15 Assessing Our Own Competence: Heuristics and Illusions

Robert A. Bjork

ABSTRACT the reading we take of our own competence is arguably as important in many real-world contexts as is our actual competence. For example, in settings where on-the-job learning can be disastrous from a personal or societal standpoint, such as air traffic control, police or military actions, and nuclear plant operation, it can be imperative that we possess the skills and knowledge we think we possess. Individuals who overestimate their own current level of skill and knowledge pose a unique hazard to themselves and others. More broadly, the reading we take of our current level of learning and knowledge in a given domain determines such important matters as how we allocate our time, whether we seek further study or practice, whether we volunteer for or avoid certain assignments, and whether we instill confidence in others.

A variety of recent findings demonstrate, however, that humans frequently misassess their own competence, and that such misassessments typically take the form of overconfidence. At the root of such overconfidence, it is argued here, is a misinterpretation of the meaning and predictive value of certain objective and subjective indices of current performance. That misinterpretation, in turn, rests on a misunderstanding of some fundamental characteristics of humans as learners. One goal of the present chapter is to summarize the types of illusions of comprehension and competence that have been identified in research studies. Another goal is to outline the implications of such results for real-world training and instruction.

Assessing our own competence has at its core predicting the future. In making such assessments, we are estimating, implicitly or explicitly, how well we might perform in some real or imagined future context, such as an athletic event, an exam, an interview, a musical performance, or a job setting of some kind. Such assessments, whether accurate or inaccurate, heavily influence our current and future behavior, such as whether we seek further study or practice, whether we volunteer or decide to compete, and the extent to which—verbally and nonverbally—we instill confidence in others.

Among the bases for such predictions are certain global judgments about ourselves, such as the extent to which we consider ourselves an expert in a given domain, or think of ourselves as a generally competent person, or consider ourselves the kind of person who performs well when it matters. Such judgments about individual differences, while influential and interesting, are not the focus of this chapter. Rather, the focus here is on how we use or misuse subjective and objective indices of our *current* performance as bases for predicting our *future* performance, or more specifically, how we, as learners, interpret—or misinterpret—objective and subjective indices of our own performance during training or instruction.

A variety of recent research findings illustrate that current performance, measured subjectively or objectively, can be a highly unreliable basis for predicting future performance. One goal of the present chapter is to characterize the types of illusions of comprehension and competence to which learners fall prey. A second goal is to outline the practical implications of such findings for the optimization of training, instruction, and performance. In providing this overview, I draw heavily on arguments that my colleagues and I have made elsewhere (Benjamin and Bjork 1996; Benjamin, Bjork, and Schwartz 1998; Bjork 1994a,b; Christina and Bjork 1991; Ghodsian, Bjork, and Benjamin 1997; and Schmidt and Bjork 1992).

15.1 INTERPRETING OBJECTIVE INDICES OF PERFORMANCE

Probably the most obvious basis for assessing our own competence at a given point in the learning or training process is our actual observable performance at that time. In general, our ability to produce answers and procedures quickly and accurately at one point in time *is* a useful guide to predicting how well we will perform in the future. The difficultly and range of questions we can answer, problems we can solve, or tasks we can execute, are other objective indices that have value in judging our preparedness for future tasks. Such objective indices, while generally reliable in our everyday experience, can be misleading, however, especially *during* training. As Christina and Bjork (1991) and Schmidt and Bjork (1992) have emphasized, certain manipulations of the conditions of training can yield good performance and rapid improvement during training, but poor long-term retention and/or transfer to related tasks or altered contexts; conversely, other manipulations that appear to impede performance and slow the learning process can enhance long-term retention and transfer. There is the potential, therefore, for learners to be fooled by their own performance during the training process.

Learning versus Performance

That performance during training is a far from perfect guide to posttraining performance emphasizes the importance of an old distinction in psychology: namely, the distinction between learning and performance. What is observable during training is *performance*—that is, the current speed or accuracy of access to the knowledge and skills that are the target of training. What is not readily observable, but must be inferred, is *learning*—that is, the relatively permanent changes in understanding, comprehension, or competence that support long-term (posttraining) retention and transfer. Changes in performance during training and changes in learning during training are far from perfectly correlated. Substantial learning can take place during periods when few, if any, changes in performance are apparent; and substantial changes in performance during training may be accompanied by little, if any, learning.

Research carried out decades ago, using both humans and animals as subjects, demonstrated that considerable learning can take place during periods when no changes in performance are apparent. With animals, a variety of "latent learning" experiments (e.g., Blodgett 1929; Tolman and Honzik 1930) demonstrated that periods of free exploration of, say, a maze, resulted in substantial learning, but that such learning was only apparent in performance once a particular target behavior was reinforced. Also, with both humans and animals, overlearning or repeated-learning trials, during which little or no additional improvement in performance was evident, were shown nonetheless to enhance learning as measured by reduced forgetting and increased speed of relearning (e.g., Ebbinghaus 1885; Krueger 1929). And in the learning of motor skills (e.g., Adams and Reynolds 1954; Stelmach 1969), periods of massed practice—during which, owing to fatigue, improvements in performance were small or nonexistent—were also shown to result in substantial learning, as measured by later performance under normal conditions.

Such results, and many related findings, led the major theorists of an earlier era (e.g., Estes 1955; Guthrie 1952; Hull 1943; Tolman 1932) to distinguish between the "habit strength" of a response and the "momentary reaction potential" of that response, to use Hull's terms, or between "habit strength" and "response strength," to use Estes's terms. Empirically, habit strength was assumed to be indexed by resistance to extinction or forgetting, whereas momentary reaction potential, or response strength, was assumed to be indexed by the probability, rate, or latency of a response.

Recently, in the context of a "new theory of disuse," Elizabeth L. Bjork and I resurrected such a distinction as part of an effort to account for a broad range of findings in research on human verbal and motor learning (Bjork and Bjork 1992). In our terms, *storage strength* represents the degree of learning of a given response and *retrieval strength* represents current ease of access to that response in memory. The retrieval strength of a given response, balanced against the retrieval strengths of competing responses, is assumed to determine completely the probability that that response is retrievable. Storage strength, on the other hand, acts as a latent variable: it is assumed to enhance the gain of retrieval strength during opportunities for study or practice—and to retard the loss of retrieval strength across time and intervening (interfering) events.

In Thorndike's original "law of disuse" (1914), the central idea was that memory representations, without continued access, decay over time. In our new theory of disuse, learned representations are assumed to remain in memory, but, without intermittent use, are assumed to become gradually inaccessible over time; that is, retrieval strength, not storage strength, is lost. Two assumptions of our theory are especially pertinent to the present analysis. First, as mentioned above, storage strength acts to slow the rate at which access to stored information or procedures is lost; training procedures that build retrieval strength rapidly, but do not build appreciable storage strength, will result in rapid posttraining loss of access to skills and procedures (that is, forgetting). Second, the growth of storage strength, as a consequence of study or practice, is assumed to be negatively accelerated function of current retrieval strength. Thus creating local conditions during training that prop up performance, so to speak—by providing a short-term basis for ready access to correct responses or procedures—will tend to impede the growth of the storage strength necessary to support long-term performance.

Illusions of Competence

Given the empirical distinction between learning and performance, and the corresponding theoretical distinction between storage strength and retrieval strength, the level of one's performance during training can be misleading as an index of one's competence. In particular, to the extent that retrieval strength is misinterpreted as storage strength, we are at risk of gaining illusions of competence. In fact, a variety of recent findings suggest that the opportunities for inflated estimates of our competence are abundant. Certain manipulations of the conditions of training—very common manipulations in real-world settings—appear to facilitate performance (retrieval strength), but not learning (storage strength). That is, whereas the early findings on latent learning, overlearning, and so forth, as cited above, demonstrate that significant learning can take place in the absence of significant changes in performance, a number of recent findings demonstrate that it is also possible for little or no learning to happen, even though there are substantial changes in performance.

Conditions of training that can mislead the learner Thus, in assessing our own competence, certain conditions of training put us at risk of overestimating the degree to which we have actually learned critical skills and information. The following are examples of manipulations of the conditions of training that have the potential to create such illusions of competence.

Massing or blocking practice on a given task One of the most robust findings in experimental psychology is that distributing or spacing practice on information or procedures to be learned enhances long-term retention (for reviews, see Dempster 1990, 1996; Glenberg 1992; and Lee and Genovese 1988). Compared to massing or blocking practice on a given task, the long-term advantages of spacing are often striking: in a number of experiments, the level of recall performance following spaced practice exceeds the level of recall following massed practice by more than two to one. In the short term, however, massed practice typically yields better performance than spaced practice. That is, the immediate consequence of massed practice is rapid improvement. Spaced practice, on the other hand, owing to the forgetting and interference that result from the events and activities that intervene between successive opportunities for study or practice, yields a slower rate of improvement *during* training. Massed practice is therefore more likely than spaced practice to provide the learner with a basis for overconfidence.

Keeping the conditions of practice constant Another robust finding is that introducing variation or unpredictability in the learning environment can enhance long-term performance, especially the transfer of learned skills or knowledge to altered tasks or novel settings. Such results have been obtained in verbal learning tasks, such as learning from text materials (e.g., Mannes and Kintsch 1987; Reder, Charney, and Morgan 1986), in the acquisition of motor skills (e.g., Catalano and Kleiner 1984), in problem solving (e.g., Gick and Holyoak 1980), and in concept learning (e.g., Homa and Cultice 1984). Even varying the incidental environmental context in which learning sessions take place can enhance long-term retention (e.g., Smith, Glenberg, and Bjork 1978; Smith and Rothkopf 1984). A related finding is that random practice, during which the trials on several tasks to be learned are interleaved in unpredictable fashion, produces better long-term retention and transfer than does blocked practice, where the tasks are learned one at a time (hence the trials on a given task are also more massed in time). The benefits of random practice are one instance of a broader set of *contextual interference* effects (Battig 1972), where interleaving the materials or tasks to be learned in ways that cause interference during training produces benefits in long-term retention and transfer (see, for example, Shea and Morgan 1979; Carlson and Yaure 1990).

Once again, however, in all those cases, the *immediate* consequence of such variation or unpredictability is to create difficulties for the learner and to slow the rate of improvement during training. When training is structured and partitioned such that conditions, materials, or tasks remain constant across a given session—that is, processing requirements stay much the same from trial to trial—the *apparent* rate of acquisition is much improved, which is likely to create an illusory sense of competence or accomplishment.

Providing continuous feedback to the learner For many years, one textbook-level generalization about the learning of motor skills was that providing external feedback to the learner is helpful, and increasing the frequency, immediacy, or accuracy of such feedback facilitates learning. Recently, however, Schmidt and his colleagues (see, for example, Schmidt 1991; Schmidt et al. 1989; Winstein and Schmidt 1990) have shown that *reducing* the frequency of feedback during training can enhance long-term posttraining retention. For example, in an experiment where subjects were asked to learn a movement pattern similar to a backhand stroke in tennis, Schmidt et al. (1989) found that providing summary feedback after every 15 trials produced better performance on a delayed test 48 hours after training than did providing feedback after every trial. During training, however, providing feedback on every trial facilitated performance. It appears, then, that providing continuous feedback is another manipulation that has the potential to mislead the learner (*and* the instructor): such feedback bolsters performan(on the short-term, but, apparently, at some actual cost to long-term learnin and performance.

Making the learner an observer-imitator It is very common in real-worl(training programs that trainees are first asked to watch an expert, often th(instructor, demonstrate a skill or procedure to be learned and are then askec to try to imitate that expert. Asking learners to attempt to produce that skil or procedure on their own is far better, from a learning standpoint, than is having them be *solely* passive observers, but such immediate imitation can again mislead the learner. Practice at producing, recalling, or generating skills and knowledge to be learned is essential for learning, but a variety of results in the literature suggest that such practice should not be made too easy: the more difficult or involved such retrieval practice, provided it succeeds, the more effective it is as a learning event (see, for example, Bjork 1988; Jacoby 1978; Landauer and Bjork 1978; Whitten and Bjork 1977). Providing a model of correct performance to be imitated can enhance current performance, but, apparently, at some cost in learning, as measured by performance at a delay. Lee et al. (1997), for example, using a keyboard task, were able to nullify both the short-term negative effects of random practice and the long-term benefits of random practice by providing, at the start of each trial, a computerized simulation of the keystroke pattern to be executed on that trial.

Evaluating training programs If such manipulations are so ineffective in terms of the long-term goals of training, why, then, are they so commonly used? The likely explanation (see Bjork 1994a; Christina and Bjork 1991; and Schmidt and Bjork 1992) is that trainees are not the only ones susceptible to being fooled by their performance during training. Instructors and other individuals responsible for training can also be misled by what they see during training, especially given that such individuals often do not have the opportunity to observe or evaluate posttraining performance. On-the-job performance frequently occurs at a time and place far removed from the training context, but even when that is not the case, rigorous, systematic evaluations of on-the-job performance are rare; and when such evaluations do exist, their purpose, typically, is not to help training personnel evaluate alternative training regimens, but rather to help management personnel evaluate their employees.

With respect to the long-term goals of training, then, instructors are at risk of choosing less effective training regimens over more effective regimens. If what instructors have to work with, so to speak, is performance during training, manipulations such as those discussed above will seem attractive. Conversely, manipulations that introduce difficulties for the learner will seem unattractive, even if those are the very manipulations necessary to produce excellent long-term performance. To make matters worse, training personnel are themselves typically evaluated by their trainees' performance during or at the conclusion of training. Finally, to make matters still worse, another questionable measure frequently plays a very important role in the evaluation of training programs: the happiness or satisfaction of the trainees themselves. It is common for trainees to be asked to fill our evaluations of the training they have received, or are receiving. Given the potential for trainees to be misled by their own performance, such evaluations, often referred to as "happy sheets" or "smile sheets," are unlikely to be the most useful of measures. For example, in an experiment where British postal workers were taught keyboard skill, Baddeley and Longman (1978) found that subjects preferred the most massed of four different schedules of training, whereas, consistent with decades of laboratory research, the most spaced of those schedules was the most effective.

Introducing difficulties for the learner In general, it seems clear that to optimize training one needs to introduce what I have referred to elsewhere as "desirable difficulties" for the learner (Bjork 1994a). Why is that the case? More specifically, what is it about responding to such difficulties that engages the types of processes that create genuine learning? Providing a full discussion of possible processes and mechanisms falls outside the scope of this chapter (for a more complete analysis, see Bjork 1994a), but one counterintuitive point merits comment: in general, the conditions that produce forgetting and errors have much in common with the conditions that produce learning.

Estes (1955) was probably the first to emphasize the relationship of learning and forgetting, which emerged in a natural way in the dynamics of his stimulus fluctuation version of stimulus-sampling theory (see also Cuddy and Jacoby 1982). In Estes's theory, the learning organism, whether human or animal, is assumed to associate responses to be learned to aspects ("elements") of the current stimulus situation. Of the total population of such stimulus elements that might characterize a given learning environment, however, only some are available for "sampling" at any given point in time. As a consequence of intervening events and the activities of the organism, the specific stimulus features or elements of the task environment available to the organism are assumed to shift and change over time; that is, individual elements "fluctuate" between being available and unavailable to the organism.

Forgetting, in Estes's theory, is a product of "conditioned" elements becoming unavailable over time, while—during the same period—elements not yet conditioned to the desired response become available. Such fluctuation, however, is also exactly what is needed for additional learning. Ultimately, long-term performance ("habit strength") depends on the proportion of the total population of stimulus elements—available and unavailable—that have become conditioned to the desired response, whereas "response strength" reflects the proportioned of conditioned elements in the currently available set. Without change ("fluctuation") in the available contextual cues over learning trials, the learning organism will continue to sample conditioned elements in the available set, which will produce a high rate of correct responding, but little new learning: elements in the unavailable set not associated to correct responding will remain that way. As contextual cues fluctuate spontaneously over time, or are induced by events, such previously unavailable cues will tend to become available, resulting in a (possibly dramatic) decrease in the rate of correct responding.

The fluctuation theory was originally proposed to account primarily for a variety of conditioning and extinction phenomena in research on animal learning. The language in which the theory is framed is therefore somewhat inapt with regard to human verbal and motor learning, but the formal properties of the theory, if not the specific terminology, are readily applicable to human learning. Bower (1972) has provided an updated version of the theory, with a focus on human learning, and Glenberg's theory (1979) of spacing phenomena in human memory incorporates cognitive dynamics that are similar, in their formal properties, to the stimulus-sampling dynamics proposed by Estes.

In the new theory of disuse (Bjork and Bjork 1992), which is not a process model, learning is dependent on forgetting in two ways. First, as mentioned earlier, increments in storage strength are a negatively accelerated function of current retrieval strength; as long as local conditions during training make the responses to be learned very accessible, little new learning will occur. Second, the act of retrieval is assumed in the theory to be a potent learning event, but the increments in storage strength (and retrieval strength) are assumed to be greater, the more difficult or involved the act of retrieval (the lower the current retrieval strength of the desired response). Thus, in that theoretical framework as well, conditions that create difficulties for the learner during training have the potential to enhance learning.

15.2 INTERPRETING SUBJECTIVE INDICES OF PERFORMANCE

Our judgments as to how well we "know" something are influenced not only by objective indices of our current performance, but also by certain subjective indices. As we learn and perform, the degree to which what we are asked to learn at a given point seems familiar or understandable influences our assessments of our progress as a learner, as does how readily information and procedures "come to mind" when we are tested. That is, beyond any objective measures of our current performance, the sense of "fluency" we have in processing or retrieving information and procedures is a factor in our assessments of our current competence. More specifically, recent findings suggest that both *perceptual fluency* and *retrieval fluency* are subjective indices that influence our decisions and behavior in significant ways.

Perceptual Fluency

The term *perceptual fluency* refers to the sense of familiarity evoked by presented materials, or the subjective ease with which those materials are perceived. Owing to the efforts of a number of researchers (e.g., Begg et al. 1989; Glenberg and Epstein 1987; Jacoby et al. 1988; Jacoby and Kelley 1987; Reder 1987; Reder and Ritter 1992; Schwartz and Metcalfe 1992; Witherspoon and Allan 1985), it is now clear that perceptual fluency affects a variety of metacognitive judgments, including assessments of our current knowledge. Reder (1987), for example, found that simply preexposing key words in a question (such as the words "golf" and "par" in the question "What is the term in golf for scoring one under par?") increased subjects' confidence that they would be able to produce the answer to a given question. Apparently, owing to repetition priming, such preexposure increased the subjects' sense of fluency or familiarity in reading a given sentence, which was then attributed, at least in part, to the subjects' own state of knowledge. Similarly, Reder and Ritter (1992), using arithmetic problems, and Schwartz and Metcalfe (1992), using paired associates, found that subjects' feeling-of-knowing judgments were increased by preexposure manipulations that made the terms of a given arithmetic problem, or the stimulus member of a paired associate, more perceptually fluent or familiar.

In general, of course, perceptual fluency *is* a useful heuristic. The more knowledge we acquire in a given domain, and the more experience or practice we have in accessing that knowledge, the more readily information in that domain is perceived and understood, and the more familiar that information seems. It is also the case, however, that perceptual fluency can result from factors unrelated to our level of knowledge, which creates a potential problem of inference and attribution for the learner. In the Reder 1987, Reder and Ritter 1992, and Schwartz and Metcalfe 1992 experiments, for example, the preexposure manipulations that increased the subjects' perceptual fluency—and their feelings of knowing—did not, in fact, alter the objective probability that they could provide the correct answer. These and other findings suggest that perceptual fluency—as a subjective index of performance—is open to misinterpretation.

Misattributions and unconscious influences Interpreting the subjective sense of perceptual fluency or familiarity poses a challenge of sorts for the learner. Prior exposures to stimuli can have long-term priming effects that influence not only the objective speed and/or accuracy of our ability to perceive or identify those stimuli, but also our subjective sense of fluency or familiarity. As Jacoby and Kelley (1987) and others have argued and demonstrated, however, such objective and subjective effects are often *not* accompanied by an awareness of the source of those effects. Various experimental findings demonstrate that priming effects can persist well past the point that subjects are either able to recall, or inclined to try to recall, the prior episode responsible for those effects. We are susceptible, therefore, to unconscious influences of prior events, which can take the form of our misattributing the source or cause of our sense of perceptual fluency.

Two types of misattribution have been demonstrated in convincing fashion by the results of recent studies. Misattributing the effects of prior exposure to aspects of the current stimulus situation is one such type. Witherspoon and Allan (1985), for example, using a perceptual identification task in which subjects were asked to try to identify words flashed very briefly on a screen, found that prior presentation of a given word (in an earlier, supposedly unrelated, phase of their experiment) increased not only the actual likelihood that subjects could correctly identify that word, but also increased subjects' estimates of the duration of exposure of a given word on the perceptual identification test. Apparently, subjects attributed their enhanced identification performance (and, we assume, increased sense of perceptual fluency) to increased exposure duration, not to the actual cause of those effects. Another example is an experiment by Jacoby et al. (1988) in which spoken sentences were played against a background of white noise, and subjects were asked to estimate the level of the background noise. When a given sentence had been presented earlier—and presumably was then more perceptually fluent and understandable, the subjects gave a lower estimate of the loudness of the background noise. That is, they attributed the resultant change in perceptual fluency not to prior exposure of a given sentence, but, falsely, to lowered background noise.

The second type of misattribution, illustrated by the Reder 1987, Reder and Ritter 1992, and Schwartz and Metcalfe 1992 experiments cited above, is particularly relevant to the issues of central concern in this chapter: that is, misattributing perceptual fluency/familiarity to the state of one's own knowledge. When something is readily perceived, or seems familiar, we are susceptible to attributing those subjective effects to prior learning, even when the effects are, in fact, a consequence of more mundane events, such as mere prior exposure.

Illusions of competence It is the case, then, that perceptual fluency attributable to causes other than one's state of knowledge can produce illusions of knowing or comprehending. In addition, the results of experiments by Glenberg, Epstein, and their colleagues (see, for example, Glenberg and Epstein 1987; Glenberg et al. 1987), suggest that even when our sense of familiarity or fluency is, in a sense, attributable to our state of knowledge, it can mislead our judgments. In their experiments, subjects were first asked to read some relatively technical material and then asked to rate the likelihood they could answer questions on that material. In general, the subjects proved to be poorly calibrated: the correlations between their judged comprehension and their actual ability to answer questions were surprisingly low. Of particular interest, however, is the finding that subjects with expertise in a given domain, such as physics or music, were even more poorly calibrated—across passages in that domain—than were subjects without expertise. Apparently, in studying text materials, the sense of familiarity or fluency resulting from one's general expertise in the relevant domain can be misattributed to an understanding of the specific content in a given passage.

The fact that Glenberg, Epstein, and their colleagues found that *all* subjects were quite poorly calibrated seems closely related to another factor that can predispose trainees to overconfidence: hindsight effects (Fischhoff 1975). Once an answer to a question is provided, or a solution to a problem demonstrated, subjects' judgments as to whether they *would have* been able to answer that question or solve that problem are seriously compromised. In general, in a wide range of situations, and for a variety of judgments, subjects exhibit hindsight biases that inflate their estimates. That is, once shown an answer or solution, subjects are more likely to think that they could have provided that answer or produced that solution.

Hindsight effects are of profound importance in real-world settings. Students and trainees, probably more often than not, judge their comprehension of the information or procedures to be learned on the basis of being exposed to that information or those procedures. Prior to an examination, for example, students tend to judge their level of preparedness, and how they should allocate their remaining time, on the basis of assessments they make while reading their notes or textbook—or, in certain courses, attending a review session and watching an instructor work problems of some type. The body of research on hindsight effects suggests that such assessments will tend both to be faulty and to overestimate one's comprehension or competence.

In effect, our being exposed to answers and solutions can deny us the type of subjective experience that might otherwise provide a valuable basis for judging our competence. Nelson and Dunlosky (1991; see also Dunlosky and Nelson 1992), for example, found that subjects, when looking at both the stimulus and response members of a pair studied earlier (such as "OAK-TURTLE"), could not predict with any substantial accuracy the likelihood that they would be able to recall the response member of that pair at a later time when cued with the stimulus member. Predictions made with only the cue present ("OAK-_?_") were much more accurate. Similarly, Jacoby and Kelley (1987) found that subjects' judgments of the relative difficulty of anagrams (such as "FSCAR-_?_"), as measured by the solution performance of other subjects, were less accurate if made with the solution present e.g., "FSCAR-SCARF").

In assessing our competence, therefore, it is important to test ourselves—that is, to give ourselves the opportunity to experience whether we can, in fact, produce answers and solutions to questions and problems. Under some circumstances, however, our sense of fluency in retrieving answers and solutions is itself an unreliable index, as I outline in the next section.

Retrieval Fluency

It is clearly not controversial to say that some things "come to mind" more readily than others and that, in general, facts and procedures that we know well are retrieved more fluently and rapidly than facts and procedures we know less well. It seems likely, then, that our perceived sense of the fluency of our access to information and procedures would influence various metacognitive judgments, such as our confidence that what we have retrieved is, in fact, correct. Aside from appeals to logic or introspection, there is now a solid research basis for such a conclusion. Recent experimental findings suggest that retrieval fluency is an important heuristic in a range of metacognitive judgments and decisions, including assessments of one's competence. Costermans, Lories, and Ansay (1992), for example, found that subjects' confidence in their answers to questions was negatively correlated with the time it took them to generate those answers, and, importantly that relationship held for incorrect answers as well as for correct answers. And the well-known research of Tversky and Kahneman on the *availability heuristic* (see, for example, Tversky and Kahneman 1974) illustrates that other judgments, such as estimating the probability or frequency of real-world events of some kind, are influenced by how easily we can retrieve representative instances of such events from our memories, even when that heuristic produces wrong, sometimes illogical, answers.

In a detailed analysis of retrieval fluency as a metacognitive index, Benjamin and Bjork (1996) define fluent retrieval in terms of three characteristics: latency, persistence, and amount. Things we know well, such as our own telephone number, are retrieved more quickly from our memories than other things we also know, but less well, such as a friend's telephone number. Things we know well also tend to be characterized by persistent access—that is, frequent and reliable recall across time and changes in setting or occasion. And things we know well tend to be characterized by recall that is rich or elaborated in the amount of information triggered by a given cue. (In terms of the speed, persistence, and elaboration of retrieval, of course, certain of us can reach the point where—from the perspective of our close friends and family members—our retrieval is too fluent in some domain, such as sports, politics, or motor vehicles.)

Our level of learning or knowledge, however, is only one of a number of possible contributors to fluent retrieval. In their analysis, Benjamin and Bjork (1996) distinguish four other "determinants" of fluent retrieval: frequency and recency of usage, episodic distinctiveness, cue sufficiency, and priming. Frequency and recency of usage play an important role because the act of retrieval is itself a learning event, in that the retrieval of information makes that information more retrievable in the future. An item of information that has been accessed frequently will tend to be characterized by fluent retrieval at the current point in time. Episodic distinctiveness refers to the collection of factors, such as salience, emotionality, and temporal isolation, that makes events distinct, hence more readily recallable. Cue sufficiency denotes the effectiveness of the cues guiding the retrieval of desired facts, information, or procedures. To the extent that those cues are underspecified—that is, associated with multiple items in memory—retrieval of any one item will be slowed, made less reliable, or both, owing to the competitive dynamics that characterize retrieval processes in human memory (cf. the cue overload principle; e.g., Watkins and Watkins 1975). Finally, as in the discussion of perceptual fluency, priming, refers to the effects of mere prior exposure of some target item. There is now abundant evidence that prior presentation of an item, even in a context nominally unrelated to some current task of interest, can increase the speed or likelihood that the item, among other possible items, is retrieved in response to a cue of some kind.

Misattributions of retrieval fluency Because it has multiple possible causes, retrieval fluency, like perceptual fluency, is open to misinterpretation. Kelley and Lindsay (1993), for example, demonstrated that one's confidence in a retrieved answer is susceptible to being misled by priming effects that make that answer, whether correct or incorrect, more retrievable. In their experiment, subjects were asked general-knowledge questions (e.g., "What was Buffalo Bill's last name?") and were also asked to rate their confidence in the correctness of their answers. When likely answers to those questions, whether correct ("Cody") or incorrect ("Hickok"), had been preexposed in an earlier phase of the experiment, the subjects not only produced those answers more frequently, they did so with increased confidence. Similarly, Shaw (1996) has demonstrated that the increased retrieval fluency that result from repeated retrieval of one's prior answers to questions can inflate one's confidence in those answers. In Shaw's experiments, subjects were asked to view slides of a staged crime scene and then asked to judge, from memory, whether certain objects had been in that scene. Whether their initial answers were correct or incorrect, subjects who were repeatedly asked to recall their prior answers, or simply asked to think about or "mull over" their prior answers, showed increased confidence in those answers.

Such results illustrate that we are prone to assume that fluent retrieval says something about our state of knowledge, even when that fluency is the product of manipulations that are orthogonal to accuracy, so to speak, such as priming (as in the Kelley and Lindsay 1993 experiment) or induced retrieval practice (as in the Shaw 1996 experiment). In general, of course, as mentioned earlier, the accuracy of retrieval processes and the fluency of those processes are highly correlated, but fluency appears to influence our confidence even when what is recalled is erroneous. An experiment by Koriat (1995) provides particularly strong support for that conclusion. Koriat developed a pool of general-knowledge questions that covaried in terms of the number of answers a given question elicited from subjects and the proportion of such answers that were accurate. Thus, in addition to questions that were high on both measures or low on both measures, there were questions that were high on one measure but low on the other measure (for example, the question "What is the capital of Australia?" is high in number of answers elicited, but low in the accuracy of those answers).

When Koriat asked his subjects to give feeling-of-knowing judgments, which took the form of asking them to predict the likelihood that they could pick the correct answer to a given question on a later recognition test, the subjects' predications were dominated by the frequency-of-access dimension. Thus subjects were highly confident of their ability to pick the right answers to questions that produced fluent retrieval, whether the consensual answers were correct or incorrect. And, for consensually incorrect cases, their confidence turned out, in fact, to be seriously misplaced: their subsequent recognition performance was somewhat below chance on such questions.

The experimental findings cited in this section are but a few of the findings that implicate retrieval fluency as an important—and occasionally misleading—factor in judgments and decisions of various types (for more complete reviews of the literature, see Benjamin and Bjork 1996 and Schwartz 1994). Of particular relevance to the concerns of the present chapter are those cases where the fluency of access to knowledge and skills at one point in time, such as during training, is used as a basis for predicting one's fluency of access at a later time.

Illusions of competence In general, of course, being able to access relevant knowledge and skills fluently at one point in time *is* a basis—indeed, a reasonable basis—for having confidence that we will be able to perform well in the future. As in the case of objective measures of performance, however, the subjective sense of fluent access to skills and knowledge during training can be a source of illusions of competence, and for much the same reasons. Our subjective sense of fluent access may be a product of local conditions that characterize the training context, but may not characterize some posttraining context of interest. The various conditions of training that inflate objective performance during training, but do not support long-term performance, such as those described earlier, are also likely to produce a misleading subjective sense of fluency during training. That is, once again, the learner is at risk of confusing retrieval strength with storage strength. Beyond such general considerations, certain specific factors can render fluent access during training an unreliable or misleading basis for confidence in our future performance. Among those factors are the following.

Changes in contextual cues It is common to distinguish between tests of *retention*, which gauge performance on the same task that was the target of training or instruction at a later time, and tests of *transfer*, which gauge performance on a different, but typically related, task. As Christina and Bjork (1994) have emphasized, however, every delayed test is a test of transfer to a greater or lesser extent. Even when there are no apparent changes in the nominal task, the conditions at the time of some posttraining test of performance may differ in subtle but significant ways from the conditions present during training. In addition to any changes in situational and environmental cues, one's mood state or body state may also differ in important ways, and the interpersonal or social context is likely to differ as well. All such changes can impair performance (see, for example, Bjork and Richardson-Klavehn 1989; Eich 1995; Smith 1988).

That performance is impaired by changes in stimulus conditions is hardly new to researchers and theorists. McGeoch (1932), for example, cited "altered stimulating conditions" as one factor in his influential three-factor theory of forgetting. And Tulving and Thomson's "encoding specificity principle" (1973) emphasizes the importance of the overlap between the cues present at the time of test and the cues present at the time of study. More recently, advocates of the influential situated-learning approach to instruction and education have argued that it is critical that learning be "situated" in the context of application (for a review and analysis of that approach, see Reder and Klatzky 1994).

There is abundant evidence, however, much of it deriving from evaluations of training programs or analyses of the causes of industrial or military accidents, that trainees do not fully understand how closely the level of performance they achieve during training may be tied to the particular stimulus conditions present during training. Individuals responsible for training programs are often surprised and dismayed by the evidence that their trainees, who had seemed to demonstrate proficient access to requisite skills and knowledge by the end of training, then prove unable to access those same skills and knowledge in real-world conditions that seem highly similar, if not identical, to the conditions of training.

Altered tasks A related factor responsible for impaired posttraining performance is that the actual tasks posed by a given posttraining real-world setting may differ in subtle, but sometimes significant, ways from the nominally identical, or highly similar, tasks embedded in training. In judging their own preparedness for real-world settings, therefore, trainees can become too reassured by their own good performance on the tasks embedded in training, either because they do not realize that the real-world setting of interest may pose altered versions of those tasks, or because they assume they are prepared for any such seemingly modest changes in task demands.

One reason that we, as learners, may be insufficiently sensitive to the differences between tasks is that we lack an appreciation for the multidimensional character of human memory. To the extent we think of our learned representations of skills and knowledge as varying along some unidimensional memory "strength," we are prone to assuming that good performance on one task ensures good performance on tasks that seem highly similar. Thus, for example, facile *recognition* of a correct response at a given point in time may provide confidence, often unjustified, in one's ability to *recall* that response at later time.

An experiment by Benjamin, Bjork, and Schwartz (1998) provides a good illustration that typical subjects may not, in fact, understand the complexity of their own memory—in particular, the differences between semantic and episodic memories. Replicating a procedure originally used by Gardiner, Craik, and Bleasdale (1973), Benjamin, Bjork, and Schwartz asked subjects a series of twenty relatively easy general-knowledge questions. Each subject's time to respond to a given question was recorded and, at the end of the experiment, the subjects were asked to free-recall as many as they could of their own prior responses. Benjamin, Bjork, and Schwartz augmented Gardiner, Craik, and Bleasdale's basic procedure by asking all subjects, after a given question was answered, to predict the likelihood that they would be able, at the end of the experiemtn, to free-recall that response.

Consistent with the findings of Gardiner, Craik, and Bleasdale (1973), answers generated with some difficulty initially—that is, associated with longer response times—were better recalled at the end of the experiment than were answers generated more readily. The most straightforward interpretation of this result is that the final test, where the questions were not presented again, was primarily a test of a subject's episodic memory, and the more involved or difficult the earlier process of answering a given question, the more salient and memorable that episode at the time of the final test. The subjects, however, predicted the opposite relationship. In direct opposition to their own subsequent free-recall performance, the subjects' predictions were a decreasing function of the time it took them to generate a given answer. Apparently, they failed to understand the differences between semantic memory and episodic memory representations, and that the final free-recall test would involve different demands and processes from those involved at the time of answering a given question. Rather, it appears that they simply interpreted initial retrieval fluency as an index of how well they "knew" a given answer and then assumed, as some kind of general rule, that the better they knew something, the better they would be able to recall that something at a later time.

Changes in relative fluency over time In any situation where there may be competing representations in memory—as in the case where multiple procedures are to be learned, each appropriate under some conditions and not appropriate under other conditions—using current fluency to predict later performance becomes an especially error-prone proposition. Owing to the counterintuitive mechanisms of interference and competition that characterize human memory, such as *unlearning* and *spontaneous recovery* (for reviews, see Anderson and Neely 1996; Postman 1971), current fluency of access to a given representation can be a particularly unreliable guide to one's later ability to access that representation. A procedure that is readily accessible from memory at one point in time, perhaps as a consequence of recent practice, can be difficult to access at a later time. Similar but inappropriate procedures learned or practiced in the interim can block or inhibit access to the procedure in question, as can spontaneous recovery of related but inappropriate procedures learned at some earlier time.

Where there are many things to remember, fluent access that is the product of recency of study or practice can be especially misleading. For materials and events of various types, there is a shift from recency to primacy over time. That is, at short delays, recent events or materials are most recallable, but at long delays first-learned materials or early events are typically the most recallable (for discussions of such effects, in animals as well as humans, see Bjork and Bjork 1992 and Wright 1989). In the immediate free recall of a list of worlds, for example, the items at the end of the list are the items recalled best (and first), but, at a delay, those items are typically the worst-recalled, whereas the items at the start of the list are then the best-recalled items (see, for example, Bjork 1975; Craik 1970).

The results of another experiment by Benjamin, Bjork, and Schwartz (1998) illustrates that subjects are largely unaware of such dynamics. Immediately after each of six lists of words, the subjects were asked to free-recall as many words as they could from that list. At the end of the experiment, they were asked to free-recall any and all words they could remember from *all* the lists they had studied. The novel aspect of the experiment was that the subjects, after each word they recalled on the immediate test, were asked to predict their likelihood of being able to recall that word again on the end-of-experiment test. Whereas the subjects predicted that the words they recalled with the most difficulty on the immediate test—that is, the last words they recalled from a given list—would be the worst-recalled on the final test, the words that were actually the worst-recalled on the final test were those the subjects recalled *first* during the tests of immediate free recall (which were mostly the easily accessible items from the recency portion of a given list).

Retrieval as memory modifer In addition to demonstrating the failure of subjects to realize that fluent retrieval may sometimes reflect transient influences, such as recency, Benjamin, Bjork, and Schwartz's results (1998) also illustrate that subjects do not understand that recalling is itself process that influences subsequence recall. More specifically, in addition to failing to realize—at least fully—that the act of retrieval can facilitate subsequent retrieval, subjects are apparently unaware that the more difficult or involved the process of retrieval, provided it succeeds, the greater its impact on subsequent recall or, conversely, that retrieval processes that are mostly effortless, such as "dumping" recency items from short-term memory on a test of immediate free recall, are also mostly useless in terms of fostering later long-term recall.

In terms of using current retrieval fluency as a basis for predicting later performance, the positive relationship between the difficulty of retrieval and its potency as a learning event poses an interesting complication for the learner. On the one hand, fluency of retrieval is one measure, if imperfect, of the extent of prior learning. On the other hand, difficult or nonfluent retrieval can enhance the likelihood of successful retrieval in the future. Thus, given those opposing considerations, using current retrieval fluency to predict future performance becomes a complex balancing act of sorts.

15.3 FOCUS ON APPLICATION

From an applied standpoint, it is obviously important to understand the relative effectiveness of alternative manipulations of the conditions of training, as measured by trainees' subsequent performance. It is also crucial, however, to understand how various conditions of training and practice may influence the learners' own subjective assessments of their learning. In fact, as my colleagues and I have argued elsewhere (Bjork 1994a; Jacoby, Bjork, and Kelley 1994), the reading we take of our own competence is arguably as important in many real-world contexts as is our actual competence. In certain special circumstances, for example, such as air traffic control, military actions, or nuclear plant operation—where on-the-job learning can be disastrous from a personal or societal standpoint—it can be imperative that we possess the skills and knowledge we think we possess. Individuals who overestimate their own competence can pose a unique hazard to themselves and others in such settings.

In less hazardous contexts, overconfidence may not pose the same perils, but it can result in errors, miscommunication, mistakes in judgments, and strained relationships among team members, colleagues, or coworkers. Also, in addition to the ways that the accuracy of our self-assessments influence our job performance, or the job performance of others, the readings we take of our current level of skill or knowledge influence our decisions and behaviors in other significant ways. As mentioned at the outset of this chapter, for example, whether we seek further study or practice, whether we plan for or aspire to certain occupations or competitive events, and even the character of our interpersonal interactions, depend on how we assess our current and potential competence.

Need to Create Desirable Difficulties for Learners

There are two reasons, then, why it is important to create desirable difficulties for the learner during training. One reason is that manipulations such as varying the conditions of practice, reducing feedback to the learner, sequencing materials and tasks to be learned in ways that cause "contextual interference," and spacing rather than massing practice sessions, enhance long-term retention and transfer. The other reason is that such manipulations also have the potential to inform the learners' own subjective experiences. In the context of the total training experience, such manipulations give learners the opportunity to experience the forgetting, errors, and mistakes that result from changes in task or contextual cues, from interference owing to interpolated or prior learning, from reduced feedback, and so forth. In an optimal training environment, the learners' successes become more informative as well—that is, become a more valid basis for confidence because they constitute more reliable evidence that skills and knowledge will remain accessible over time and in altered circumstances.

It is equally important, for similar reasons, to *avoid* conditions of training that enhance local performance but do not support long-term performance. Conditions such as massed practice, blocked practice, unvarying task demands, continuous feedback, and so forth act as crutches that support performance during training. In addition to not supporting actual learning, as measured by long-term retention and transfer, such conditions can also misinform learners' subjective experiences. That is, as emphasized in this chapter, such conditions can result in objective and subjective indices of performance that produce illusions of competence.

Potential to Upgrade Training

Typical real-world training programs are far from optimal. Partly because training personnel are themselves misled by trainees' performance during training, and partly because certain fundamental characteristics of humans as learners and rememberers are poorly understood, the types of counterproductive procedures that prop up performance during training are heavily represented in training programs. In general, then, the potential to upgrade training by introducing more productive conditions of training is substantial.

In this era, there is considerable emphasis on the use of physical or engineering technologies, especially computer and video technologies, to enhance training and instruction. Such tools, however, though potentially of great value in instruction and training, can be used well or poorly in terms of the considerations outlined in this chapter. Based on the accumulated body of psychological research, the real potential to upgrade training lies in structuring the conditions of training optimally, and that is as true in the design and use of hardware and software in computer-assisted instruction as it is in the context of more traditional instruction and training.

From a theoretical standpoint, we now understand better than ever before the encoding, storage, and retrieval processes that characterize human learning and memory. From an empirical standpoint, we also know more than ever before about the conditions that do and do not support learning, retention, and transfer. That knowledge provides a basis to revamp and upgrade training in some very substantial ways.

Attitudes and Assumptions That Impede Effective Training

In actual practice, however, there are some serious obstacles to introducing the types of innovations and changes necessary to upgrade training. Among those obstacles are institutional attitudes toward training and teaching that are counterproductive, and organizational structures that can impede communication and cooperation (for a discussion of these factors, see Bjork 1994b). More prevalent and subtle obstacles, however, are two assumptions about humans as learners. One such assumption is that performance equals learning. That is, the distinction between learning and performance, so fundamental to the optimization of training, does not tend to be understood and recognized in actual training contexts. Doing anything during training that increases errors or decreases the rate of improvement will not, therefore, tend to be well received—not by management, not by instructors, and not by trainees themselves.

A second assumption, which interacts in certain ways with the first assumption, is that differences in performance between individuals reflect innate differences in ability or aptitude. In business and other applied settings there is an unwarranted emphasis on individual differences. The role of aptitude is overappreciated and the role of training, practice, and experience is underappreciated. One consequence of that state of affairs is that time, energy, and resources are spent on selecting individuals who supposedly have the "right stuff," rather than on creating, optimizing, or evaluating training programs.

In combination, those two assumptions produce, among other things, a misunderstanding of the meaning and role of errors. There is not only a failure to realize that people learn by making and correcting mistakes, there is also a tendency to assume that errors and mistakes made during training reflect fundamental inadequacies of the learner. Those two assumptions also result in a failure to distinguish between testing as a pedagogical device and testing as assessment. Tests, as learning events, can retard forgetting, potentiate subsequent study, and inform the learners' subjective experiences. The primary use of testing, however, is assessment—to measure and evaluate individuals.

In short, being able to specify the changes and innovations that have the potential to upgrade training is only part of the battle. To actually upgrade training broadly and significantly requires not only an institutional and organizational culture that is receptive to change, but training and management personnel who better understand the characteristics and potential of humans as learners.

15.4 CONCLUDING COMMENT

In a rapidly changing and ever more complex world, the ultimate survival tool for individuals and organizations is knowing how to learn. To fully engage our remarkable capacity to learn requires an understanding of the complex, multifaceted, and counterintuitive processes that underlie our encoding, retention, and retrieval of information and skills. Our intuitions and introspections are clearly a poor guide to that understanding: for reasons that are not entirely clear, our intuitions about ourselves as learners are misguided, and the trials and errors of everyday living and learning do not seem to inform or correct our intuitions in any substantial way. Nor does custom and standard practice in training and education seem to be informed by any such understanding, as I have emphasized in this chapter. The potential to maximize our ability to learn and change does not lie in intuition or standard practice, but rather in the interaction of research and application, as illustrated by this volume.

REFERENCES

Adams, J. A., and Reynolds, B. (1954). Effect of shift in distribution of practice conditions following interpolated rest. *Journal of Experimental Psychology, 47,* 32–36.

Anderson, M. C., and Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork and R. A. Bjork (Eds.), *Handbook of perception and cognition.* Vol. 10, *Memory,* pp. 237–313. New York: Academic Press.

Baddeley, A. D., and Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21,* 627–635.

Battig, W. F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R. F. Thompson and J. F. Voss (Eds.), *Topics in learning and performance,* pp. 131–159. New York: Academic Press.

Begg, I., Duft, S., Lalonde, P., Melnick, R., and Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language, 28,* 610–632.

Benjamin, A. S., and Bjork, R. A. (1996). Retrieval fluency as a metacognitive index. In L. M. Reder (Ed.), *Implicit memory and metacognition: The Twenty-Seventh Carnegie Symposium on Cognition,* pp. 309–338. Hillsdale, NJ: Erlbaum.

Benjamin, A. S., Bjork, R. A., and Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General, 127,* 55–68.

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Eds.), *Information processing and cognition: The Loyola Symposium,* pp. 123–144. Hillsdale, NJ: Erlbaum.

Bjork, R. A. (1988). Retrieval practice and the maintenance of knowledge. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues.* Vol. 1, *Memory in everyday life,* pp. 396–401. New York: Wiley.

Bjork, R. A. (1994a). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing,* pp. 185–205. Cambridge, MA: MIT Press.

Bjork, R. A. (1994b). Institutional impediments to effective training. In D. Druckman and R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance,* pp. 295–306. Washington, DC: National Academy Press.

Bjork, R. A., and Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, and R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes,* vol. 2, pp. 35–67. Hillsdale, NJ: Erlbaum.

Bjork, R. A., and Richardson-Klavehn, A. (1989). On the puzzling relationship between environmental context and human memory. In C. Izawa (Ed.), *Current issues in cognitive processes: The Tulane Flowerree Symposium on Cognition,* pp. 313–344. Hillsdale, NJ: Erlbaum.

Blodgett, H. C. (1929). The effect of introduction of reward upon the maze performance of rat *University of California Publications in Psychology, 4*, 113–134.

Bower, G. H. (1972). Stimulas-sampling theory of encoding variability. In A. W. Melton and I Martin (Eds.), *Coding processes in human memory*. Washington, DC: Winston.

Carlson, R. A., and Yaure, R. G. (1990). Practice schedules and the use of component skills i problem solving. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 484-496.

Catalano, J. F., and Kleiner, B. M. (1984). Distant transfer in coincident timing as a function o variability of practice. *Perceptual and Motor Skills, 58*, 851–856.

Christina, R. W., and Bjork, R. A. (1991). Optimizing long-term retention and transfer. In D Druckman and R. A. Bjork (Eds.), *In the mind's eye: Enhancing human performance*, pp. 23–56 Washington, DC: National Academy Press.

Costermans, J., Lories, G, and Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 142–150.

Craik, F. I. M. (1970). The fate of primary memory items in free recall. *Journal of Verbal Learning and Verbal Behavior, 9*, 143–148.

Cuddy, L. J., and Jacoby, L. L. (1982). When forgetting helps memory: Analysis of repetition effects. *Verbal Learning and Verbal Behavior, 21*, 451–467.

Dempster, F. N. (1990). The spacing effect: A case study in the failure to apply the results of psychological research. *American Psychologist, 43*, 627–634.

Dempster, F. N. (1996). Distributing and managing the conditions of encoding and practice. In E. L. Bjork and R. A. Bjork (Eds.), *Handbook of Perception and Cognition*. Vol. 10, *Memory*, pp. 317–344. New York: Academic Press.

Dunlosky, J., and Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory and Cognition, 20*, 374–380.

Ebbinghaus, H. (1885). *Memory: A contribution to experimental psychology*. Translated from the German by H. A. Ruger and C. E. Bussenius, 1913. Reprint, New York: Dover, 1964.

Eich, E. (1995). Searching for mood dependent memory. *Psychological Science, 6*, 67–75.

Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review, 62*, 369–377.

Fischoff, B. (1975). Hindsight is not equal to foresight: The effects of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288–299.

Gardiner, J. M., Craik, F. I. M., and Bleasdale, F. A. (1973). Retrieval difficulty and subsequent recall. *Memory and Cognition, 1*, 213–216.

Ghodsian, D., Bjork, R. A., and Benjamin, A. S. (1997). Evaluating training *during* training: Obstacles and Opportunities. In M. A. Quinones and A. Dutta (Eds.), *Training for a rapidly changing workplace: Applications of psychological research*, pp. 63–88. Washington, DC: American Psychological Association.

Gick, M. L., and Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355.

Glenberg, A. M. (1979). Component-level theory of the effects of spacing of repetitions on recall and recognition. *Memory and Cognition, 7*, 95–112.

Glenberg, A. M. (1992). Distributed practice effects. In L. R. Squire (Eds.), *Encyclopedia of learning and memory*, pp. 138–142. New York: Macmillan.

Glenberg, A. M., and Epstein, W. (1987). Inexpert calibration of comprehension. *Memory and Cognition, 15*, 84–93.

Glenberg, A. M., Sanocki, T., Epstein, W., and Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General, 116*, 119–136.

Guthrie, E. R. (1952). *The psychology of learning*. New York: Harper and Row.

Homa, D., and Cultice, J. (1984). Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 83–94.

Hull, C. L. (1943). *The principles of behavior*. New York: Appleton-Century-Crofts.

Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior, 17*, 649–667.

Jacoby, L. L., Bjork, R. A., and Kelley, C. M. (1994). Illusions of comprehensions and competence. In D. Druckman and R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance*, pp. 57–80. Washington, DC: National Academy Press.

Jacoby, L. L., Allan, L. G., Collins, J. C., and Larwill, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 240–247.

Jacoby, L. L., and Kelley, C. M. (1987). Unconscious influences of memory for a prior event. *Personality and Social Psychology Bulletin, 13*, 314–336.

Kelley, C. M., and Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general-knowledge questions. *Journal of Memory and Language, 32*, 1–24.

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124*, 311–333.

Krueger, W. C. F. (1929). The effect of overlearning on retention. *Journal of Experimental Psychology, 12*, 71–78.

Landauer, T. K., and Bjork, R. A. (1978). Optimal rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, and R. N. Sykes (Eds.), *Practical aspects of memory*, pp. 625–632. London: Academic Press.

Lee, T. D., and Genovese, E. D. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport, 59*, 277–287.

Lee, T. D., Wishart, L., Cunningham, S., and Carnahan, H. (1997). Modeled timing information during random practice eliminates the contextual interference effect. *Research Quarterly for Exercise and Sport, 68*, 100–105.

Mannes S. M., and Kintsch, W. (1987). Knowledge organization and text organization. *Cognition and Instruction, 4*, 91–115.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39*, 352–370.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19,* 111–138.

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

Reder, L. M., and Klatzky, R. L. (1994). Transfer: Training for performance. In D. Druckman and R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance,* pp. 25–56. Washington, DC: National Academy Press.

Reder, L. M., and Ritter, F. E. (1992). what determines initial feeling of knowing? Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 435–451.

Schmidt, R. A. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. In G. E. Stelmach and J. Requin (Eds.), *Tutorials in motor neuroscience,* pp. 59–75. Dordrecht: Kluwer.

Schmidt, R. A., and Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207–217.

Schmidt, R. A., Young, D. E., Swinnen, S., and Shapiro, D. C. (1989). Summary knowledge of results fro skill acquisition: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 352–359.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin and Review, 3,* 357–375.

Schwartz, B. L., and Metcalfe, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1974–1083.

Shaw, J. S. (1996). Increases in eyewitness confidence resulting from postevent questioning. *Journal of Experimental Psychology, 2,* 126–146.

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

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

Smith, S. M., Glenberg, A. M., and Bjork, R. A. (1978). Environmental context and human memory. *Memory and Cognition, 6,* 342–353.

Smith, S. M., and Rothkopf, E. Z. (1984). Contextual enrichment and distribution of practice in the classroom. *Cognition and Instruction, 1,* 341–358.

Stelmach, G. E. (1969). Efficiency of motor learning as a function of intertrial rest. *Research Quarterly, 40,* 198–202.

Thorndike, E. L. (1914). *The psychology of learning.* New York: Teachers College.

Tolman, E. C. (1932). *Purposive behavior of animals and men.* New York: Century.

Tolman, E. C., and Honzik, C. H. (1930). Introduction and removal of reward and maze performance of rats. *University of California Publications in Psychology, 4,* 257–275.

Tversky, A., and Kahneman, D. (1974). Judgments under uncertainty: Heuristics and biases. *Science, 185,* 124–131.

Tulving, E., and Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Watkins, O. C., and Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory, 104,* 442–452.

Whitten, W. B., II, and Bjork, R. A. (1977). Learning from tests: Effects of spacing. *Journal of Verbal Learning and Verbal Behavior, 16,* 465–478.

Winstein, C. J., and Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 677–691.

Witherspoon, D., and Allan, L. G. (1985). The effects of prior presentation on temporal judgments in a perceptual identification task. *Memory and Cognition, 13,* 103–111.

Wright, A. A. (1989). Memory processing by pigeons, monkeys, and people. In G. H. Bower (Ed.), *The psychology of learning and motivation,* vol. 24, pp. 25–70. New York: Academic Press.

Attention and Performance XVII

Cognitive Regulation of Performance: Interaction of Theory and Application

edited by Daniel Gopher and Asher Koriat

This book is based on the papers that were presented at the Seventeenth International Symposium on Attention and Performance held at Beit Oren, Haifa, Israel, July 7–12, 1996

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

Attention and Performance

Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience, edited by David E. Meyer and Sylvan Kornblum, 1992

Attention and Performance XV: Conscious and Nonconscious Information Processing, edited by Carlo Umiltà and Morris Moscovitch, 1994

Attention and Performance XVI: Information Integration in Perception and Communication, edited by Toshio Inui and James L. McClelland, 1996.

Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application, edited by Daniel Gopher and Asher Koriat, 1998.

This book was set in Palatino on the Monotype "Prism Plus" PostScript Imagesetter by Asco Trade Typesetting Ltd., Hong Kong, and was printed and bound in the United States of America.

ISSN: 1047
ISBN: 0-262-09033-3