of the similarity of the context and how easily it affords recognition of the cueing elements. If the simulation or training elements are not recognized as cues for response, the simulation will not be effective for later performance.

In principle, one can analyze the correspondence between cues of the learning task or simulation, and those of the transfer task, thereby determining the level of fidelity required. Salvendy and Pilitsis (1980) compared training programs for teaching suturing techniques to medical students and found that the best performing groups were those groups that used an electromechanical device, allowing them to puncture simulated tissue and providing auditory and visual feedback on correctness of technique. Those that merely heard lectures about suturing or watched a movie of the techniques did not perform as well.

Certainly, there are tasks where added simulation fidelity helps training. This has been heavily documented with respect to aircraft simulators. On the other hand, one must be cautious in making general statements about the degree of fidelity required for optimal training, because seemingly minor differences in transfer contexts may produce substantial differences in simulator effectiveness. For example, motion cues in flight simulators sometimes improve training, at other times not. Ince, Williges and Roscoe (1975) found that different kinds of motion (rough-air simulation vs. cockpit motion during turns) have different effects. Further, whether motion cues are important during training also depends on the nature of the task required of the pilot during transfer. That is, transfer from the training environment to the performance environment was not a simple function of the overlap in identical elements. Some were more important to be represented in the training task, and which ones were more important depended on what performance features had to be recalled.

4. Conclusions

In this section we re-visit the questions raised in the report and formulate conclusions based on each one.

1. Is transfer of training difficult to achieve?

There is broad consensus that transfer is typically very specific to the context in which training has occurred -- but this outcome is not inevitable. Failures of transfer, being more noticeable, tend to mask the many successes.

2. Is transfer optimal if training is situated in the context of application?

As a general principle, having identical elements between the training and performance context facilitates transfer. However, it is impossible in most cases to fully anticipate the performance context, rendering fully situated learning impossible. Further, many factors may reduce the effectiveness of situated learning, including, for example, high costs of imposing training fidelity, reduction of time spent on acquisition and practice of underlying procedures, and disruption to the target context from the presence of novices.

3. Does traditional training on abstract concepts and rules facilitate transfer?

Although concrete experience is critically important, the teaching of abstract principles has been shown to play a role in acquiring skills over a broad domain of tasks.

4. Must task-irrelevant elements be duplicated for transfer to occur?

Mismatches between task-irrelevant elements of the training and performance context could produce slight decrements in transfer; however, this can be ameliorated by varying the training context with respect to these elements.

5. Should training mimic the experiences that are anticipated?

The principle that training and transfer should have identical elements suggests this is generally true. The positive outcomes of variable practice and part-whole training suggest strong constraints on this conclusion.

6. Does longer training lead to better transfer?

The relation between length of training and performance is not a simple one. Clearly, longer training has facilitatory effects, but it can also lead to inflexibility when the training fails to anticipate variability in the performance context.

7. Should feedback be given during training? If so, when and how much?

Somewhat surprisingly, giving immediate and constant feedback may fail to optimize training; delayed and intermittent feedback may produce superior results because it allows learners to detect and self-correct errors and it diminishes reliance on extrinsic sources.

8. How can one predict when transfer will occur?

Few generalizations can be made about transfer as a whole; what is needed is a task taxonomy that characterizes the nature of the critical elements that should be held identical between the training and performance contexts. For example, conceptually based tasks may be very different from rote motor tasks with respect to the need to maintain superficial, physical elements of the performance context.

9. How important is fidelity to simulation?

The level of fidelity that is required depends on the nature of the elements that must be preserved. In many cases, a high degree of fidelity is not required and may be detrimental.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, 117: 288-318.

Anderson, J. R. (1976). *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum.

Anderson, J. R. (1983) *The Architecture of Cognition*. Cambridge, Mass.: Harvard Univ. Press.

Anderson, J. R. (1993) *Rules of the Mind*. Hillsdale, N.J.: Erlbaum.

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP Tutor. *Cognitive Science*, 13: 467-506.

Andrews, D. H. Relationships among simulators, training devices, and learning: A behavioral view. *Educational Technology*, January, 1988, pg. 48-54.

Bailey, J. S., Hughes, R., G., & Jones, W. E. (1980). Application of Backward Chaining to Air-to-Surface Weapons Delivery Training (AFHRL-TR-79-63, AD-A085 610). Williams AFB, AZ: Operations Training Division, Air Force Human Resources Laboratory.

Battig, W. F. (1956). Transfer from verbal pretraining to motor performance as a function of motor task complexity. *Journal of Experimental Psychology*, 51: 371-378.

Battig, W. F. (1979). The flexibility of human memory. In L. S. Cermak & F.I.M. Craik (Eds.), *Levels of processing in human memory* (pp. 23-44). Hillsdale, N. J.: Erlbaum.

Biederman, I., & Shiffrar, M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual learning task. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13: 640-645.

Bilodeau, E. A., & Bilodeau, I.M. (1961). Motor-skills learning. *Annual review of psychology*, 12: 243-280.

Bjork, R. A. & Richardson-Klavehn, A., (1989). On the Puzzling Relationship Between Environment Context and Human Memory. In C. Izawa (Ed.) *Current Issues in cognitive processes: The Tulane Flowerree Symposium on cognition*. New Jersey: L. Erlbaum.

Bovair, S., Kieras, D. E., & Polson, P. G. (1990) The acquisition and performance of text-editing skill: A cognitive complexity analysis. *Human Computer Interaction,* 5: 1-48.

Bower, G.H., & Trabasso, T. R. (1964). Concept identification. In R. C. Atkinson (Ed.), *Studies in Mathematical Psychology.* Stanford, CA: Stanford University Press, pp. 32-94.

Brown, J.S., Collins, A. & Duguid, P. Situated Cognition and the Culture of Learning. 1988 Tech Report from Institute for Research on Learning, No. IRL88-0008.

Bugelski, B. R., & Cadwallader, T. C. (1956). A reappraisal of the transfer and retroaction surface. *Journal of Experimental Psychology*, 52: 360-366.

Carnahan, H., & Lee, T. D. (1989). Training for transfer of a movement timing skill. *Journal of Motor Behavior*, 21: 48-59.

Caro, P. W., Corley, W. E., Spears, W. D., & Blaiwes, A. S. Training Effectiveness Evaluation and Utilization Demonstration of a Low Cost Cockpit Procedures Trainer (Report No. NAVTRAEQUIPCEN 78-C-001301). Pensacola, Fla.: Seville Training Systems, 1984.

Carson, L. M., & Wiegand, R. L. (1979). Motor schema formation and retention in young children: A test of Schmidt's schema theory. *Journal of Motor Behavior*, 11: 247-251.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15: 1147-1156.

Charney, D. H. and Reder, L. M. (1987) Initial skill learning: an analysis of how elaborations facilitate the three components. In P. Morris (Ed.) *Modelling cognition*. Chichester, England: J. Wiley & Sons.

Cheng, P. W. Restructuring versus automaticity: Alternative accounts of skill acquisition. *Psychological Review*, 92: 414-423.

Chipman, S. F., Segal, J.W. and Glaser, R. (Eds.). (1985) Thinking and learning skills. *Research and open questions*, Volume 2. Hillsdale, NJ: Lawrence Erlbaum.

Christina, R. W. & Corcos, D .M. (1988) *Coaches guide to teaching sport skills*. Champaign, IL: Human Kinetics.

Cognition and Technology Group at Vanderbilt. (1990) Anchored instruction and its relationship to situated cognition. *Educational Research*, 19: 2-10.

Cormier, S. M. (1984). Transfer of training: An interpretive review. Technical Report 608. Alexandria: Virginia: US Army Research Institute.

Cormier, S.M. (1987). The Structural Processes Underlying Transfer of Training. Transfer of Learning, Contemporary Research and Applications. Alexandria, VA: Army Research Institute.

Cox, J.A., Wood, R.D., Jr., Boren, L.M. and Thorpe, H.W. (1965). Functional and Appearance Fidelity of Training Devices for Fixed Procedures Tasks. HUMRRO Technical Report 65-4. Alexandria, VA: Human Resources Research Office.

Crawford, A.M. & Crawford, K. S. (1978). Simulation of operational equipment with a computer-based instructional system: A low cost-training technology. *Human Factors*, 20: 215-224.

Druckman, D., and Bjork, R. A., (Eds.) (1991). *In the Mind's Eye* : *Enhancing Human Performance.* Committee on Techniques for the Enhancement of Human Performance, National Research Council. Washington, D. C.: National Academy Press.

Fernandez, A., & Glenberg, A. M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition*, 13: 333-345.

Fisk, A. D., Lee, M. D., & Rogers, W. A. (1991). Recombination of automatic processing components: The effects of transfer, reversal, and conflict situations. *Human Factors*, 33: 267-280.

Gagne, R. M. (1954). Training devices and simulators: Some research issues. *American Psychologist*, 9: 95-107.

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

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15: 1-38.

Godden, D.R., & Baddeley, A.D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66: 325-331.

Goldstein, M. & C. H. Rittenhouse (1954). Knowledge of results in the acquisition and transfer of a gunnery skill. *Journal of Experimental Psychology*, 48: 187-196.

Gray W. D. & Orasanu, J. M. (1987) Transfer of cognitive skills. In Cormier, S. M. & Hagman, J. D. (Eds.), *Transfer of Learning: Contemporary Research and Applications.* New York: Academic Press

Greeno, J.G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex Information Processing: The Impact of Herbert A. Simon*, pp. 285-318. Hillsdale, NJ: Erlbaum.

Greeno, J. G., Smith, D. R., Moore, J. L. (1993.) Transfer of situated learning. In D. K. Detterman & R. J. Sternberg (Eds.) *Transfer on Trial: Intelligence, Cognition, and Instruction*, pp. 99-167. Norwood, N.J.: Ablex

Grimsley, D. L. (1969) Acquisition, Retention and Retraining: The Effects of High and Low Fidelity in Training Devices. Technical Report 69-1, HUMRRO. Alexandria, VA: Human Resources Research Office.

Hayes, J.R. and Simon, H.A. (1977) Psychological differences among problem isomorphs. In Castellan, N. Jr., Pisoni, D. and Potts, G. (Eds.). *Cognitive Theory*, Volume II. Potomac, MD: Lawrence Erlbaum.

Hendrickson, G. & Schroeder, W. H. (1941) Transfer of training in learning to hit a submerged target. *Journal of Educational Psychology*, 32: 205-213

Holding, D. H. (1976). An approximate transfer surface. *Journal of Motor Behavior*, 8: 1-9.

Holyoak, K.J. (1984). Analogical thinking and human intelligence. In R. J. Sternberg (Ed.), *Advances in the Psychology of Human Intelligence* (Vol. 2, pp. 199-230). Hillsdale, NJ: Erlbaum.

Ince, F., Williges, R.C. and Roscoe, S.N. (1975). "Aircraft simulator motion and order of merit of flight attitude and steering guidance displays." *Human Factors*, 17: 388-400.

Independent Commission on the Los Angeles Police Department (1991). Report of the independent commission on the Los Angeles Police Department.

Johnson, P. (1984). The acquisition of skill. In M. M. Smyth & A. M. Wing (Eds.), *The Psychology of Human Movement* (pp. 215-240). London: Academic Press.

Johnson, S.L. (1981). Effect of training device on retention and transfer of a procedural task. *Human Factors*, 23: 257-272.

Judd, C. H. (1908). The relation of special training to general intelligence. *Educational Review*, 36: 28-42.

Katona, G. (1940). *Organizing and Memorizing*. New York: Columbia University Press.

Kausler, D. H. (1974). *Psychology of Verbal Learning and Memory.* New York: Academic.

Kerr, R., & Booth, B. (1978). Specific and varied practice of a motor skill. *Perceptual and Motor Skills*, 46: 395-401.

Kessler, C. M. (1988). Transfer of Programming Skills in Novice LISP Learners. Doctoral Dissertation, Carnegie-Mellon University, Pittsburgh.

Klahr, D., & Carver, S. M. (1988). Cognitive objectives in a LOGO debugging curriculum: Instruction, learning, and transfer. *Cognitive Psychology*, 20: 362-404.

Knerr, C. M., Morrison, J. E., Mumaw, R. J., Stein, D. J., Sticha, P. J., Hoffman, R. G., Buede, D. M., & Holding, D. M. (1987). Simulation-Based Research in Part-Task Training (AF HRL-TR-86-12, AD-B107 293). Brooks Air Force Base, Texas: Air Force Human Resources Laboratory.

Langley, D. J., & Zelaznik, H. N. (1984). The acquisition of time properties associated with a sequential motor skill. *Journal of Motor Behavior*, 16: 275-301.

Lave, J. (1988). *Cognition and practice: Mind, mathematics, and culture in everyday life.* Cambridge: Cambridge University Press.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge, England: Cambridge University Press.

Levine, M. (1966): Hypothesis behavior by humans during discrimination learning. *Journal of Experimental Psychology*, 71: 331-338.

Lintern, G., & Roscoe S. N. (1980). Adaptive perceptual-motor training. In S. N. Roscoe (Ed.), *Aviation psychology*. Ames, IA: Iowa State University Press.

Lintern, G. (1991). An informational perspective on skill transfer in human-machine systems. *Human Factors*, 33: 251-266.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95: 492-527.

Logan, G. D., & Klapp, S. T. (1992). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory & Cognition*, 17: 179-195.

MacKay, D.G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, 89: 483-506.

Mane, A., Adams, J. A., & Donchin, E. (1989). Adaptive and part-whole training in the acquisition of a complex perceptual-motor skill. The Learning Strategies Program: An Examination of the Strategies in Skill Acquisition. *Acta Psychologica* (Special Issue), 71: 179-196.

Mayer, R. E., & Greeno, J. G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology*, 63: 165-173.

McGeoch G. O., & Irion, A. L. (1952). *The Psychology of Human Learning* (2nd ed.) New York: Longmans.

McKendree, J. E., & Anderson, J. R. (1987). Frequency and practice effects on the composition of knowledge in LISP evaluation. In J. M. Carroll (Ed.), *Cognitive Aspects of Human-Computer Interaction*. Cambridge, MA: MIT Press.

Miller, R. B. (1954). Psychological Considerations in the Design of Training Equipment (WADC-TR-54-563). Wright Air Development Center, Wright Patterson AFB, Ohio, December 1954. (AD 71202)

Miller, R.B. (1953). Handbook on Training and Training Equipment Design. Report No. 53-136, Wright-Patterson Air Force Base, Wright Air Development Center.

Moore, J. L., & Greeno, J. G. (1991). Implicit understanding of functions in quantitative reasoning. *Proceedings of the Thirteenth Annual conference of the Cognitive Science Society*, Hillsdale, N.J.: Lawrence Erlbaum, 1991.

Naylor, J. C. (1962). Parameters affecting the relative efficiency of part and whole practice methods: A review of the literature (NAVTRADEVCEN 950-1, AD-275 921). Port Washington, NY: US Naval Training Device Center.

Newell, K. M. (1985). Skill learning. In D. H. Holding (Ed.), *Human skills* (pp. 203-226). Chichester: John Wiley & Sons.

Novick, L.R., & Holyoak, K.J. (1991). Mathematical problem solving by analogy. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17: 398-415.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review*, 56: 132-143.

Patrick, J. (1992). *Training: Research and practice*. San Diego, CA: Academic Press.

Polson, P. & Kieras, D. E. (1985). A quantitative model of learning and performance of text editing knowledge. In L. Bormann and B. Curtis (Eds.) *Proceedings of CHI '85 Human Factors in Computing Systems Conference*. New York: Association for Computing Machinery.

Prophet, W. W. & Boyd, H.A. (1970) Device-task fidelity and transfer of training: Aircraft cockpit procedures training. Technical Report 70-10. Alexandria, VA: Human Resources Research Organization.

Reder, L. M., Charney, D. H. & Morgan, K. I. (1986). The role of elaborations in learning a skill from an instructional text. *Memory and Cognition*, 14: 64-78.

Restle, F. (1962). The selection of strategies in cue learning. *Psychological Review*, 69: 329-343.

Robinson, E. J. (1927). The "similarity" factor in retroaction. *American Journal of Psychology*, 30: 297-312.

Ross, B. H. (1989). Distinguishing types of superficial similarities: Effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15: 456-468.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (in press). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their Training and Performance*. Norwood, NJ: Ablex.

Salmoni, Schmidt & Walter (1984). Knowledge of results and motor learning. A review and critical reappraisal. *Psychological Bulletin,* 95(3): 355-386.

Salvendy, G., & Pilitsis, J. (1980). The development and validation of an analytical training program for medical suturing. *Human Factors*, 22: 753-770.

Sandberg, J., & Wiedlinga, B. (1992) Situated Cognition: A paradigm shift? *Journal of Artificial Intelligence in Education*, 3: 129-138

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82: 225-260.

Schmidt, R. A. (1988). *Motor Control and Learning: A Behavioral Emphasis*. Champaign, IL: Human Kinetics Publishers.

Schmidt, R. A. (1991) *Motor learning and performance*. Champaign, Il: Human Kinetics, 188-190.

Schmidt, R. A., & Young, D. E. (1987). Transfer of movement control in motor skill learning. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of Learning*, pp. 47-49. Orlando, FL: Academic Press.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, 84: 1-66.

Segal, J.W., Chipman, S. F., & Glaser, R. (Eds.). (1985). Thinking and learning skills. Volume 1. *Relating instruction to research*. Hillsdale, NJ: Lawrence Erlbaum.

Shapiro, D. C., & Schmidt, R. A. (1982). The schema theory: Recent evidence and developmental implications. In J.A.S. Kelso & J. E. Clark (Eds.), *The Development of Movement Control and Coordination* (pp. 113-150). New York: Wiley.

Shea, J. B., & Morgan, R. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology; Human Learning and Memory*, 5: 179-187.

Shea, J. B., & Zimny, S. T. (1983). Context effects in memory and learning movement information. In R. A. Magill (Ed.), *Memory and Control of Action* (pp. 345-366). Amsterdam: North-Holland.

Simon, H. A. (1980). Problem solving and education. In D. T. Tuma & F. Reif (Eds.), *Problem Solving and Education* (pp. 81-92). Hillsdale, N.J.: Lawrence Erlbaum

Singley, M. K. (1986). Developing Models of Skill Acquisition in the Context of Intelligent Tutoring Systems. Doctoral Dissertation, Carnegie Mellon University.

Singley, M. K., & Anderson, J. R. (1989). *The Transfer of Cognitive Skill*. Cambridge: Harvard Univ. Press.

Smith, E. (1990). Content and process specificity in the effects of prior experiences. In R.K. Srull and R.S. Wyer, Jr. (Eds.), *Advances in Social Cognition*, Vol. III, pp. 1-59. Hillsdale, NJ: Erlbaum.

Smith, E. R., Branscombe, N.R., & Bormann, C. (1988). Generality of the effects of practice on social judgment tasks. *Journal of Personality and Social Psychology*, 43: 385-395.

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, 5: 460-471.

Smith, S. M. (1988). Environmental context-dependent memory. In D. M. Thomson & G. M. Davies (Eds.), *Memory in Context: Context in Memory*. New York: Wiley, pp. 13-34.

Smith, S. M., & Vela, E. (1986). Outshining: The Relative Effectiveness of Cues. Paper presented at the annual meeting of the Psychonomic Society, New Orleans, LA.

Stucky, S.U. Situated cognition: A strong hypothesis. To appear in F.L. Engel, D.G. Bouwhuis, T.Bosser, & G. Adewalk (Eds.), *Cognitive modeling and interactive environments*. Proceedings of the NATO ARW.

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

Trabasso, T. R., & Bower, G. H. (1966). Presolution dimensional shifts in concept identification: A test of the sampling with replacement axiom in all-or-none models. *Journal of Mathematical Psychology*, 3: 163-173.

Trollip. S. R. (1979) The evaluation of a complex computer-based flight procedures trainer. *Human Factors*, 21: 47-54.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80: 352-373.

Valverde, H. H. (1973). A review of flight simulator transfer-of-training studies. *Human Factors*, 15: 510-523.

Vela, E. (1989). Environmental Context Dependent Memory: A Meta-Analytic Review. Paper presented at the annual meeting of the Psychonomic Society, Atlanta, Georgia.

Vera, A. H., & Simon, H. A. (1993). Situated action: A symbolic interpretation. *Cognitive Science*, 17: 7-48.

Wheaton, G. R., Rose, A. M., Fingerman,P. W., Korotkin, A. L., and Holding, D. H., (1976) Evaluation of the Effectiveness of Training Devices: Literature Review and Preliminary Model. Research Memorandum 76-6. Army Research Institute.

Whitehead, A. N. (1929). *The Aims of Education*. New York: MacMillan.

Wiering, G. (1992). A Diamond in the Rough: The National Training Center. Army Research, Development & Acquisition Bulletin, March-April, 18-21.

Wightman, D. C. (1983). Part-Task Training Strategies in Simulated Carrier Landing Final Approach Training (NAVTRAEQUIPCEN IH-347). Orlando, FL: Naval Training Equipment Center.

Wightman, D. C., & Lintern, G. (1985). Part-task training for tracking and manual control. *Human Factors*, 27: 267-283.