Reder, L. M. & Klatzky, R. (1994) Transfer: Training for
Performance. In Druckman, D. & Bjork, R. A. (Eds.) Learning,
Remembering, Believing: Enhancing team and individual
performance. Washington, D.C.: National Academy Press.

Transfer: Training for Performance

Probably the most critical issue in any type of learning is how well the learning transfers from one situation to another, particularly to the actual performance of a task. 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 chapter we focus particularly on situations in which there is some period of training prior to on-the-job execution of a task. We focus exclusively on individuals in the training context. That context may or may not be a group setting; Chapters 5-7 consider group learning.

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 critical because most 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; 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 in task performance, and even an antici-pated circumstance may be impossible to fully simulate in training. Ideally, a training program should produce the ability to accommodate some degree

of variability—both within the task environment and between the training and task environments—as well as establish basic skills required for the task itself.

One principal question underlies this chapter: How much does the training context have to incorporate the performance context in order to produce effective transfer? This question applies to training on mathematics problems at school, flying a plane, training for combat, controlling airplane traffic at an 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. How close in fidelity to the performance situation must a device be in order to be effective? Simulating unnecessary aspects of the performance context means that training will be inadequate. The fundamental issue of specificity of training (fidelity of simulation in training) is determining the factors that produce transfer. Only when those factors are understood can resources be balanced against outcomes to design optimal training programs.

An issue related to training fidelity is the relative importance of training individuals on abstract generalizations, in contrast 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 chapter discusses a strong position in regard to this issue, which is 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 in comparison with learning without a prior task? Transfer can be 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 is between tasks that are similar in complexity and do not have an inclusion relation.

TRANSFER BY IDENTICAL ELEMENTS

As noted above, a major issue in developing training programs is the required similarity between the training and performance contexts. If the training context exactly simulates the performance context, transfer should

be perfect, but this situation rarely exists. Theories of transfer have stressed the requirement of learning and 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 a person's mental faculties accounted 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 defines 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, the 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.

One example of transfer of identical elements comes from Singley and Anderson (1989), who taught people three different computer text editing programs, varying the order of acquisition of the three programs across subjects. They developed a set of 107 rules capable of simulating editing in the three programs. Some of the rules were shared by all three programs, some by two, and some were unique to a specific program. A given editing task using one program might or might not share many rules with the comparable task using a different program. Their expectation was that there would be savings in performance with the second program to the degree that the rules had already been acquired. The data were in strong support of their predictions.

A different example—or transfer between nonidentical elements—comes from MacKay (1982), who reported data from a group of English-German bilingual speakers who were asked to produce the same sentence in the same language 12 times, with a 20-second pause between sentences. They were to produce the sentence as quickly as possible. The production time declined regularly over the 12 repetitions, to an asymptote of about 2 seconds. In the last 20-second 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 the previous ones, despite a complete change in the motor movements necessary for output.

An example of very little transfer comes from Logan and Klapp (1992), who had subjects solve alphabetic arithmetic problems (e.g., if $A = 1$, $B = 2$, etc., does $A + 2 = D$?) with one set of 10 letters and the digits 2-5 for 12 sessions of nearly 500 trials each. Initially, the subjects' response times increased markedly with a new digit (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 milliseconds (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 new digit increased dramatically—to nearly the value it had had in the initial sessions. That is, there was only a small amount of transfer.

These three examples 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 these differences be explained? The answer lies in what elements are and are not identical.

Cognitive Abstractions

Singley and Anderson (1989) and Bovair et al. (1990) have developed formal theories that specify the "elements" of cognitive tasks. They 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 from these theories, as well as the model of Polson and Kieras (1985) are similar to Anderson's (1983) ACT (Adoptive Control of Thought) model, so ACT is 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 procedural or skill memory. Many of the tasks that people perform are originally encoded in memory in 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 by associations acquired through 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 the associations and nodes.

Procedural memories are condition-action, or production, rules: condition is the "lefthand side" of the production; action is the "righthand side" of the production. The righthand side of a production specifies an action, and it is executed when all of the elements of the lefthand condition side are met, that is, 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 (or conflict resolution, as it is called) of a single production to fire is determined by the specificity and the strength of the competing productions. Specificity refers to the number of condition elements that match on the lefthand side. More specific (or complex) productions are easier to fire than more general productions, all else being equal. A production can vary enormously in complexity. For example, it can set up new goals to be achieved, place new elements in working memory, or initiate a simple motor response.

According to models 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 another task to the extent that memory chunks and productions are shared. Anderson (1993) 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 muscula-
ture. In this model transfer occurs between tasks that share mental struc-
tures. Practice on one task strengthens linkages that are necessary for the
other.

The MacKay 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 model explains the 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 then all
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.

The transfer will not be perfect, however, if the unshared structures
between two tasks are not well practiced, for example, when one attempts to
write with the nondominant hand. Only the abstract level of structure per-
taining to the written words is shared in this case; the muscle-movement
structures for the nondominant hand are not well practiced, and being unshared,
they do not benefit from practice with the dominant one. Furthermore, in
this case the shared mental structure—the name—is already well practiced
and cannot benefit further from the commonalities between learning and
performance.

Exemplars

A theory of Logan (1988) makes the fundamental assumptions that people
performing in a task store instances of past performance in memory and that
each instance is stored as an independent copy or "exemplar." On their first
encounter with the task, having no stored instance, people will use whatever
strategic, rule-based tools they have available; this constitutes a task "algo-
rithm." Subsequently, however, they will have available not only the algo-
rithm, but memory of the past instance of performance, or as many in-
stances as 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, the algorithm or a retrieved instance. The time to retrieve each
past instance is assumed to vary stochastically (the probability being a func-
tion of previous instances) 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, past instances of that same task are 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 consti-
tutes not only the general type of problem that must be solved, but also the
specific parameters that are provided. Learning is item-specific, and re-
trieval is of the same items that were previously used in the task. It is
assumed that transfer between distinct items within the same task does not
occur because presentation of a novel item does not lead to retrieval of an
item used previously in training. This theory explains why so little transfer
was obtained in the Logan and Klapp experiment of letters and numbers
when the specific letters changed. What little transfer did occur was attrib-
uted to better learning of an algorithmic solution (e.g., with practice, sub-
jects became able to count faster from the new letter).

It should be noted that Logan emphasizes tasks in which conceptual
structures are minimally important and instance retrieval provides an effec-
tive solution. Without identical elements at the instance level, there is little
transfer. Singley and Anderson (1989), and MacKay (1982) as well, em-
phasize tasks in which conceptual structures are very important and re-
sponses are either extremely well practiced or minimally taxing parts of the
task. Hence, nonidentical elements at the output level matter little. (We
return below to the issue of how transfer may be governed by different
principles in different tasks.) 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 virtu-
ally no transfer.

The Ease or Difficulty of Transfer

The cases of successful transfer discussed above, along with much of
the research discussed below, make clear that transfer is not invariably
difficult to achieve. However, there is also a surprising amount of research
demonstrating the difficulties of achieving transfer. In the difficult cases,
performance often seems to be overembedded in the training context, so
that the identical elements across contexts are not perceived and so do not
have an 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 et al., 1985; Segal et al., 1985). Hayes

and Simon (1977), for example, showed little transfer between two structur-ally or formally identical problems that differed in superficial characteris-tics. 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 of the similarity of elements was given.

Charney and Reder (1987) proposed that an imp . ant 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 the training context always makes obvious whic. procedure should be practiced, then an important element of the skill—the ability to recognize when to use each 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 more general (less specific) productions when there was variable practice. Un-fortunately, there was little evidence that people actually generalized from variable practice to novel situations if the new situations were very differ-ent 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 only because the context or situation is very different, but also because the task has dramatically changed. For example, Kneer et al. (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 some functions of the programming language LISP facilitated performance on evaluation of other LISP func-tions (compared to a control), but not on generating LISP functions, and vice versa.

In theory, when productions are initially formed they are very condi-tion- and action-specific. The specificity of the production rules would be unnoticed in everyday situations because one does not usually practice only half of a skill, such as learning to evaluate LISP functions but not to gener-ate them. In order to achieve transfer, a 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, suc-cessful 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 "re-name" 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.

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 is called situated learning (e.g., Lave, 1988; Lave and Wenger, 1991), although closely associated terms are situated cognition and situated action. The situated approach is taken by researchers from several fields, including psychology, anthropology, and philosophy. Its fundamental tenets include an emphasis on contextual de-terminants 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 do not necessarily agree on all of its details, four general principles characterize the approach:

(1) Action is grounded in the concrete situation in which it occurs. A potential 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 informa-tion structures and processes are inadequate or 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 con-text 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 under-stand performance, it is necessary to understand the social situation in which it occurs, including the way in which social interactions affect 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 perfor-mance is affected by context, as has been amply demonstrated. In 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 (Tulving and Thomson, 1973). 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 nonmatching situations.

What makes the situated-learning approach different is the degree to which learning is claimed to be context specific and the implications of this claim for education and training. For example, Greeno et al. (1993:99) state:

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 onl... ...sed there. Lave and others reject the premise that school constitutes a neutral setting in which things are learned that can later be applied in the real world. Brown et al. (1988) argue that success with schooling has little bearing on performance elsewhere. So, for example, rather than being ...ight mathematics as an abstract skill, a person should learn the mathematic... techniques relevant for his or her trade in the situations in which they will be needed. Some theorists of situated learning also argue that people tend not to remember skills or be able to apply them if they are taught in an abstract manner. Only when they are learned "on the job," embedded in the performance 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, that people perceive the environment directly and use that perception to support thought. Stucky (in press) argues that rather than using contextual cues to represent relevant aspects of a 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 the consequences of it, like stacking. Transfer from one environment to another then depends on common properties that produce invariance of interaction.

It cannot be contested that one of the most important goals of training is improving teaching techniques so 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 performance context in order to be effective. This is of course a fundamental premise of the situated-learning approach. In support of the premise, there are many cases of superior performance with more specific, relevant practice. And because it is difficult to anticipate all of the features of a performance context, it would seem appropriate to train in that context to ensure that training covers all of those features.

However, a strong position that basic training is a waste of time, because learning must be situated, is false. There are many domains in which fundamental skills are critical to acquire before more specific training can occur, such as learning to catch a ball before playing any sport with a ball. Furthermore, 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 mathematics will necessarily be better if it is learned in real-world contexts is debatable, and there have been no studies that address this complex issue. We suspect that for people with strong mathematics skills, embedding mathematics in real-life contexts takes away valuable time from the acquisition and practice of fundamental concepts and procedures.

The theory that training in the performance context is optimal led the advocates of situated learning to propose 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. Training in the context of applica- tion 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.

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. And there are many situations in which one cannot imagine placing a novice directly in the work context: playing in an orchestra, copiloting an airplane, or serving in a tank crew are perhaps obvious examples, but even providing nursing care or fighting a fire do not seem feasible.

Yet another problem with training in the target context is the potential

for variability within that context. Because one cannot always anticipate the future contexts that will be required of the learner, a major instructional goal is to devise a training procedure that will optimize performance in various contexts.

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, for example, soldiers have been trained in simulated battles at the National Training Center at Ft. Irwin in the Mojave Desert, California (see Wiering, 1992). Over the course of a 4-week rotation at the center, trainees become demonstrably better at fighting simulated battles. There is reason to believe that this training, using real armored vehicles, negotiating in real deserts, improved performance in the transfer task of fighting real desert battles in the Middle East. (Simulation is discussed further below.)

One point that can be well taken from situated-learning theorists is the stress on evaluating the role of the performance context of—particularly the social context—when designing a training regimen. The 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, 1991:125). 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 the training program may often ignore them. A workplace analysis may reveal such characteristics and lead to their being dealt with during training to the extent that it is possible to do so.

GENERAL PRINCIPLES OF TRANSFER

A general principle of transfer seems to be that identical elements are necessary. But which elements? And how much identity is necessary? Must learning be situated in the transfer context in .der to be effective? Or does learning in one context become so situated that it cannot be generalized to other contexts? The rest of this chapter discusses many aspects of these broad questions.

Role of Abstract Concepts and Rules

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 and 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 mathematics, should be taught by using physical models rather than using symbolic ics, should be taught by using physical models rather than using symbolic formulas and algorithms. As noted above, proponents of situated learning emphasize the importance of knowledge embedded in the performance situation and to deemphasize the role of manipulating internal symbols in accomplishing a task. One problem with this view is that it fails to explain cases where the abstract concept is developed before the real-world model. For a classic example, G. F. B. Riemann developed his non-Euclidean geometry as an abstract theory much before Albert Einstein used it to provide an explanation of the universe based on general relativity.

Greeno et al. (1993) propose a more general view that transfer can occur to the extent that the situations share characteristics or similarities for interaction, which can be perceived without there being a symbolic representation of the properties that specify the similarities. In the example given above, children can directly perceive flatness of a surface without having a symbolic "resentation of flatness. Vera and Simon (1993:22-23) responded:

> 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 many examples in which 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, such as 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 1," "part 2," 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 than a principle, 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 deduce the formula from principles and examples, but only on solving problems similar to those encountered during training. The other 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 and Judd, as described by Judd, 1908). If a thrower aims directly at the target under water, the dart will go beyond the actual location of the target, because of the misleading cues from the refraction of light. One group of fifth- and sixth-grade students received an explanation of light refraction, while another group did not. On the init~¹ 'ask, where the depth of water was 12 inches, both groups performe. ..t equally well; however, when the water level was changed to 4 inches, the difference between groups became striking. the students without the abstract instruction were confused and made large and persistent errors, while those who had received the abstract instruction corrected their ai~ quickly. When the depth of water was changed again, 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, has also been shown to 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 sexing photographs of chickens, while an additional part of an expert's skill involves knowing how to pick up a chicken to determine its sex.)

These studies provide strong support for the benefits of abstract instruction, but it is important to emphasize that abstract instruction in the absence of concrete examples rarely results in transfer. 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.

Environmental Context

The question of whether the environmental context of learning matters—whether task-irrelevant elements have to be duplicated for transfer to occur—has been addressed primarily in research on verbal learning. That research focuses on whether changes in the environmental context from learning to remembering affect the amount that is remembered.

This issue refers to the effects of incidental context, 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 (Bjork and Richardson-Klavehn, 1989). For example, the presence of posters citing cancer statistics in a room where people receive an anti-smoking lecture would be considered an influential rather than incidental context.

In a prototypical experiment, 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).

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 and Glenberg, 1985). This inconsistency in outcomes can be better understood if one considers two factors that have been suggested to modulate the degree of environmental influence: the availability of memory cues other than the environment, and 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 and 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 in a recognition situation. It can also occur because the original learning situation promoted self-cuing, such as instruction at the time of learning that encouraged subjects to think about items conceptually and relate them to one another (e.g., forming words into a story). Bjork and Richardson-Klavehn (1989) suggested that this hypothesis should be augmented by a "reinstatement hypothesis," that physically reinstating the context of learning will be useful only when subjects cannot mentally do so for themselves. For example, if a person studied in a pink room but is tested in a green one, imagining the green walls to be pink may compensate for the change in environment (see, e.g., Smith, 1979). According to these hypotheses, 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 more than 50 studies and calculating the size of the context effect (defined as the performance difference between no-context-change and context-change groups), Vela found overall a moderate, statistically significant advantage for the same context in learning and test. This result 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 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 cuing 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 broad range of performance settings.

Fidelity of Training to Anticipated Experience

It is interesting that the traditional literature on transfer gives many examples in which 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

Almost all tasks are variable in some aspects of their context, and this variability is generally not predictable. This is true for tasks like flying a fighter plane in combat or seemingly invariant tasks like performing the broad jump, what Schmidt (1988) has called open and closed tasks, respectively. 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 (see reviews in Cormier, 1984; Shapiro and Schmidt, 1982; Johnson, 1984; Newell, 1985; Schmidt, 1988; and Shea and Zimny, 1983). 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 (e.g., Kerr and Booth, 1978; see also Chapter 4 for a description of this experiment).

The committee's previous book (Druckman and 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 and Young, 1987; Schmidt, 1988), and we treat them together. 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 (Reder et al., 1986; Shea and 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 and Wiegand, 1979). Again, such variation has been found to facilitate transfer, although not without exception (Van Rossum, 1990, see also Chapter 4 in this volume).

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 vari-

able, rather than rote and repetitive, encoding of task information. Such a 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 promotes richer and more varied encoding and reduces the contextual elements of the encoded information.

Elaborations of these strengthening and "decontextualizing" 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 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 any 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 selection 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 et al. (1986) found that when exercises were not grouped (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 of the elements that must be identical for transfer to occur. In Schmidt's (1975) schema theory for motor learning, it is assumed 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 can 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 is then 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, when variability is high and new instances of the task are frequently introduced. Under the assumption that a schema is more stable or retrievable in memory than are isolated instances, this approach might also explain why performance in a closed task benefits from variations in training (see Shapiro and 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 problems (Holyoak, 1984; Novick and 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 and Holyoak, 1989; Gick and 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 explanation is suggested by studies using a variant on a contextual-interference paradigm (Carnahan and Lee, 1989; Langley and Zelaznik, 1984; after Shea and 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), and 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. 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 while

the duration-trained 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 ~~~r during learning (as would be expected if they had higher context-~~~ ...terference), 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 is interesting 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, and 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 varies 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, for which transfer is minimal even after extended training with multiple instances (Logan and Klapp, 1992). In contrast, Smith et al. (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

There are many everyday situations in which one learns a task in parts, then transfers to a whole situation. This is called part-task training. Learning to drive a car transfer, or in applied contexts, part-task training. 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 (for reviews see Knerr et al., 1987; Naylor, 1962; Wightman and Lintern, 1985). A McGeoch and Irion, 1952; Naylor, 1962; Wightman and Lintern, 1985). 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. (1987) 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 by simulation. Training transitions, such as updating techniques in response to new equipment, may also be facilitated by part-task retraining. In addition, many applied tasks involve working in teams, and training of individuals outside of the team constitutes a form of part-task training (Salas et al., 1993). 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 opposite outcomes (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., longer than 10 seconds), practicing subcomponents will produce positive transfer. One could potentially identify boundaries between such subcomponents by looking for points at which temporal aspects of the task vary the most from one performance to another. But 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 its structure in the whole-task context.

Tasks may be divided into parts not only sequentially, but on the basis of the cognitive or perceptual processes involved. A successful part-task training program based on task division into cognitively distinct components was devised by Mane et al. (1989). They devised a space fortress video game in which the objective was to destroy a space fortress by firing missiles from a spaceship. The part-task training prior to whole-task training was cost-effective in that the savings in later training was greater than the amount spent in the part-task training. 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, that 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 decomposed during training and how it is reconstructed at transfer. Three general methods for dec... using 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. 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 reverse order, from those performed latest to those performed earliest. This technique, called "backward chaining," was used successfully by Bailey et al. (1980) 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 pilots given whole-task training.

The evidence 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 and 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, but 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 for a task taxonomy that identifies the variables that predict the potential success of a part-task training program and, for a given task, suggests techniques for decomposition and recombination that optimize part-task 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.

Length of Training

The relationship between level of initial learning and transfer is complex. Although higher levels of initial learning lead to longer retention, they do not necessarily produce greater benefits in a transfer situation. This result reflects the fact that differences between the training and transfer contexts may be more difficult to accommodate after higher levels of initial learning.

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 then were given new lists to learn. Performance of the experimental subjects were compared with subjects who learned the second list without prior training on another list. Learning in this task can be subdivided into (at least) four components: learning the stimulus terms, learning the response terms, learning the stimulus-response associations, and learning the response-stimulus asso-

ciations. 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 did produce positive and negative transfer, respectively.

The effect of the amount of prior learning depends in part on the rate of learning of the various components. For example, in a transfer situation in which 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 one cannot be sure that items are learned faster than associations; indeed, it depends greatly on the nature of the paired associates themselves. Thus, it is not surprising that in a review of the verbal transfer literature, Kausler (1974:232) termed amount 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 the responses. Longer training periods should produce greater learning of all of these components, but some of that learning may not be productive in the trai... context.

The picture is further complicated when one considers that the nature of processing can change qualitativ..ly over the course of learning and so affect transfer. Such changes include modification of the stimulus cues that control responses, for example, changes from visual control to proprioceptive (internal muscle sensations) 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 and Shiffrin, 1977). All of these changes over the course of learning have effects on transfer that interact with the relationship between the training and transfer contexts.

Automatic 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 automatic responses given a change in task conditions. Fisk et al. (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 negative transfer 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 ea... r theories of transfer in paired-associate learning tasks. Automaticity is assumed to result in a high strength in memory for consistent targets and a low strength for consistent distractors. Negative transfer results when target and distractor roles are reversed either 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 easily rejected. These changes in strength resulting from consistent training, particularly of target items, are not easily overcome.

An important conclusion from this research 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. But against the potentially negative effects of automaticity, with concomitant overspecificity 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.

The Role of Feedback

A surprising general finding from research on the importance of feedback to ultimate performance is that it rarely helps and sometimes actually hurts long-term learning (for reviews, see, e.g. Salmoni et al. 1984; Wheaton et al., 1976). This finding holds for both motoric and cognitive tasks in nature.

Although the overall finding is clear, it is important to distinguish between intrinsic and extrinsic feedback, concurrent and terminal feedback, immediate and delayed feedback, and separate and accumulated feedback. Intrinsic feedback refers to the knowledge one gets immediately and automatically when performing a task. It is obvious when one hits the tennis ball. Extrinsic feedback 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, the feedback is concurrent; in contrast, the outcome of a tennis stroke is 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 can be made is between knowledge of results and knowledge of performance: 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 research on feedback has looked at the effectiveness of knowledge of results. 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 on results 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 et al. (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 but no long-term benefit. In contrast to some o.. e results that Schmidt reports, however, Anderson et al. did not fin.. ..ng-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. V. 'en the feedback is removed during performance, those trained with the feedback are at a disad-

vantage. 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 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 and Trabasso, 1964). In these tasks, there is usually an arbitrary stimulus-to-response mapping, and no intrinsic feedback. Feedback then apparently enables subjects to prune the set of possible hypotheses about the correct rule (Levine, 1966; Trabasso and Bower, 1966), and its delay only imposes 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 not arbitrary.

Predicting Transfer

Transfer is a problem of high dimensionality. Much of the controversy over transfer may 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 not yet been studied.

The notion of a transfer space was suggested for the learning of lists of verbal stimulus-response associates by Osgood (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. Positive transfer was predicted in situations with semantically similar stimuli and responses; 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. Predictions of the Osgood model were not always accurate, particularly with respect to negative transfer (Bugelski and 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 from, 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 (1976:8) noted: "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 the research on transfer surfaces. That work stresses the similarity of the training and task environment, but there are also influences from the nature of the training, such as training schedules and the availability of feedback. Situated learning theorists stress social variables; one could also consider motivational variables.

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, memory, or 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.

The Role of Fidelity in 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, for example, 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 use

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.

There are several reasons for simulating a performance situation rather than actual training in it: 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, 1992:18). There is a very high priority placed on realism: if a commander wants a ditch dug, he sends his soldiers out and they 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 were actually American troops trained to fight according to Soviet military doctrine. There are key differences between this simulation and real warfare that make the Ft. Irwin training an effective teaching device, including observer controllers who monitor performance during each simulated battle. An important learning component is intense after-action review, in which strategic mistakes are discussed and soldiers are encouraged to discover their own correct actions and mistakes. Soldiers who are "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 task for which 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, which requires fuel and maintenance. Errors can be corrected during training in a simulator that might cost a life in a real airplane. Furthermore, specific types of errors can be identified and enable a learner to practice those subcomponents of the task that need most attention. Another advantage of simulators is that difficult or emergency situations (e.g., failure of an indicator) can be simulated at will to see whether or not the learner responds correctly.

Others suggest that high levels of fidelity are not cost-effective and may even be detrimental (e.g., Andrews, 1988; Patrick, 1992). Too much fidelity to an actual complex system can sometimes be worse than a simpler representation of the environment. As Andrews (1988:49) 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.

On the basis of ideas of Miller (1954) and Gagne (1954), Patrick (1992:503) suggested that different types of simulation be used for different levels of proficiency of 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 no-menclature 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 operation-al 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.

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 the surrounding environment in which the learner will ultimately be required to perform. 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, beyond certain levels, increasing the fidelity of the simulation device will yield only small improvements in perfor-mance over a simpler device. 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 for cognitive tasks and procedures, high transfer can be achieved with simula-tions of low physical fidelity. 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 et al., 1984; Cox et al., 1965; Crawford and Crawford, 1978; Grimsley, 1969; Johnson, 1981; Prophet and Boyd, 1970; Trollip, 1979; see Valverde, 1973, for a review of early studies). For example, Cox et al. (1965) trained Army personnel to operate a 92-step procedure on a control panel. They compared training using real equipment (costing $11,000) with training using a realistic simulator (costing $1,000) and with cheap cardboard mock-ups and photographs. There were no ob-served differences in training time or in long-term performance or retention. Prophet and Boyd compared acquisition of the start-up and shut-down pro-

cedures for a Mohawk aircraft when trained with the real thing or a cheap mock-up, and 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 "cuing potential" (Cormier, 1987). In order to execute the appropriate actions, a person needs to be able to recognize the cue in the target task as the same cue that was used for training. 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 cuing 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 determin-ing 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 elec-tromechanical device, allowing them to puncture simulated tissue and pro-viding 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. But 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, they do not. Ince et al. (1975) found that different kinds of motion (rough-air simulation or cockpit motion during turns) have different effects. Furthermore, 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 is not a simple function of the overlap in identical elements. Some are more important to be represented in the training task, and which ones are more important depend on what performance features had to be recalled.

CONCLUSIONS

There is broad consensus that positive transfer is typically very specific to the contexts in which training has occurred, but this result is not invari-ant. Failures of transfer, being more noticeable, tend to overshadow many successes.

Identical Elements

As a general principle, having identical elements between training and performance contexts facilitates transfer. However, it is impossible in most cases to fully anticipate the performance contexts, re˙ ing fully situated learning impossible. Furthermore, many factors may ʌeduce the effectiveness of situated learning, including, for example, high costs of achieving training fidelity, reduction of time spent on acquisition and practice of underlying procedures, and disruption of other performers from the presence of novices.

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

The principle that training and transfer should have identical elements suggests training should mimic the experiences that are anticipated. The positive outcomes of variable practice and part-whole training suggest strong constraints on this conclusion.

The level of fidelity that is required in simulators in training 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.

Type and Length of Training

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.

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.

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.

In summary, 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, ent from rote motor tasks with respect to the need to maintain superficial, physical elements of the performance context.

4

Illusions of Comprehension, Competence, and Remembering

People's performances as individuals, and their contributions as members of groups, depend not only on their actual competence, but also on their assessment of that competence. The reading people take of their own state of knowledge or level of skill determines how they allocate their time and energy and the influence they have on others. Many things depend on people's subjective assessment of what they know and do not know: whether they volunteer for certain roles or tasks, whether they seek further practice or instruction, and whether they instill confidence in others, as well as the answers they give to questions from superiors and subordinates and the affect they induce in others by facial expressions and body language (see Chapter 10). Recent evidence suggests, however, that under certain conditions, people's assessments of what they know or remember can be seriously flawed, particularly when they use one index, such as familiarity, recognition, or fluency, to predict something else, such as unaided recall or production.

There are a variety of ways in which people can be fooled. Consider some of the assessments college students typically make in a course. They must decide if they are prepared for an upcoming test. In order to allocate study time, they must monitor the state of their own learning and comprehension across the topics for which they are responsible. If they end up doing poorly on the test, they may take a reading of their memories for how much time they spent studying. At the end of the course, they may also monitor their learning and comprehension in the course—when asked to judge the effectiveness of the instructor, for example, or deciding whether to recommend the course to a friend.