Journal of Experimental Psychology: General
2001, Vol. 130, No. 4, 701–725

0096-3445/01/$5.00 DOI: 10.1037//0096-3445.130.4.701

On the Fragility of Skilled Performance: What Governs Choking Under Pressure?

Sian L. Beilock and Thomas H. Carr
Michigan State University

Experiments 1–2 examined generic knowledge and episodic memories of putting in novice and expert golfers. Impoverished episodic recollection of specific putts among experts indicated that skilled putting is encoded in a procedural form that supports performance without the need for step-by-step attentional control. According to *explicit monitoring* theories of choking, such proceduralization makes putting vulnerable to decrements under pressure. Experiments 3–4 examined choking and the ability of training conditions to ameliorate it in putting and a nonproceduralized *alphabet arithmetic* skill analogous to mental arithmetic. Choking occurred in putting but not alphabet arithmetic. In putting, choking was unchanged by dual-task training but eliminated by self-consciousness training. These findings support explicit monitoring theories of choking and the popular but infrequently tested belief that attending to proceduralized skills hurts performance.

Why does the execution of a well-learned skill fail under pressure? Research investigating skill and expertise has produced a number of important findings regarding the variables that mediate optimal skill performance (Allard & Starkes, 1991; Anderson, 1982, 1987; Ericsson, Krampe, & Tesch-Romer, 1993; Keele, 1986; Logan, 1988; Reimann & Chi, 1989). Nevertheless, the phenomenon of "choking under pressure" remains unexplained—and feared by skilled performers across many domains. *Performance pressure* has been defined as an anxious desire to perform at a high level in a given situation (Hardy, Mullen, & Jones, 1996) and is thought to vary as a function of the personally felt importance of a situation (Baumeister, 1984). *Choking,* or performing more poorly than expected given one's level of skill, tends to occur in situations fraught with performance pressure. This phenomenon seems particularly visible in sensorimotor or action-based skills, where it has garnered interest in both experimental and real-world settings. People often speak of the "bricks" in basketball free throw shooting or the "yips" in golf putting, and a majority of the experimental research on choking done to date has used sensorimotor tasks of one kind or another (Baumeister, 1984; Lewis & Linder, 1997; Masters, 1992).

Two competing theories have been proposed to account for decrements in skilled performance under pressure. *Distraction* theories propose that pressure creates a distracting environment that shifts attentional focus to task-irrelevant cues, such as worries about the situation and its consequences (Wine, 1971). This shift of focus changes what was single-task performance into a dual-task situation in which controlling execution of the task at hand and worrying about the situation compete for attention. *Self-focus* theories (perhaps more appropriately termed *explicit monitoring* or *execution focus* theories, as they are concerned with attention to skill execution) suggest that pressure raises self-consciousness and anxiety about performing correctly, which increases the attention paid to skill processes and their step-by-step control (Baumeister, 1984; Lewis & Linder, 1997). Attention to execution at this step-by-step level is thought to disrupt well-learned or proceduralized performances (Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997; Masters, 1992).

Distraction and explicit monitoring theories appear to be competing alternatives—indeed, they are complete opposites in their proposed mechanisms. However, it is important to note that they may have different domains of applicability and hence could turn out to be complementary rather than mutually exclusive. Distraction theory holds that the mechanisms of choking operate on task control structures that are attended during performance. Thus, under distraction theory, breakdowns under pressure are most likely in skills that rely on working memory for storage of decision and action-relevant information that might be vulnerable to corruption or forgetting as a result of dual-task interference. This calls to mind skills based on fact retrieval as possible test cases. In contrast, explicit monitoring theory suggests that the mechanisms of choking operate on task control structures that are proceduralized—based on mental or motor programs that run largely unattended, without the services of working memory, and might best remain outside the scrutiny of introspection. This calls to mind sensorimotor skills as test cases.

Given these differences in potential domain of applicability, our first two experiments were aimed at identifying a particular skill that had the right properties to be susceptible to choking according to one of these theories but not the other. We chose golf putting, which is a complex sensorimotor task that is thought to become proceduralized with practice and hence falls into the domain of explicit monitoring theory. Because a proceduralized skill ought not to require constant on-line attentional control (e.g., Fitts & Posner, 1967; Proctor & Dutta, 1995), it should be relatively robust

---

Sian L. Beilock, Departments of Psychology and Kinesiology, Michigan State University; Thomas H. Carr, Department of Psychology, Michigan State University.

Correspondence concerning this article should be addressed to Sian L. Beilock, Department of Psychology, Michigan State University, Psychology Research Building, Room 236, East Lansing, Michigan 48824. Electronic mail may be sent to beilocks@pilot.msu.edu.

against conditions that draw attention away from the primary task as in distraction theory. However, this type of skill should be sensitive to the kind of attention-induced disruptions of fluent execution envisioned by explicit monitoring theory.

To confirm the proceduralized status of golf putting, in Experiments 1 and 2 we compared reports of generic, schematic, or prescriptive knowledge about putting with episodic memories of particular putts in expert and novice golfers. The goal was to document a particular property of the cognitive substrate of this sensorimotor skill—the declarative accessibility, or openness to introspection, recollection, and report, of the skill's processes and procedures at different levels of expertise. In particular, we sought to use as a diagnostic tool the well-documented dependence of explicit episodic memory on the presence of attention (e.g., Craik, Govini, Naveh-Benjamin, & Anderson, 1996). If golf putting becomes proceduralized with practice and, as a consequence, task control structures are largely unattended during skill execution, then episodic memory for the step-by-step unfolding of particular instances of performance should be impoverished. Observing such a pattern would earmark practiced golf putting as a skill that should be susceptible to choking under pressure according to explicit monitoring theory.

After documenting the proceduralized status of practiced golf putting, in Experiment 3 we trained novices to an asymptotic level of achievement and then created a high-pressure test situation intended to induce choking. Participants performed either our chosen sensorimotor task of golf putting or a comparison task whose practiced control structure has already been shown to depend on fact retrieval rather than proceduralized motor programs. The comparison task was Zbrodoff and Logan's (1986) *alphabet arithmetic* task.

Training took place under one of three different regimens. Choice of regimens followed what we called a "vaccination strategy," intended to test the theories of choking by determining whether practice at dealing with the particular causal mechanism proposed by each theory would reduce choking in performers who would have been likely to choke had the training not been received. One regimen was ordinary single-task practice, which provided a baseline measure of the occurrence of choking. The other two regimens exposed performers to the particular aspects of high-pressure situations that have been proposed by the two theories of choking to cause performance decrements. In the "dual-task distraction" regimen, practice took place under dual-task conditions (while monitoring an auditory word list for a target word) in order to expose performers to being distracted from the primary task by execution-irrelevant activity in working memory. In the "execution-oriented self-consciousness" regimen, practice took place while being videotaped for subsequent analysis by experts in order to expose performers to having attention called to themselves and their performance in a way intended to induce explicit monitoring of skill execution.

Finally, Experiment 4 replicated and extended Experiment 3. The goal was to test the distraction and explicit monitoring theories' predictions at lower levels of practice in the golf putting task. Most conceptions of skill acquisition, including motor program theories applicable to golf putting, propose that early in learning, performance is supported by unintegrated control structures that are held a few steps at a time in working memory (Anderson, 1987, 1993; Fitts & Posner, 1967; Keele, 1986; Keele & Summers, 1976; Proctor & Dutta, 1995). With practice, control is thought to evolve toward the integrated procedures that are the objects of the explicit monitoring theory. One might imagine, then, that in tasks that follow such a developmental trajectory, choking due to distraction might be observed early in learning whereas choking due to explicit monitoring would occur in a more practiced state. Experiment 4 explored this idea by imposing high-pressure tests on golf putting performance at two different points in practice.

We now turn to a more detailed discussion of differences in the knowledge representations controlling the execution of various skills at different levels of expertise and of how these differences may aid in our examination and understanding of the two competing theories of choking under pressure.

Generic, Episodic, and Procedural Skill Representations

Skill knowledge can be declaratively accessible in two different forms. *Generic* knowledge captures schema-like or prescriptive information about how a skill is typically done. *Episodic* knowledge, on the other hand, captures a specific memory—an autobiographical record of a particular performance. According to current theories of skill acquisition and automaticity, changes in expertise should affect these two types of declaratively accessible representations very differently.

First, declaratively accessible generic knowledge should increase with increasing expertise—experts have more explicitly available general knowledge about the domain in which they are skilled than do their novice counterparts (for reviews, see Proctor & Dutta, 1995; Van Lehn, 1989). It would be shocking to discover that experts could not describe the dos and don'ts of their skill in as much detail or explain its ideal execution as competently as novices. Thus, experts' off-line generic or prescriptive accounts of their skill should provide a more extensive and systematic chronicle of how, in general, performance should be accomplished than the generic accounts of novices.

Second, declaratively accessible episodic memories of any particular performance should decrease with increasing expertise. Why should this be? It is widely believed that highly practiced, overlearned performances are *automated*—meaning they are controlled in real time by procedural knowledge that requires little attention, operates largely outside of working memory, and is substantially closed to introspection (Anderson, 1987, 1993; Fitts & Posner, 1967; Keele & Summers, 1976; Kimble & Perlmuter, 1970; Langer & Imber, 1979; Proctor & Dutta, 1995; Squire & Knowlton, 1994). Because of the well-established relation between attention and episodic memory, this belief carries implications for recollecting one's performances. In both short-term memory (Daneman & Carpenter, 1980; Muter, 1980; Peterson & Peterson, 1959; Posner & Rossman, 1965) and long-term memory (Craik et al., 1996; Naveh-Benjamin, Craik, Guez, & Dori, 1998), diverting or reducing the amount of attention paid to material being encoded for storage reduces subsequent explicit memory for that material. The impact of reducing attention is greatest in recall and is present in cued recall and recognition as well. To the extent that practiced tasks are indeed carried out with less attention to processes, procedures, and the control structures that govern them, real-time performance ought to leave impoverished episodic memories of the performance's execution.

In contrast, the relatively unpracticed performances of novices are thought to be controlled by declarative knowledge that is held in working memory and attended to step-by-step during performance (Anderson, 1987, 1993; Fitts & Posner, 1967; Kimble & Perlmuter, 1970; Proctor & Dutta, 1995; Squire & Knowlton, 1994). Attending to such knowledge should leave an explicitly retrievable episodic record of task execution—a declaratively accessible memory of the performance as an autobiographical experience that includes the step-by-step operations by which the performance was implemented.

"Expertise-Induced Amnesia"

Thus, current theories of skill acquisition and automaticity suggest that increasing expertise through practice will create a kind of domain-specific amnesia. If a skill is controlled by declarative knowledge that is attended to during performance, episodic memory for skill execution processes should be explicitly retrievable. However, if a skill is supported by procedural knowledge that automates real-time performance, then episodic memory for this performance should be minimized.

The idea of "expertise-induced amnesia" may seem uncontroversial to some investigators of skilled performance. To others, however, a problem will come immediately to mind: There is well-known evidence suggesting that expertise serves to enhance episodic recollection, not degrade it. For example, in their classic chess study, Chase and Simon (1973) found that chess masters were better able to recall briefly presented chess positions than were less experienced players (for confirmatory data, see De Groot, 1946/1978; and for similar results from computer programmers, see Soloway & Ehrlich, 1984). Analogous evidence comes from studies of reading. Real-time deployment of world knowledge, creating superior comprehension of a situation described in a narrative text, leads to better recall of the text's wording (Bransford & Johnson, 1972; Dooling & Christiaansen, 1977; McCandliss & Carr, 1994, 1996). In light of such evidence, one anonymous reviewer of an earlier attempt to report this work called the idea of expertise-induced amnesia "otherworldly" and "patently false," claiming instead that "experts have exquisite episodic recall of the most arcane minutiae in their area of competence" (personal communication, August 24, 1999).

However, problems exist in using results such as those mentioned above as evidence against the prediction of expertise-induced amnesia. The chess studies focused on memory for the kinds of stimuli that are operated on by chess players, not memory for the operations themselves. That is, experts were asked to recreate the positions of specific pieces on the board. Experts were not asked for the steps or processes by which the situation was assessed, how a move appropriate to that stimulus configuration was chosen, or how a chosen move was physically implemented. The same applies to studies of reading, where people able to deploy greater world knowledge were asked to remember the stimulus material they read, not the sequence of reading operations that took place. Thus, the above-mentioned studies can be taken to support the notion that experts have better episodic recollection for the stimuli to which they apply their knowledge. However, these studies do not demonstrate that experts have superior recollection for the sequence of cognitive processes involved in formulating specific plans of action or the sequence of cognitive processes by which actions are implemented in real time. For this reason, it remains a reasonable idea, despite the existing literature on experts' episodic memories, that because expert knowledge runs automatically during real-time skill execution, experts may neither attend to nor later remember the step-by-step unfolding of their performances.

Consistent with this possibility is the finding from the chess literature that in both on-line and retrospective verbal-report protocols, experts report having explicitly considered fewer alternatives in making any given move than do novices. The experts report that the best move or a small number of good moves just popped into their heads, whereas the novices report a serial process of generating and evaluating several possible moves in succession (Ericsson & Smith, 1991).

Furthermore, it is not always true that highly practiced experts demonstrate superior episodic memory for the stimuli on which they have operated. Fisk and Schneider (1984) studied the acquisition of expertise at searching through arrays of visually presented words for members of a target category. They found that after a great deal of practice, during which speed and accuracy of finding targets greatly increased and sensitivity of performance to the number of words in each array greatly decreased, recognition memory for the words that had been searched through was markedly worse than it had been at lower levels of practice. Fisk and Schneider argued that practice automated performance and that automating performance increased real-time skill but decreased subsequent episodic memory. In light of Fisk and Schneider's finding, it should be noted that the literature on skill acquisition and automaticity from which we derived the prediction of expertise-induced amnesia has been dominated by studies of the speeded performance of reaction time tasks with significant sensorimotor components. In contrast, much of the work on expertise that suggests good rather than poor episodic memory has focused on cognitive tasks that are based on a great deal of factual knowledge and whose real-time sensorimotor demands are minimal (e.g., chess, computer programming, physics problem solving). The memorial results of performing such fact-reliant cognitive tasks may be different from those of tasks that rely more on sensorimotor knowledge. More generally, fact-reliant tasks and sensorimotor tasks may diverge in the nature of their underlying representations and control structures and hence may differ in many ways, not just memorial consequences. An argument for such differences in underlying representations and control structures has been made by Klapp, Boches, Trabert, and Logan (1991), to which we return later in the article.

Experiments 1 and 2: Declaratively Accessible Knowledge of Golf Putting

This brings us to the first two experiments in the present study, which document the prediction of expertise-induced amnesia in the sensorimotor task of golf putting. Putting was chosen because it is a complex task in which considerable time and effort is required to become an expert performer. Even at the highest levels, putting is not easy and success depends heavily on extensive past experience. This is in some ways similar to chess, where expert chess masters hone their skills over a long period of time, developing a large knowledge base consisting of many relevant chess piece configurations and game scenarios (De Groot, 1946/1978). Nevertheless, putting is a sensorimotor task in a way that chess is not. In addition, putting's discrete nature enables straightforward trial-by-trial measurement of accuracy, so that differences in expertise can be readily verified.

Experiment 1

In Experiment 1 expert golfers' generic knowledge of golf putting and episodic recollection of specific putts were compared to the generic knowledge and episodic recollection of novice golfers within the context of a laboratory golf putting task. If on-line well-learned golf putting is supported by procedural knowledge, as theories of sensorimotor skill acquisition would predict, then expert golfers should give longer, more detailed generic descriptions of the steps involved in a typical putt compared with the accounts given by novices, but shorter, less detailed episodic recollections of a particular putt. Because proceduralization reduces the need to attend to the specific processes by which skill execution unfolds, experts' episodic recollections of step-by-step real-time performance should be impoverished.

Method

Participants

Participants ($N = 48$) were undergraduate students enrolled at Michigan State University and consisted of intercollegiate golf team members ($n = 16$), intercollegiate athletes with no golf experience ($n = 16$), and introductory psychology students with no golf experience ($n = 16$). An equal number of male and female participants were recruited from each of the three populations. The two groups of novices were included in order to examine the possibility that the intense training and practice engaged in by elite athletes may alter their strategic approach to skill acquisition, even in a new domain outside their already-acquired expertise, causing differences in performance and knowledge representation in comparison with nonathletes.

Procedure

After giving informed consent and filling out a demographic sheet concerning previous golf experiences, participants were told that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were instructed that the object of the task was to putt a golf ball as accurately as possible, making it stop at a target located 1.5 m away, marked by a square of red tape on a carpeted indoor putting green (3 × 3.7 m). A standard golf putter and golf ball were supplied. All groups participated in identical pretest, practice, and posttest conditions, though the participants were not made aware of the separate conditions. To the participant, the golf putting task appeared to involve three blocks of putts with a short break after each block during which a questionnaire was filled out.

Pretest condition. Participants were set up 1.5 m from the target. They were asked whether they preferred to putt right-handed or left-handed and were then given the appropriate putter. Participants took a series of 20 putts. After completing the putts, participants filled out a questionnaire eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Practice condition. Participants were again set up 1.5 m from the target. Participants took a series of 30 putts. After completing the putts, participants filled out an identical questionnaire to the one that they had previously filled out in the pretest condition eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Posttest condition. Participants were set up 1.5 m from the target. Participants then took a series of 20 putts. Immediately following the posttest condition, participants filled out a questionnaire designed to access their episodic recollection of the last putt they had just taken (Appendix A, second paragraph).

Results

Putting Performance

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target at which the ball stopped after each putt. The mean distance from the target of the last 10 putts in the pretest condition was used as a measure of pretest golf putting skill. The mean distance from the target of the middle 10 putts in the practice condition was used as a measure of practice putting skill. The mean distance from the target of the last 10 putts in the posttest condition was used as a measure of posttest golf putting skill. Means and standard errors for putting performance appear in Figure 1.

As Figure 1 makes clear, golf team members showed superior putting performance in comparison to the two novice groups, who did not differ. This was true both before and after the practice phase, during which the three groups all improved by approximately the same amount. This pattern was confirmed by a 3 (undergraduate, athlete, golf team) × 2 (pretest, posttest) analysis of variance (ANOVA), which revealed significant main effects of experience, $F(2, 45) = 16.23$, $p < .001$, $MSE = 56.88$, and test, $F(1, 45) = 29.21$, $p < .001$, $MSE = 14.91$, with no interaction ($F < 1$).

Thus the expertise of the golf team members transferred substantially to the somewhat novel task demands of making the ball stop on a target rather than drop into a hole. However, the golfers did improve with practice, indicating that there were still some elements of the present task left for them to learn. In contrast to the skill displayed by the golfers, the nongolf sensorimotor expertise of the athlete group did not transfer to putting whatsoever. Athletes enjoyed no advantage in putting accuracy over the nonathlete undergraduates at any point in practice.

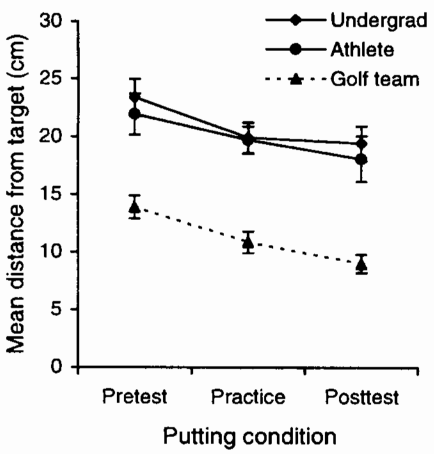

Figure 1. Mean (± *SE*) distance from the target at which the ball stopped after each putt in the pretest, practice, and posttest conditions for each group. Undergrad = undergraduate.

Generic and Episodic Memory Protocols

Questionnaire responses were analyzed quantitatively, in terms of the number of golf putting steps included in each type of protocol, and qualitatively, in terms of the relative frequencies of different categories of steps.

Quantitative analysis. Three expert golfers and the how-to golf putting book *Classic Instruction in Golf* (Jones, Davis, Crenshaw, Behar, & Davis, 1998) were used in establishing a master list of steps involved in a successful golf putt (Appendix B). The statements in each participant's protocol were compared with this master list. If a step given by a participant referred to the same action or the same biomechanism as a step on the master list, it was counted as one step. For example, the step given by a participant, "I swung the club back behind me," and Step 13 on the expert golfer list, "Backswing—swing the club straight back," were coded as a match because they both refer to the same action (i.e., taking a backswing). Similarly, the step given by a participant, "I kept my hips still," and Step 21 on the list developed by the expert golfers, "Head/trunk/hips/legs—should remain still during the stroke," were deemed a match because they both refer to the same biomechanism (i.e., motion of the hips). If two steps given by participants both described one step on the list developed by the expert golfers, they were combined and counted as one step. For example, if a participant reported the two steps "I held the putter with two hands" and "My right hand was above my left hand," these steps were combined to match the step on the list developed by the expert golfers that referred to the grip of the putter. If a step given by a participant did not match a step on the master list, yet did refer to a necessary part of the participant's putting process (e.g., "I brushed my hair out of my face so I could see the target"), it was counted as a step. Because the master list was thorough and detailed, these "nonmatch steps" were quite rare. Finally, if a step given by a participant did not match a step on the master list, and was not part of the putting process itself (e.g., "I thought about the fact that I needed golf lessons"), it was not counted as a step. Although such nonprocess commentary is legitimately part of the autobiographical record, it is not part of the specific object of prediction in testing for expertise-induced amnesia. However, nonprocess commentary was also quite rare and, if included, would not have changed the results in any way.

The order in which the steps were recorded was not taken into account in determining the number of steps given by participants. Two experimenters independently coded the data. Interexperimenter reliability was extremely high ($r = .97$). Table 1 gives representative generic and episodic golf putting protocols for all three participant groups.

Table 2 and Figure 2 present the results. Mean number of steps did not differ significantly between the two generic protocols in any of the groups, as confirmed by a 3 (undergraduate, athlete, golf team) × 2 (first generic protocol, second generic protocol) ANOVA in which the main effect of test and the interaction of Expertise × Test produced $F\text{s} < 1$. The second generic protocol was used in a 3 (undergraduate, athlete, golf team) × 2 (second generic protocol, episodic protocol) ANOVA to compare the lengths of the generic and episodic protocols produced at each level of expertise. The analysis revealed an interaction of Expertise × Protocol type, $F(2, 45) = 24.30$, $p < .001$, $MSE = 2.10$. Direct comparisons of the number of generic versus episodic steps within each group showed that the undergraduates gave significantly more steps in their episodic than in their generic protocols, $t(15) = 4.29$, $p < .001$. The athletes produced a difference in the same direction as the undergraduates, but it was not significant, $t(15) = 0.29$, $p < .78$. In contrast, golf team members gave significantly more steps in their generic than in their episodic protocols, $t(15) = 4.70$, $p < .001$. Thus, as can be seen in Figure 2, golfing expertise was associated with longer generic descriptions and shorter episodic recollections.

Qualitative analysis: Types of steps. The first qualitative analysis divided steps into three categories (see Table 3). Assessment or planning referred to deciding how to take a particular putt and what properties the putt ought to have. Examples are "read the green," "read the line" (from the ball to the hole or target), "focus on the line," and "visualize the force needed to hit the ball." Mechanics or execution referred to the components of the mechanical act that implements the putt. Examples are "grip the putter with your right hand on top of your left," "bring the club straight back," and "accelerate through the ball," all of which deal with the effectors and the kinesthetic movements of the effectors required to implement a putt. Ball destinations or outcomes referred to where the ball stopped or landed and hence to degree of success. A 3 (undergraduate, athlete, golf team) × 2 (generic protocol, episodic protocol) ANOVA was conducted on the number of steps given in each of these three categories.

The analysis of assessment produced a significant interaction between expertise and type of protocol, $F(2, 45) = 14.56$, $p < .001$, $MSE = 1.07$, which is displayed in the left panel of Figure 3. Assessment steps appeared more often in the generic descriptions of golf team members than anywhere else. A simple effects test confirmed a difference among groups in the generic protocol, $F(2, 45) = 13.75$, $p < .001$, $MSE = 2.14$, and Fisher's least significant difference (LSD) test showed that the golf team gave significantly more assessment steps in their generic descriptions than did either the undergraduates or the athletes, who did not differ. Furthermore, the golf team gave more assessment steps in their generic descriptions than they did in their episodic recollections, $t(15) = 4.90$, $p < .001$, whereas the undergraduate and athlete groups did not differ in the number of assessment steps included in the two kinds of protocols, $t(15) = 0.64$ and $t(15) = 0.00$, respectively ($ps > .10$).

As an adjunct to the analysis of assessment, those steps that involved mental imagery (i.e., imagining some aspect of how a putt ought to look or feel before executing the action) were counted. Mental imagery is a topic of considerable interest in sports psychology and has been defined in that literature as "the imagined rehearsal of skill processes, procedures, and possible outcomes prior to task performance" (Woolfolk, Murphy, Gottesfeld, & Aitken, 1985). In the undergraduate group, 0.0% of generic steps and 0.7% of episodic steps referred to mental imagery. In the athlete group, 0.7% of the generic steps and 0.0% of the episodic steps referred to imagery. In the golf team group, 7.0% of the generic steps and 2.0% of the episodic steps referred to imagery. Thus almost all of the reports of imagery were from golfers, and most of these were part of the generic descriptions.

One might worry that the experts' exclusion of assessment steps from their episodic recollections was merely an artifact of our very simple and highly repetitive situation, in which assessment was not much needed by the time episodic memory was measured, which

Table 1
Representative Generic and Episodic Putting Descriptions

Generic putting description	Episodic putting description
Undergraduates	
1. Feet apart	1. Feet apart
2. Lean forward	2. Knees not locked
3. Aim ball	3. Leaning forward
4. Swing	4. Positioning hands
	5. Lining putter up with the ball
	6. Look at the hole
	7. Aim ball
	8. Swing
	9. Follow through
Athletes	
1. Estimating distance	1. Estimate distance to target
2. Bending knees	2. I placed my feet a comfortable distance apart
3. Looking back at target	3. Bent my knees
4. Relaxed backswing	4. Line up the putter with the target
5. Follow through	5. Slowly pulled the putter back
	6. Follow through lightly
	7. Using straight arms
Golf team members	
1. Walk behind the ball and look at the putt	1. Look up at putt
2. Read the green from behind the ball	2. Place putter behind ball with the head square at the target
3. Make sure nothing is in its path	3. Look at target
4. Look at distance of putt	4. Look at putter and ball
5. Pick a target to aim at	5. Take putter back
6. Place putter behind ball lined up with the target	6. Swing through ball
7. Move putter closer to you of the ball and line up at target	7. Look up at target
8. Take a practice swing	
9. Move putter back to behind the ball	
10. Line up squarely with target	
11. Move feet and body square with putter head	
12. Look at target	
13. Look down at the ball	
14. Swing the putter head straight back	
15. And straight through	
16. Look up at ball	

was after the 70th putt. To guard against this alternative explanation, we performed a reanalysis of the golf team members' protocols, dropping from each generic protocol all assessment steps that (a) did not appear in the corresponding episodic protocol and (b) were likely to be unnecessary once 69 putts had been taken in our laboratory situation. Excluded were steps such as "read the green" and "read the lie of the ball," because neither the green nor the lie of the ball changed during the experiment. Steps such as "taking aim," that would always be necessary in order to execute a putt, were maintained. This reanalysis of assessment produced the same-shaped interaction between expertise and type of protocol as the original, $F(2, 45) = 3.34$, $p < .05$, $MSE = 0.68$.

Table 2
Questionnaire Responses: Number of Steps (Experiment 1)

	Generic 1		Generic 2		Episodic	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*
Undergraduate	5.19	0.39	5.63	0.38	7.69	0.58
Athlete	5.94	0.54	6.25	0.55	6.75	0.77
Golf team	8.63	0.94	8.44	0.97	5.56	0.60

Turning to mechanics, this analysis also produced an interaction between expertise and protocol type, $F(2, 45) = 7.96$, $p < .001$, $MSE = 1.68$, but of a very different nature, as can be seen in the middle panel of Figure 3. Undergraduates gave significantly more mechanics steps in their episodic descriptions than in their generic descriptions, $t(15) = 3.34$, $p < .005$, and athletes produced a nonsignificant difference in the same direction, $t(15) = 0.36$. In contrast, the golfers gave more mechanics steps in their generic descriptions than in their episodic descriptions, though the difference was only marginally significant, $t(15) = 1.75$, $p < .10$. The greater number of mechanics steps in the episodic protocols of undergraduates compared with golfers was significant, $t(30) = 2.13$, $p < .05$. In sum, mechanics was a category of steps that for experts tended to appear more often in generic descriptions than in episodic descriptions, but for novices appeared more often

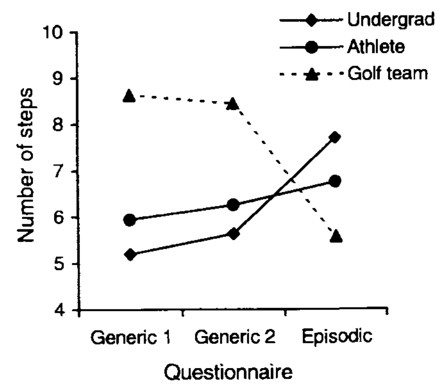

Figure 2. Mean number of steps for the first and second generic questionnaires and the episodic questionnaire for each group. Undergrad = undergraduate.

in episodic descriptions than in generic descriptions. Athletes were intermediate.

The analysis of ball destinations produced two main effects but no interaction. Overall, more ball destinations were included in the episodic protocols than in the generic protocols, $F(1, 45) = 9.36$, $p < .004$, $MSE = 0.22$, and the golf team included more destination information than either the undergraduates or the athletes, $F(2, 45) = 3.98$, $p < .026$, $MSE = 0.21$. Thus, as shown in the right panel of Figure 3, ball destinations were more likely to appear in the episodic recollections of experts than anywhere else, though even there they were relatively infrequent, accounting for only 13% of the steps that were included.

A second qualitative analysis looked for steps present in both protocols that referred to the same action or biomechanism but provided more detail in one type of protocol than in the other. For instance, a step in the episodic description of one participant was stated as "I positioned my feet so that they were shoulder length apart." This was scored as an elaboration of a step in the same participant's generic description that was stated as "feet positioning." Overall, elaborations were more likely to occur in episodic descriptions relative to generic descriptions than vice versa. Therefore, greater detail in the episodic description was scored as a "positive" elaboration whereas greater detail in the generic description was scored as a "negative" elaboration. In the undergraduate group, 14% of the steps in the episodic descriptions were elaborations of steps in the generic descriptions. In the athlete group, −1% of the episodic steps were elaborations of generic steps. In the golf team group, 5% of the episodic steps were elaborations of generic steps. A one-way ANOVA on these data produced a significant effect of expertise, $F(2, 45) = 3.53$, $p < .038$, $MSE = 1.23$. Fisher's LSD test showed that undergraduates elaborated their episodic recollections relative to their generic descriptions significantly more often than the athletes and marginally more often than the golf team ($p < .06$). Athletes and golfers did not significantly differ from one another.

Although the athlete group consisted of novice golfers, their elaborations were more similar to the golf teams' than to the undergraduates'. Similar to the athletes' pattern of mechanics steps, the athletes' pattern of elaborations suggests that sport training and participation lead athletes to approach novel skill situations in certain ways that resemble the approach of more experienced performers. This occurs despite the fact that the athletes' measured achievements in golf putting performance are no better than those of other novices.

Discussion

The results of Experiment 1 demonstrate an effect of level of expertise on the content of generic knowledge and episodic memories of golf putting. Experts gave longer, more detailed generic descriptions of the steps involved in a typical putt compared with the accounts given by novices and shorter, less extensive episodic recollections of a particular putt. These quantitative differences were accompanied by qualitative differences between experts and novices in the nature of the steps included in each type of description.

Expert golfers' generic descriptions dealt considerably more with assessing and planning a putt than did novices'. This finding is consistent with research on expert performers across a wide range of task domains (Chi, Feltovitch, & Glaser, 1981; Lesgold et al., 1988; Priest & Lindsay, 1992; Proctor & Dutta, 1995; Voss & Post, 1988). In areas as diverse as physics problem solving and radiological X-ray diagnosis, experts spend more time evaluating a situation and deciding how to approach or formulate a problem before they actually begin to work on it than do novices.

Expert golfers' episodic recollections included fewer assessment steps than did their generic descriptions. Expert golfers also made fewer references to putting mechanics in their episodic recollections than did novices. This pattern follows the prediction of expertise-induced amnesia derived from current theories of skill acquisition and automaticity. According to this idea, experts' extensive generic knowledge of putting is declaratively accessible during off-line reflection, but it is not used during real-time per-

Table 3
Assessment, Mechanic, and Destination Descriptions by Questionnaire Type—Experiment 1

	Generic								Episodic							
	Assessment steps		Mechanics steps		Destination description		Total steps		Assessment steps		Mechanics steps		Destination description		Total steps	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Undergraduate	1.44	0.32	4.19	0.56	0.00	0.00	5.63	0.38	1.62	0.24	5.88	0.74	0.19	0.14	7.69	0.58
Athlete	1.25	0.27	5.00	0.58	0.00	0.00	6.25	0.55	1.25	0.27	5.12	0.63	0.38	0.15	6.75	0.77
Golf team	3.69	0.48	4.50	0.84	0.25	0.11	8.44	0.97	1.37	0.30	3.63	0.75	0.56	0.16	5.56	0.60

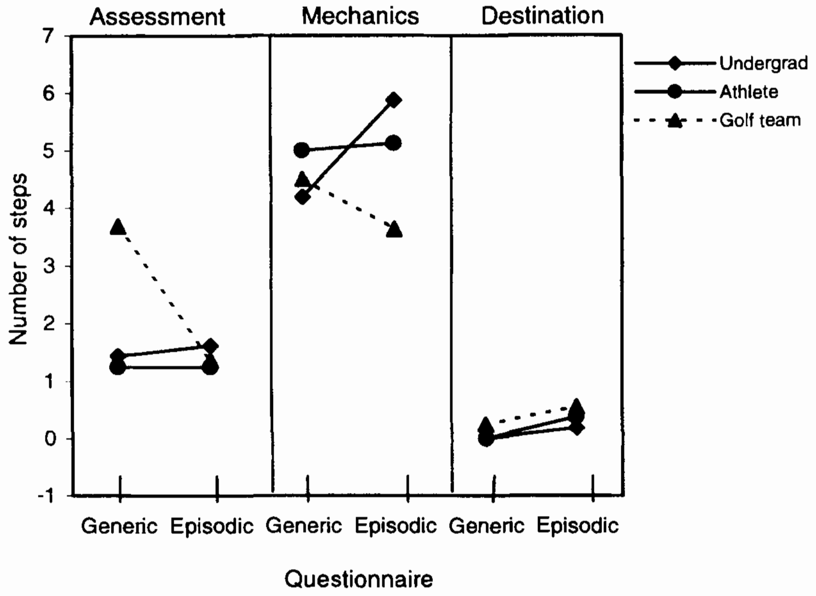

Figure 3. Mean number of steps in each category for the second generic questionnaire and the episodic questionnaire for each group. Undergrad = undergraduate.

formance, which is controlled by automated procedural knowledge. Because proceduralization reduces the need to attend to the processes by which skill execution unfolds, episodic recollection of step-by-step real-time performance is impoverished.

How are the details of these declarative reports related to the accuracy of performance? A significant negative correlation was found between the length of the undergraduates' generic descriptions and their pretest putting accuracy ($r = -.52$, $p < .03$). Because the measure of golf putting accuracy in the present study was an error score (i.e., mean distance from the target), it appears that the more detailed the generic descriptions supplied by the undergraduate novices early in practice, the better they performed. This correlation is consistent with an additional idea derived from current theories of skill and automaticity stipulating that novices' real-time performance is controlled by declaratively accessible knowledge concerning skill execution. Furthermore, this correlation was the only significant individual-differences relationship found between the contents of the declarative protocols and the accuracy of putting within any of the groups. This pattern suggests that a more extensive generic representation aids putting performance in the very earliest stages of skill learning but loses its impact as practice proceeds, once again consistent with expectations generated from theories of skill and automaticity. The disappearance of the correlation between undergraduates' generic knowledge and their performance accuracy appears to have occurred rapidly in the present situation, disappearing by 60–70 putts in the posttest scores. Thus procedural control structures may be established and begin to come to the fore quite quickly, at least in certain task domains (see Brown & Carr, 1989; Klapp et al., 1991; Raichle et al., 1994).

Although experts gave less elaborate episodic recollections of putting mechanics than did their novice counterparts, they gave more extensive recollections of ball destinations. This result suggests that performance outcomes are more salient to expert golfers, paralleling findings in other, more cognitive domains. It has been shown that expert physicists allocate more attentional resources to assessing and monitoring specific goal outcomes during problem solving than do less experienced physicists (Voss & Post, 1988). Of course, in our simple and repetitive situation, outcomes were generally similar to one another in both form and importance. The very low rate of inclusion of outcome information, even by experts, should increase as competitive motivations and consequences of success or failure become greater. We now turn to the second experiment in the present study, which was designed to replicate and extend the findings of Experiment 1.

Experiment 2

In Experiment 2 expert golfers' generic knowledge of golf putting and episodic recollection of specific putts were again compared with the generic knowledge and episodic recollection of novice golfers. Knowledge and recollection were assessed during either a standard golf putting task using a normal putter (i.e., the same task as in Experiment 1) or an altered putting task using a "funny putter" that consisted of a regular putter head attached to an S-shaped curved and arbitrarily weighted putter shaft. The design of the funny putter required experienced golfers to alter their well-practiced putting form in order to compensate for the distorted club, forcing them to allocate attention to the new skill execution processes. If experts' golf putting skill is proceduralized, then the disruption caused by the novel putter should not only lead to a lower level of performance in comparison to regular putter use but should also produce more elaborated episodic memory protocols—possibly similar to those of the novice golfers—as a result of the need to attend to the specific processes of skill execution under the constraints of the new putter. However, novice performers should not be affected by the funny putter in the same way as more experienced golfers. Because novices have not yet adapted to putting under normal putter constraints, performance should not depend as heavily on the type of putter used. Furthermore, according to the theories of skill acquisition we have reviewed, novices' on-line representations of golf putting are explicitly monitored in real time. Therefore, attending to novel putter constraints should not produce different episodic memory protocols in comparison with regular putter use, because in both cases novices attend to their performances in a way that should support explicit episodic memory.[1]

The design of Experiment 2 was similar to that of Experiment 1, with three exceptions. First, in order to ensure that individuals were not adapting to the highly repetitive task of putting from one specific spot on the green, all participants alternately putted from nine different spots, located at varying angles and distances from the target. Second, the experienced golfers in Experiment 2 were university students with 2 or more years of high school varsity golf experience rather than intercollegiate golf team members. Last, in Experiment 2 participants filled out two episodic protocols. As in Experiment 1, the first episodic questionnaire was unexpected. Prior to the last putt taken before the second episodic questionnaire, however, individuals were instructed to monitor their performance carefully for later recall.

Method

Participants

Participants ($N = 72$) were undergraduate students enrolled at Michigan State University and consisted of experienced golfers with 2 or more years of high school varsity golf experience ($n = 36$) and introductory psychology students with no golf experience ($n = 36$). Participants were randomly assigned within skill level to either a regular putter or funny putter condition in a 2 (novice golfer, experienced golfer) × 2 (regular putter, funny putter) experimental design with 18 participants in each group.

Procedure

After giving informed consent and filling out a demographic sheet concerning previous golf experiences, participants were told that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were instructed that the object of the task was to putt a golf ball as accurately as possible from nine locations on a carpeted indoor putting green (3 × 3.7 m) that were either 1.2, 1.4, or 1.5 m away from a target, marked by a square of red tape, on which the ball was supposed to stop. All participants followed the same random alternation of putting from the nine different locations. A standard golf putter and golf ball were supplied for those participants who took part in the regular putter condition, and the funny putter and a standard golf ball were supplied for those participants in the funny putter condition.

All groups participated in identical pretest, practice, and posttest conditions, though the participants were not made aware of the separate conditions. To the participant, the golf putting task appeared to involve four blocks of putts with a short break after each block during which a questionnaire was filled out.

Pretest condition. Participants were set up at the first putting spot. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. Participants were then informed that they would be putting from nine different locations on the green, each with a corresponding number. The experimenter reviewed the numbers associated with each putting location and asked participants to repeat back the numbers corresponding to each putting spot. Participants were informed that the experimenter would call out a number corresponding to a particular spot on the green from which they were to execute their next putt. Participants then took a series of 20 putts. After completing the putts, participants filled out a questionnaire eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Practice condition. Participants were again set up at the first putting spot. Participants took a series of 30 putts. After completing the putts, participants filled out an identical questionnaire to the one that they had previously filled out in the pretest condition eliciting a description of the steps involved in a typical golf putt (Appendix A, first paragraph).

Posttest 1 condition. Participants were set up at the first putting spot. Participants then took a series of 20 putts. Immediately following the first posttest condition, participants filled out a questionnaire designed to access their episodic recollection of the last putt they had just taken (Appendix A, third paragraph).

Posttest 2 condition. Participants were again set up at the first putting spot. Participants then took a series of 10 putts. Immediately prior to the 10th putt in the trial block, the experimenter instructed participants that they should pay close attention to the processes involved in their next putt because after it was complete, they would be asked to fill out another questionnaire, identical to the one they had just filled out, regarding their memories of this next putt. Immediately following the second posttest condition, participants filled out a questionnaire designed to access their episodic recollection of the last putt they had just taken (Appendix A, third paragraph).

Results

Putting Performance

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target at which the ball stopped after each putt. As in Experiment 1, the mean distance from the target of the last 10 putts in the pretest condition was used as a measure of pretest golf putting skill. The mean distance from the target of the middle 10 putts in the practice condition was used as a measure of practice putting skill. The mean distance from the target of the last 10 putts in the first posttest condition was used as a measure of Posttest 1 golf putting skill. The mean distance from the target of the 10 putts in the second posttest condition was used as a measure of Posttest 2 golf putting skill. Means and standard errors for putting performance appear in Figure 4.

As can be seen from Figure 4, the experienced golfers showed superior putting performance in comparison with the novice golfers, regardless of type of putter used. This was true both before and after the practice phase. This pattern was confirmed by a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (pretest, Posttest 1) ANOVA, which revealed significant main effects of experience, $F(1, 68) = 42.73$, $p < .001$, $MSE = 51.55$, and test, $F(1, 68) = 4.04$, $p < .048$, $MSE = 25.25$; no significant main effect of putter, $F(1, 68) = 1.47$, $p < .229$, $MSE = 51.55$; and no interaction of Test × Experience × Putter ($F < 1$).

In order to assess putting performance from the pretest condition to the second posttest condition, a three-way ANOVA similar to the one reported above was computed using the mean distance from the target of the last 10 putts in the pretest condition as a measure of pretest skill and the mean distance from the target of the 10 putts in the second posttest condition as a measure of

---

[1] We thank Claudia Carello for suggesting the funny putter as a diagnostic tool.

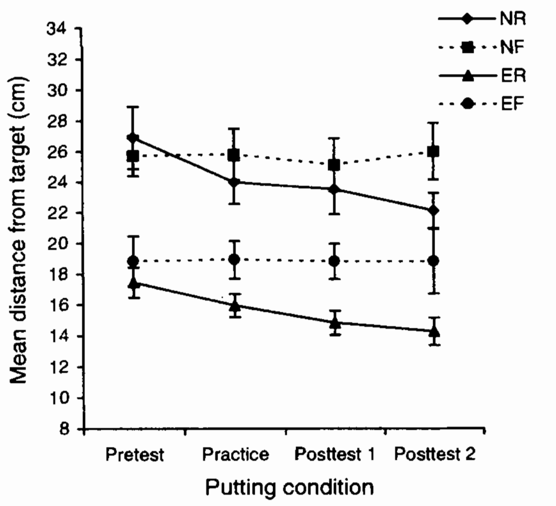

Figure 4. Mean (± *SE*) distance from the target at which the ball stopped after each putt in the pretest, practice condition, Posttest 1, and Posttest 2. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

Posttest 2 golf putting skill. The results of this analysis did not differ from those reported above.

Thus, as can be seen from Figure 4, the experienced golfers, regardless of type of putter used, outperformed the novice golfers at all stages of practice. In addition, experienced golfers using the funny putter were less accurate than the regular putter–experienced golfers—especially during the practice condition and posttests. Independent sample *t* tests within the experienced golfers revealed no significant differences between putter type during the pretest, $t(34) = 0.74$, $p > .47$, but significant differences during the practice condition, $t(34) = 2.08$, $p < .05$, and the first posttest, $t(34) = 2.87$, $p < .007$, and marginally significant differences during the second posttest, $t(34) = 2.0$, $p < .054$. In contrast, the novice golfers did not significantly differ by putter type at any point in the experiment, although novices using the funny putter generally performed at a slightly lower level than their regular putter counterparts. Thus, although the funny putter produced differences in performance within higher levels of experience, it did not significantly affect the less experienced golfers. It should be noted that although experienced golfers using the funny putter performed at a lower level than regular putter experts during the pretest, this difference was not statistically significant. It may be that in the pretest condition, expert golfers—regardless of putter type—were adjusting to the novel experimental demands of having to land the ball on the target rather than in a hole. Thus, regular putter experts were not performing up to their potential in the pretest. The difference between the regular and funny putter experts widened quickly, however, appearing as early as the practice condition. Because experienced golfers often encounter novel putting green environments and must adapt to these situations in order to maintain a low handicap, it is not surprising the regular putter experts were able to rapidly adjust to our indoor green. In fact, several of the experienced golfers mentioned adjusting to the "fast green" or having to "land the ball on the tape" in their episodic protocols, suggesting that these individuals were able to identify and adapt to our somewhat irregular putting environment. In contrast, as can be seen from Figure 4, those experts using the funny putter were unable to adapt to the demands of the new putter within the time frame of the experiment, performing at a similar level of accuracy across experimental conditions.

Generic and Episodic Memory Protocols

As in Experiment 1, questionnaire responses were analyzed quantitatively, in terms of the number of golf putting steps included in each type of protocol, and qualitatively, in terms of the relative frequencies of different categories of steps.

Quantitative analysis. Analysis of number of golf putting steps given by participants was performed in the exact same manner as in Experiment 1. Two experimenters independently coded the data. Interexperimenter reliability was extremely high ($r = .95$).

Table 4 and Figure 5 present the results. A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (first generic protocol, second generic protocol) ANOVA on the two generic protocols revealed a marginally significant main effect of test, $F(1, 68) = 3.03$, $p < .086$, $MSE = 1.67$, and no interaction of Expertise × Putter × Test ($F < 1$). Thus, as in Experiment 1, the second generic protocol was used in a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (second generic protocol, first episodic protocol) ANOVA to compare the lengths of the generic and episodic protocols produced at each level of expertise. This analysis revealed an interaction of Experience × Putter × Questionnaire, $F(1, 68) = 9.63$, $p < .003$, $MSE = 2.77$.

A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the second generic

Table 4
Questionnaire Responses: Number of Steps (Experiment 2)

	Generic 1		Generic 2		Episodic 1		Episodic 2	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
NR	6.39	0.51	6.69	0.59	9.28	0.94	9.78	0.96
NF	6.11	0.65	6.67	0.75	9.11	0.72	9.83	0.81
ER	8.17	0.81	8.79	0.76	7.11	0.57	8.60	0.72
EF	10.22	0.81	10.30	0.85	11.89	0.75	11.78	0.89

Note. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

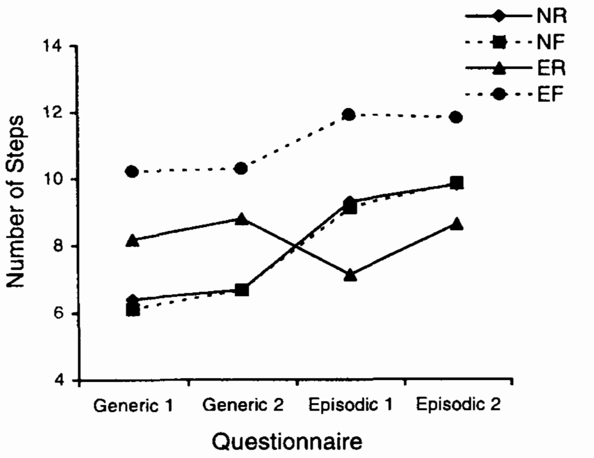

Figure 5. Mean number of steps for the first and second generic questionnaires and the first and second episodic questionnaires for each group. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

protocol produced a main effect of expertise, $F(1, 68) = 14.72$, $p < .001$, $MSE = 10.01$, with the experienced performers giving longer generic protocols than the novices; no main effect of putter, $F(1, 68) = 1.01$, $p > .318$, $MSE = 10.01$; and no Experience × Putter interaction, $F(1, 68) = 1.01$, $p > .318$, $MSE = 10.01$. In contrast, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the first episodic questionnaire produced an Experience × Putter interaction, $F(1, 68) = 10.70$, $p < .002$, $MSE = 10.28$. Independent sample t tests revealed that whereas the two novice groups did not differ in terms of the number of steps given in their episodic protocols, $t(34) = 0.14$, $p > .89$, the funny putter–experienced golfers gave significantly more steps in their episodic protocol than the regular putter–experienced golfers, $t(34) = 5.09$, $p < .001$. In addition, both novice groups and the funny putter–experienced group gave significantly more steps than the regular putter–experienced golfers in the episodic questionnaire ($ps < .05$).

Direct comparisons of the number of generic versus episodic steps within each group showed that, similar to Experiment 1, both the regular and funny putter novices gave significantly more steps in their episodic than their generic protocols, $t(17) = 4.10$, $p < .001$, and $t(17) = 6.27$, $p < .001$, respectively. In addition, the experienced golfers using the funny putter gave significantly more steps in their episodic than in their generic protocols, $t(17) = 2.64$, $p < .017$. In contrast, the experienced golfers using the regular putter gave significantly more steps in their generic than in their episodic protocols, $t(17) = 3.04$, $p < .007$. Furthermore, as can be seen from Figure 5, the experienced golfers using the funny putter gave longer generic and episodic putting descriptions than any other group. Increased attention to the novel constraints of the funny putter most likely prompted these golfers to allocate more attention to skill execution processes, enhancing generic descriptions and leaving explicit episodic memory traces of performance.

If the funny putter–experienced golfers gave more elaborate episodic descriptions as a result of increased attention to the specific processes involved in novel skill execution, then instructing these individuals to pay close attention to a particular instance of a putt, as did the instructions given prior to filling out the second episodic questionnaire, should not significantly change episodic descriptions in comparison to the first unexpected episodic questionnaire. That is, if the constraints of the funny putter serve to increase attention to skill execution, instructing experienced golfers to explicitly monitor performance should not alter attentional allocation and thus should not affect episodic memory protocols. In contrast, if those experts using the regular putter are asked to monitor performance for a later recall test, their episodic descriptions should increase in comparison to their first episodic protocol—especially if the first recollection was truly based on an unmonitored proceduralized instance of performance.

A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (first episodic protocol, second episodic protocol) ANOVA was performed, producing a significant Protocol × Experience × Putter interaction, $F(1, 68) = 4.08$, $p < .047$, $MSE = 1.86$. As can be seen in Figure 5, the novices, regardless of putter type, gave marginally longer putting descriptions in the second episodic questionnaire than in the first, $t(17) = 1.7$, $p < .108$, and $t(17) = 1.83$, $p < .085$, respectively. The experienced golfers using the funny putter did not differ in putting description length from the first to second episodic questionnaire, $t(17) = 0.11$, $p > .92$. In contrast, the experienced golfers using the regular putter gave longer protocols in the second episodic questionnaire in comparison to the first, $t(17) = 2.82$, $p < .012$. Furthermore, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the second episodic questionnaire produced a marginally significant Experience × Putter interaction, $F(1, 68) = 3.35$, $p < .071$, $MSE = 12.99$. Thus, instructing participants to monitor skill execution did not affect the funny putter experts' episodic recollections and only marginally influenced the novice golfers' episodic descriptions. However, although instructing regular putter experts to monitor performance did increase their episodic recollections, they still did not reach the level of either the novice group or the funny putter–experienced golfers.

Qualitative analysis: Types of steps. The qualitative analysis was performed in the same manner as in Experiment 1. Steps were divided into three categories (assessment, mechanics, and ball destinations), and a 2 (experienced golfers, novice golfers) × 2 (funny putter, regular putter) × 2 (second generic protocol, first episodic protocol) ANOVA was conducted on the number of steps given in each of these three categories (see Table 5).

The analysis of assessment steps produced a significant interaction of expertise and type of protocol, $F(1, 68) = 14.53$, $p < .001$, $MSE = 1.2$, which is displayed in the left panel of Figure 6, along with a nonsignificant interaction of Expertise × Protocol × Putter ($F < 1$). A one-way ANOVA on the generic protocol with putter collapsed within skill level produced a main effect of experience, $F(1, 70) = 23.47$, $p < .001$, $MSE = 2.98$. Assessment steps appeared more often in the generic descriptions of experienced golfers, regardless of putter type, than anywhere else. In terms of the episodic protocol, a one-way ANOVA with putter collapsed within skill level produced a marginally significant main effect of experience, $F(1, 70) = 3.58$, $p < .063$, $MSE = 1.71$, with experienced golfers continuing to give more assessment steps in their episodic recollections than the novices. Paired sample t tests further revealed that the regular putter–experienced golfers gave

Table 5

Assessment, Mechanic, and Destination Descriptions by Questionnaire Type—Experiment 2

	Assessment		Mechanics		Destination		Total	
Group	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
				Generic				
NR	1.70	0.32	5.00	0.54	0.00	0.00	6.69	0.59
NF	1.56	0.29	5.11	0.74	0.00	0.00	6.67	0.75
ER	3.54	0.41	5.26	0.63	0.00	0.00	8.79	0.76
EF	3.65	0.57	6.66	0.82	0.00	0.00	10.30	0.85
				Episodic 1				
NR	1.72	0.37	7.33	0.71	0.22	0.10	9.28	0.94
NF	1.94	0.26	7.00	0.80	0.17	0.09	9.11	0.72
ER	2.06	0.19	4.94	0.58	0.11	0.08	7.11	0.57
EF	2.78	0.37	8.72	0.66	0.39	0.16	11.89	0.75
				Episodic 2				
NR	1.89	0.25	7.61	0.88	0.28	0.14	9.78	0.96
NF	1.67	0.26	7.89	0.90	0.28	0.11	9.83	0.81
ER	2.35	0.23	5.96	0.61	0.28	0.11	8.60	0.72
EF	2.56	0.35	9.06	0.70	0.17	0.09	11.78	0.89

Note. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

significantly more assessment steps in their generic descriptions than in their episodic recollections, $t(17) = 4.03, p < .001$, and the funny putter–experienced golfers gave somewhat more assessment steps in their generic descriptions, though the difference was not significant, $t(17) = 1.72, p < .104$. In contrast, the regular putter novices did not differ in terms of the number of assessment steps given in the generic and episodic protocols, $t(17) = 0.00$, while the funny putter–novice group gave more assessment steps in their episodic recollections than in their generic protocols, $t(17) = 2.12$, $p < .049$. Thus, similar to Experiment 1, assessment steps decreased in number from the generic to episodic protocol for the two experienced groups, regardless of type of putter used, whereas the opposite pattern occurred in the two novice groups.

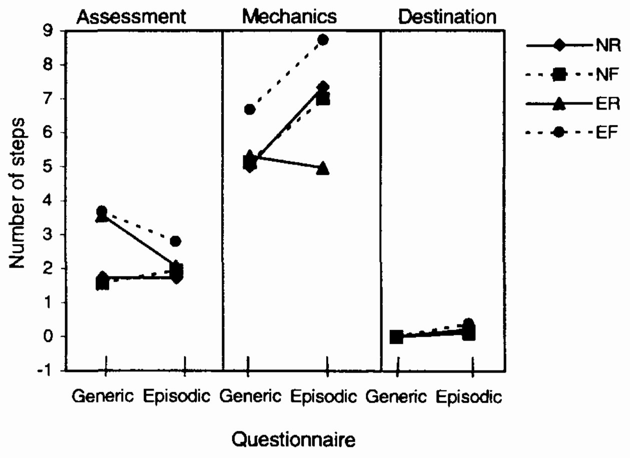

Figure 6. Mean number of steps in each category for the second generic questionnaire and the first episodic questionnaire for each group. NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

As an adjunct to the analysis of assessment, those steps that involved mental imagery (i.e., imagining some aspect of how a putt ought to look or feel before executing the action) were counted. In the regular putter–novice group, 2.7% of generic steps and 1.5% of episodic steps referred to imagery. In the funny putter–novice group, 0.6% of generic steps and 1.9% of episodic steps referred to imagery. In the regular putter–expert group, 2.2% of generic steps and 5.0% of episodic steps involved imagery. Finally, in the funny putter–expert group, 1.2% of generic steps and 1.4% of episodic steps referred to imagery. Thus, as in Experiment 1, golf putting steps involving imagery were predominantly reported by experienced golfers using the regular putter.

Turning to mechanics, this analysis produced an interaction of Experience × Protocol × Putter, $F(1, 68) = 5.26, p < .025$, $MSE = 3.43$, as can be seen in the middle panel of Figure 6. A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA on the generic protocol produced a nonsignificant main effect of experience, $F(1, 68) = 1.77, p < .188$, $MSE = 8.55$ (though experienced golfers did give more mechanics steps in their generic protocols than novices in terms of absolute number); no main effect of putter, $F(1, 68) = 1.18, p > .280$, $MSE = 8.55$; and no interaction of Experience × Putter ($F < 1$). In the episodic protocol, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA produced an Experience × Putter interaction, $F(1, 68) = 8.82, p < .004$, $MSE = 8.63$. As can be seen from the middle panel of Figure 6, the experienced golfers using the funny putter gave more mechanics steps than any other group, while the regular putter–experienced golfers gave fewer mechanics steps than the other three groups. The two novice groups did not differ.

In addition, both regular and funny putter novices gave significantly more mechanics steps in their episodic recollections than in their generic protocols, $t(17) = 2.20$, $p < .001$, and $t(17) = 1.84$, $p < .001$, respectively. Similarly, the experienced golfers using the funny putter gave significantly more mechanics steps in their episodic recollections as compared with their generic protocols, $t(17) = 4.27$, $p < .001$. In contrast, the regular putter experienced golfers gave fewer mechanics steps in their episodic recollections than in their generic descriptions, although this difference was not significant, $t(17) = 0.556$, $p < .585$.

The analysis of ball destinations produced a main effect of protocol type, $F(1, 68) = 15.43$, $p < .001$, $MSE = 0.12$; no main effect of experience or putter ($Fs < 1$); and no interaction of Protocol × Experience × Putter, $F(1, 68) = 2.17$, $p > .145$, $MSE = 0.12$. Thus, as shown in the right panel of Figure 6, regardless of putter type or expertise, ball destinations were more likely to appear in episodic recollections than generic protocols.

As in Experiment 1, a second qualitative analysis looked for elaborations—steps present in both protocols that referred to the same action or biomechanism but provided more detail in one type of protocol than in the other. Because elaborations were more likely to occur in episodic descriptions relative to generic descriptions than vice versa, greater detail in the episodic description was scored as a positive elaboration whereas greater detail in the generic description was scored as a negative elaboration. In the regular putter–novice group, 11.1% of the steps in the episodic description were elaborations of steps in the generic descriptions. In the funny putter–novice group, 7.5% of the episodic steps were elaborations of generic steps. In the regular putter–experienced group, −0.3% of episodic descriptions were elaborations of generic steps. Finally, in the funny putter–experienced group, 3.5% of episodic recollections were elaborations of generic steps. A 2 (experienced golfer, novice golfer) × 2 (regular putter, funny putter) general factorial ANOVA on these data produced a significant main effect of experience, $F(1, 68) = 8.33$, $p < .005$, $MSE = 1.28$; a nonsignificant effect of putter ($F < 1$); and no Experience × Putter interaction, $F(1, 68) = 1.97$, $p > .165$, $MSE = 1.28$. Regardless of putter used, the novice golfers gave more elaborations in their episodic protocols than the more experienced golfers. Furthermore, all groups gave more elaborations in their episodic descriptions than in their generic protocols, with the exception of the regular putter–experienced golfers, who gave fewer elaborations in their episodic recollections.

In order to assess qualitative differences between the first and second episodic protocols, a 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) × 2 (first episodic protocol, second episodic protocol) ANOVA was also performed on each of the three categories of steps (assessment, mechanics, destination; see Table 5). The analysis of assessment steps produced a main effect of experience, $F(1, 68) = 6.00$, $p < .017$, $MSE = 2.34$; no main effect of protocol or putter; and no Experience × Putter × Protocol interaction ($Fs < 1$). Thus, as can be seen from the left panel of Figure 7, the experienced golfers gave more assessment steps in both episodic questionnaires than did either group of novices, who did not differ.

As an addition to assessment steps, the percentage of second episodic protocol steps involving references to mental imagery was also assessed. In the regular putter–novice group, 1.3% of second episodic steps referred to imagery. In the funny putter–novice group, 2.2% of second episodic steps referred to imagery. In the regular putter–expert group, 4.8% of second episodic steps involved imagery. Finally, in the funny putter–expert group, 1.3% of second episodic steps referred to imagery. Thus, as in the first episodic questionnaire, golf putting steps involving mental imagery were more likely to be found in the regular putter–experienced golfers' protocols than anywhere else.

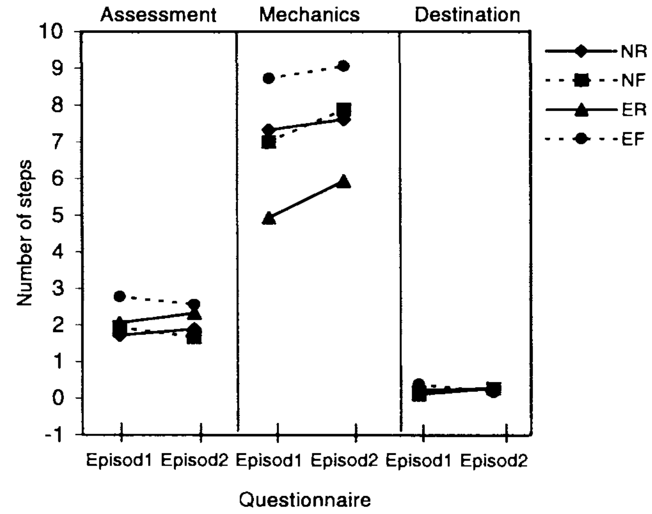

Figure 7. Mean number of steps in each category for the first and second episodic questionnaires for each group. Episod = Episodic; NR = novice golfer–regular putter; NF = novice golfer–funny putter; ER = experienced golfer–regular putter; EF = experienced golfer–funny putter.

The analysis of mechanics produced a main effect of protocol, $F(1, 68) = 8.13$, $p < .006$, $MSE = 1.73$, and an Experience × Putter interaction, $F(1, 68) = 6.07$, $p < .016$, $MSE = 17.87$. A 2 (experienced golfer, novice golfer) × 2 (funny putter, regular putter) general factorial ANOVA was then performed on the combined average of mechanics steps involved in the first episodic questionnaire and second episodic questionnaire, revealing an Experience × Putter interaction, $F(1, 68) = 6.07$, $p < .016$, $MSE = 8.93$. As can be seen in the middle panel of Figure 7, mechanics steps increased from the first to second episodic protocol across all groups. Furthermore, the funny putter–experienced golfers gave more mechanics steps than any other group in either episodic questionnaire, whereas the regular putter–experienced golfers gave fewer mechanics steps than any other group. The two novice groups did not differ. Thus, although the regular putter–experienced golfers knew that they would be asked to give an episodic recollection of the last putt they took in the second episodic questionnaire, these golfers still gave fewer mechanics steps than both groups of inexperienced golfers and the experienced golfers using the funny putter.

Analysis of ball destinations was not interpretable because of inhomogeneity of variance across groups for destination steps reported in the episodic questionnaires. However, as can be seen from the right panel of Figure 7, all groups gave similar absolute numbers of ball destination steps in both the first ($M = 0.22$, $SD = 0.48$) and the second ($M = 0.25$, $SD = 0.47$) episodic questionnaires.

Discussion

Experiment 2 yielded results similar to those of Experiment 1. Experts using the regular putter gave more elaborate generic representations of putting than novices using either type of putter, in parallel with diminished episodic accounts of particular instances of performance. Those experts who were asked to use the funny putter also gave more detailed generic representations than did novices. However, in contrast to experts using the regular putter in both experiments, experts using the funny putter did not show diminished episodic memories for specific performances as would be expected if on-line performance was executed automatically and without leaving an explicit memory trace. In fact, funny putter experts gave more elaborate episodic descriptions of particular instances of skill execution than did the experts using the regular putter and both novice groups. The results of Experiment 2 suggest that real-time putting performance for experienced golfers is supported by proceduralized knowledge that may be disrupted through the addition of novel task constraints. When this disruption occurs and experts are forced to attend to step-by-step performance in the same way as novices, their expertise allows them to remember more of what they have attended to than do less skilled performers. This outcome resembles findings of superior episodic memory in chess and computer programming experts for the stimuli to which they have applied their knowledge (Chase & Simon, 1973; Soloway & Ehrlich, 1984). Furthermore, regardless of the type of putter used, novice golfers in the present study produced similar putting performances and generic and episodic putting descriptions, thus suggesting that in contrast to experienced golfers, novel skill performance in novice golfers is not based on a proceduralized, practice-specific skill representation.

The results of Experiments 1 and 2 speak to the nature of skill representations at various levels of expertise. With respect to the sensorimotor skill of golf putting, it appears that highly practiced, well-learned task components are encoded in a procedural form that supports effective real-time performance without requiring step-by-step attentional control. Reduced attention leads to a reduction in declaratively accessible episodic memory for details of the performance. However, if task constraints (e.g., a funny putter) are imposed that force experienced golfers to alter execution processes in order to adjust to the novel environment, the proceduralized skill knowledge that once drove normal execution is disrupted. The consequence is a reduction in putting accuracy, an extended period of adaptation in which learning appears to proceed rather slowly, and a more detailed episodic memory trace for specific instances of skill execution. The notion that well-learned sensorimotor skill performance is governed by a proceduralized representation carries implications for how this type of skill will behave in pressure or attention-demanding situations. Specifically, the two main theories that have been proposed to account for choking make different predictions concerning the types of skills that will be susceptible to performance decrements under pressure. Next we review these theories and how the cognitive mechanisms hypothesized by each theory to account for choking under pressure may be related to the type of knowledge representation governing task performance.

Experiments 3 and 4: Choking Under Pressure

As outlined in the introduction, two types of theories have attempted to explain choking under pressure. Distraction theory proposes that pressure influences task performance by creating a distracting environment. Distraction-based accounts of suboptimal performance propose that performance pressure shifts attentional focus to task-irrelevant cues—such as worries about the situation and its consequences. In essence, this shift of focus changes what was single-task performance into a dual-task situation in which controlling the task at hand and worrying about the situation compete for attention. The most notable arguments for the distraction hypothesis come from research involving academic test anxiety (Eysenck, 1979; Kahneman, 1973; Wine, 1971). Individuals who become highly anxious during test situations, and consequently perform at a suboptimal level, are thought to divide their attention between task-relevant and task-irrelevant thoughts more so than those who do not become overly anxious in high-pressure situations (Wine, 1971).

The distraction explanation for performance decrements under pressure is consistent with the idea that complex performances are attention or capacity demanding and that removing attention will disrupt performance (Humphreys & Revelle, 1984; Norman & Bobrow, 1975; Proctor & Dutta, 1995). However, as demonstrated in Experiments 1 and 2, there are skills that become automated or proceduralized with extended practice and thus may not require constant on-line attentional control during execution (Anderson, 1987, 1993; Fitts & Posner, 1967; Kimble & Perlmuter, 1970; Proctor & Dutta, 1995; Squire & Knowlton, 1994). Such skills should be able to withstand the attentional demands of a dual-task environment in that explicit attention to step-by-step skill procedures is not mandatory for successful performance. The well-learned sensorimotor task of golf putting may be one such task. However, skills that rely on declaratively accessible knowledge even in their practiced state may behave quite differently under pressure. A potential example is Zbrodoff and Logan's (1986) alphabet arithmetic task.

Alphabet arithmetic is a laboratory task analogous to mental arithmetic in which skilled performance is thought to be supported by the retrieval of stored instances of particular equations to which the performer has been exposed. Answers to an alphabet arithmetic problem such as "$A + 2 = ?$," whereby individuals must count two units down the alphabet to obtain the answer "C," may be achieved in two ways: either by using a rule-based system or algorithm to solve the equation or by the stimulus-driven retrieval of past instances of the problem from memory. Logan (1988) assumed that solutions are derived by a race between these two processes. As exposure to examples of problems increases, instances stored in memory increase as well. The larger the base of instances stored in memory, the higher the probability that memory retrieval will provide an answer to the problem before the rule-based algorithm reaches a solution. In either case, the answer enters working memory and hence is declaratively accessible—what differs is how it gets there (Klapp et al., 1991). Logan's model is supported by changes in speed and accuracy of performance with practice on the alphabet arithmetic task and other rapidly performed tasks involving judgment and choice, such as lexical decision and semantic categorization.

Answers to Logan's (1988) alphabet arithmetic task are thought to be declaratively accessible at all stages of skill learning. If choking is due to distraction of attention, one might imagine that choking would be a more imminent danger in tasks based on declarative knowledge that often enters working memory during the course of performance, because distraction of attention is a primary antecedent of corruption of information and forgetting in working memory (Daneman & Carpenter, 1980; Muter, 1980; Peterson & Peterson, 1959; Posner & Rossman, 1965). Furthermore, if the distraction hypothesis is valid, then it is possible that training in a dual-task environment would enable performers to adapt to distraction and the concurrent allocation of attention to something other than the primary task, alleviating the negative impact of pressure (Hirst, Neisser, & Spelke, 1978; Spelke, Hirst, & Neisser, 1976).

It has also been proposed that pressure situations raise anxiety and self-consciousness about performing correctly and successfully. The resulting focus on the self prompts individuals to turn their attention inward on the specific processes of performance in an attempt to exert more explicit monitoring and control than would be applied when a high-achievement outcome is less desired and its consequences are less important (Baumeister, 1984; Lewis & Linder, 1997). Note that the essence of this proposal is exactly the opposite of the distraction hypothesis. The main idea behind the self-focus, or what we would like to term the explicit monitoring hypothesis, is that although close attention and control may benefit novice performers in the initial stages of task learning, it will become counterproductive as practice builds a more and more automated performance repertoire. This is due to the fact that explicit monitoring of step-by-step skill processes and procedures is thought to disrupt well-learned or proceduralized skill execution processes (Kimble & Perlmuter, 1970; Langer & Imber, 1979; Lewis & Linder, 1997). Masters (1992) proposed that performance disruption occurs when an integrated or compiled real-time control structure that can run as an uninterrupted unit is broken back down into a sequence of smaller, separate, independent units—similar to how the performance was organized early in learning. Once broken down, each unit must be activated and run separately, which slows performance and, at each transition between units, creates an opportunity for error that was not present in the integrated control structure.

In addition to the differences in the types of knowledge that may govern task performance, variations in complexity of skills may also mediate the pressure–performance relationship. That is, it may be that complex skills, involving the integration and sequencing of multiple steps or parts, are more prone to breakdowns and performance deficits than less complex one-step skills. Certainly the skill of golf putting involves such complexity.

According to the explicit monitoring theory, then, the complex, proceduralized sensorimotor skill of golf putting analyzed in Experiments 1 and 2 should be extremely susceptible to performance decrements under pressure, as it is unaccustomed to being explicitly attended in real time. However, alphabet arithmetic, in which answers to particular problems are thought to be declaratively accessible at all stages of skill learning, should not be negatively affected by pressure-induced attention to performance processes. Furthermore, if the explicit monitoring hypothesis is valid, then training in an environment that heightens self-consciousness and achievement anxiety is likely to alleviate the negative impact of pressure, by adapting performers to conditions that entice them to pay too much attention to step-by-step execution.

Experiment 3

As a first step toward determining whether type of task knowledge and/or complexity might influence susceptibility to choking, we conducted Experiment 3. In this experiment participants learned either the sensorimotor task of golf putting or Zbrodoff and Logan's (1986) more declaratively based alphabet arithmetic task under single-task, dual-task, or self-consciousness-raising training conditions. Following training, participants were exposed to single-task low-pressure and high-pressure posttest situations in their task.

Testing the hypotheses concerning the distraction and explicit monitoring theories of choking under pressure requires control over the training environment. To ensure that our manipulation was the major source of each participant's golf putting or alphabet arithmetic experience, we recruited novice golfers and individuals with no exposure to alphabet arithmetic and taught them these tasks in the laboratory. However, despite the predictions we have made with respect to novice performance, choking as a concept is primarily aimed at individuals who can be expected to perform at a relatively accomplished level. Therefore, in order to examine performance at the later stages of skill acquisition, we trained participants rather heavily. Participants performed more than 280 golf putts or alphabet arithmetic trials in our laboratory prior to being exposed to a high-pressure situation. This number of task repetitions was chosen because pilot testing revealed a leveling off in performance with this amount of practice, suggesting that performance on the practiced putts or alphabet arithmetic problems was reaching asymptote.

Method

Participants

Undergraduate students ($N = 108$) with little or no golf experience who were enrolled in an introductory psychology class at Michigan State University served as participants. Participants were randomly assigned to either a single-task, self-consciousness, or dual-task distraction training group in either the golf putting or alphabet arithmetic task. Eighteen participants took part in each training group.

Procedure: Golf Putting Task

After giving informed consent and filling out a demographic sheet concerning previous golf experiences, participants were told that the purpose of the study was to examine the accuracy of golf putting over several trials of practice. Participants were instructed that the object of the task was to putt a golf ball as accurately as possible from nine locations on a carpeted indoor putting green (3 × 3.7 m) that were either 1.2, 1.4, or 1.5 m away from a target, marked by a square of red tape, on which the ball was supposed to stop. All participants followed the same random alternation of putting from the nine different locations. A standard golf putter and golf ball were supplied. Participants took part in a 270-putt training condition followed by an 18-putt low-pressure posttest and an 18-putt high-pressure posttest described below.

Single-task group. Participants were set up at the first putting spot. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. Participants were then informed that they would be putting from nine different locations on the green, each with a corresponding number. The experimenter reviewed the numbers associated with each putting location and asked participants to repeat back the numbers corresponding to each putting spot. Participants were informed that the experimenter would call out a number corresponding to a particular spot on the green from which they were to execute their next putt. Participants then completed a total of 270 putts consisting of three training blocks of 90 putts each, with a short break after each set of putts. On completion of the training condition, participants were given a short break during which the experimenter computed the mean distance from the target of their last 18 putts.

Participants then completed an 18-putt single-task low-pressure posttest, though they were not made aware of the test situation. To the participant, the low-pressure posttest appeared to be just another series of putts. Participants were then informed of their mean putting performance for the last 18 putts in the training condition and given a scenario designed to create a high-pressure situation. Specifically, participants were told that if they could improve their accuracy by 20% in the next set of putts, they would receive $5. However, participants were also informed that this monetary award was a "team effort." Participants were told that they had been randomly paired with another participant, and in order to receive their $5, not only did they themselves have to improve by 20% but the participant that they had been paired with had to improve by 20% as well. Next, participants were informed that the individual they had been paired with had already completed the experiment and had improved by 20%. Therefore, if the present participant improved by 20%, both participants would receive $5. However, if the present participant did not improve by the required amount, neither participant would receive the money. Participants then took another 18 putts constituting the high-pressure posttest. Following these putts, the experimenter computed the participants' putting average and informed them of their performance. Finally, participants were fully debriefed and given the monetary award regardless of their performance.

Distraction group. Participants were set up at the first putting spot. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. The experimenter informed participants that they would be putting from nine locations on the green. The experimenter then directed participants' attention to a tiny light that had been set up next to each putting spot. Participants were informed that the lights were hooked up to a switchboard controlled by the experimenter. Participants were told that before every putt, a light would illuminate beside the location from which they were to take their next putt. The experimenter then explained to participants that while they were putting they would be listening to a recorded list of spoken words being played from a tape recorder. Participants were told to monitor the words carefully, and, each time they heard the word *cognition,* to repeat it back to the experimenter. Words were played at the rate of one every 2 s. The target word occurred randomly once every four words. Participants then completed a total of 270 putts consisting of three training blocks of 90 putts each, with a short break after each set of putts during which the tape recorder was turned off. When participants completed the training condition, the tape recorder was turned off and participants were given a short break during which the experimenter computed the mean distance from the target of their last 18 putts. Participants then took part in an 18-putt low-pressure posttest and an 18-putt high-pressure posttest identical to that of the single-task group.

Self-consciousness group. Participants were set up at the first putting location. They were asked whether they preferred to putt right-handed or left-handed and were given the appropriate putter. The experimenter then explained that participants would be putting from nine different locations on the green, each with a corresponding number. Once the participants understood the number–putting spot relationships, the experimenter informed participants that they would be filmed by a video camera while putting. The video camera was set up on a tripod that stood on a table directly in front of participants, approximately 1.8 m away. Participants were told that they would be videotaped so that a number of golf teachers and coaches at Michigan State University could review the tapes in order to gain a better understanding of how individuals learn a golf putting skill. The experimenter adjusted the camera and turned it on. Participants then completed a total of 270 putts consisting of three training blocks of 90 putts each, with a short break after each set of putts during which the video camera was turned off. When participants completed the training condition, the video camera was turned off and faced away. Participants were then given a short break in which the experimenter computed the mean distance from the target of their last 18 putts. The participants then took part in an 18-putt low-pressure posttest and an 18-putt high-pressure posttest identical to that of the single-task and distraction groups.

Procedure: Alphabet Arithmetic Task

After giving informed consent and filling out a demographic sheet concerning previous golf and alphabet arithmetic experiences, participants were told that the purpose of the study was to examine how individuals learned the alphabet arithmetic task. Participants were set up in front of a monitor controlled by a standard laboratory computer. Participants were informed that they would be solving alphabet arithmetic equations such as "A + 2 = C" by counting two units down the alphabet to C. Next, the experimenter verbally presented three alphabet arithmetic equations to participants and instructed them to solve the equations out loud in order to ensure proper understanding. Participants were then shown a small keyboard containing two buttons labeled *True* and *False* and told to press the appropriate button when they derived the answer to an equation presented on the screen. Participants were instructed to try to judge the validity of the equations as quickly as possible without sacrificing accuracy.

The stimuli were capital letters, digits, the plus symbol (+), and the equal sign (=). All participants were presented with the same random order of nine different equations, consisting of the letters (A, G, and S) and the digits (2, 3, and 4) that were equally randomly repeated. Each trial began with a fixation point exposed for 500 ms in the center of the screen. The fixation point was immediately replaced by an equation, which remained on the screen until the participant pressed the *True* or *False* button on the keyboard. When the participant responded, the equation was extinguished, and the screen remained blank for a 1.5-s intertrial interval. All participants took part in a 270-equation training condition in which each equation was randomly repeated 30 times, followed by an 18-equation low-pressure posttest and an 18-equation high-pressure posttest, to be described below, in which each equation appeared twice in a random order.

Single-task group. Participants were set up in front of the monitor and given the alphabet arithmetic instructions. Participants then completed a total of 270 equations consisting of three training blocks of 90 equations each, with a short break after each set.

Following completion of the training condition, participants took part in an 18-equation single-task low-pressure posttest, similar to the low-pressure posttest in the golf putting task. Participants then took a short break and were given a scenario designed to create a high-pressure situation. Participants were told that the computer used a formula that equally took into account reaction time and accuracy in computing an "alphabet arithmetic performance score." The experimenter then described the same high-pressure scenario used in the golf putting task. Participants next completed the 18 equations constituting the high-pressure posttest, were fully debriefed, and were given the monetary award regardless of their performance.

Distraction group. Participants were set up in front of the monitor and given the alphabet arithmetic instructions. Participants were also told that while they were performing the arithmetic task they would be listening to a series of words being played through a headset. Participants were instructed to monitor the words carefully and, each time they heard the word *cognition,* to press a foot pedal that was located near their feet. The experimenter then instructed participants to put on the headphones, move the foot pedal to a comfortable location, and practice pressing it a few times. Participants then completed a total of 270 equations consisting of three training blocks of 90 equations each, with a short break after each set during which the headset was taken off.

Following completion of the training condition, participants were instructed to remove the headset and move the foot pedal away from their feet. Participants then took part in an 18-equation low-pressure posttest and 18-equation high-pressure posttest identical to that of the single-task group.

Self-consciousness group. Participants were set up in front of the monitor and given the alphabet arithmetic instructions. Participants were also told that they would be filmed by a video camera while solving the equations. The video camera was set up on a tripod directly to the left of participants, approximately 0.9 m away. Participants were told that they would be videotaped so that a number of math teachers at Michigan State University could review the tapes in order to determine how quickly and accurately individuals learn the arithmetic skill. The experimenter adjusted the camera and turned it on. Participants then completed a total of 270 equations consisting of three training blocks of 90 equations each, with a short break after each set during which the video camera was turned off.

Following completion of the training condition, the video camera was turned off and faced away from the participants. Participants then took part in an 18-equation low-pressure posttest and 18-equation high-pressure posttest identical to that of the single-task and distraction groups.

Results

Putting Performance

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target that the ball stopped after each putt. All three groups improved significantly with practice as demonstrated by a 3 (single-task, distraction, self-consciousness) × 2 (mean distance from target of first 18 putts in training condition, mean distance from target of last 18 putts in training condition) ANOVA revealing a main effect of practice, $F(1, 51) = 85.03, p < .001, MSE = 27.57$; a nonsignificant main effect of training group, $F(2, 51) = 0.658, p > .522, MSE = 90.65$; and no interaction, $F(2, 51) = 0.214, p > .808, MSE = 27.57$. As can be seen in Figure 8, although there was not a significant effect of training group, the distraction group's performance was slightly, but not significantly, degraded in comparison to the single-task and the self-consciousness groups both early and late in the training condition. These results coincide with research in the skill performance literature demonstrating that dual-task performance may lead to a decrement in the performance of a primary task that has not become fully automatized (Proctor & Dutta, 1995).

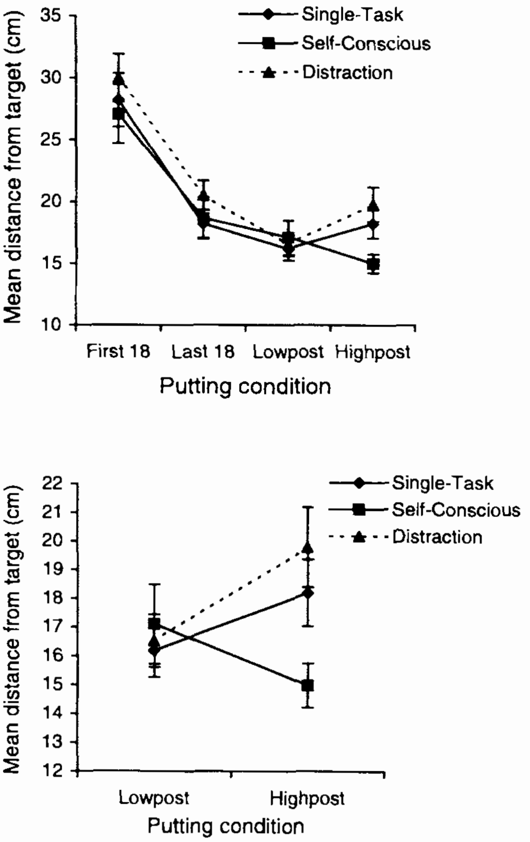

Figure 8. Mean (± *SE*) distance from the target at which the ball stopped after each putt in the training and posttest conditions for each group in the golf putting task (top) and posttest performance only (bottom). Lowpost = low-pressure posttest; Highpost = high-pressure posttest.

In the low-pressure posttest (measured by the mean distance from the target of the 18 putts in the low-pressure test), putting accuracy was nearly identical across groups ($F < 1$). However, this homogeneity disappeared in the high-pressure posttest (measured by the mean distance from the target of the 18 putts in the high-pressure test) as confirmed by a one-way ANOVA revealing a significant difference between the three groups, $F(2, 51) = 4.57, p < .015, MSE = 23.57$. The above results are further supported by a 3 (single-task, distraction, self-consciousness) × 2 (low-pressure posttest, high-pressure posttest) ANOVA revealing a significant interaction of Training Group × Posttest, $F(2, 51) = 7.37, p < .002, MSE = 9.71$. Direct comparisons of putting performance within each group showed that both the single-task group and the distraction group significantly declined in putting accuracy from the low-pressure posttest to the high-pressure posttest, $t(17) = -2.21, p < .04$, and $t(17) = -3.24, p < .005$, respectively. In contrast, the self-consciousness group improved in putting accuracy from the low-pressure posttest to the high-pressure posttest, although this improvement was only marginally significant, $t(17) = 1.81, p < .09$. Thus, as can be seen in Figure 8, whereas both the single-task and distraction groups were adversely affected by the high-pressure situation, the self-consciousness group actually improved.

Alphabet Arithmetic Performance

Accuracy (shown in Table 6) was relatively high and did not differ significantly between groups both early (as measured by the mean number of correct judgments of the first 18 equations in the training condition), $F(2, 51) = 0.178, p > .838, MSE = 5.94$, and late (as measured by the mean number of correct judgments of the last 18 equations in the training condition), $F(2, 51) = 0.735, p > .485, MSE = 2.80$, in the training phase. Early in training, the single-task, distraction, and self-consciousness groups' accuracy was 90%, 89%, and 87% correct, respectively, and late in training the single-task and distraction groups' accuracy was 93% correct and the self-consciousness group's was 96% correct. Similarly, there were no significant differences in accuracy between groups during either the low-pressure posttest (as measured by the mean

Table 6
Mean Accuracy (Percentage Correct) in Training and Posttest Conditions of Alphabet Arithmetic Task

Group	Training 1 (first 18 equations)		Training 2 (last 18 equations)		Low-pressure posttest		High-pressure posttest	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Single task	89.51	2.00	93.21	2.96	96.30	1.56	95.06	1.61
Self-consciousness	87.04	4.42	96.30	1.19	96.60	1.20	94.75	1.38
Distraction	89.20	2.64	92.90	2.05	94.14	1.52	94.75	1.05

number of correct judgments of the 18 equations in the low-pressure test), $F(2, 51) = 0.879$, $p > .421$, $MSE = 1.2$, or the high-pressure posttest (as measured by the mean number of correct judgments of the 18 equations in the high-pressure test), $F(2, 51) = 0.017$, $p > .983$, $MSE = 1.09$. During the low-pressure posttest, accuracy for the single-task, distraction, and self-consciousness groups was 96%, 94%, and 97% correct, respectively, and accuracy during the high-pressure posttest was 95% correct for all three groups.

Reaction times were computed for only those equations that were answered correctly. Mean reaction times and standard errors both early and late in the training condition, as well as for the low- and high-pressure posttests, are illustrated in Figure 9. All three groups significantly decreased their reaction times across the training condition as shown by a 3 (single-task, distraction, self-consciousness) × 2 (mean reaction time of first 18 equations in the training condition, mean reaction time of last 18 equations in the training condition) ANOVA revealing main effects of practice, $F(1, 51) = 171.63$, $p < .001$, $MSE = 6.2 \times 10^5$, and training group, $F(2, 51) = 5.91$, $p < .005$, $MSE = 9.6 \times 10^5$, and a marginally significant interaction, $F(2, 51) = 2.88$, $p < .066$, $MSE = 6.2 \times 10^5$ (see Figure 9). Tukey's honestly significant difference (HSD) tests on the main effect of training group further revealed that the distraction group performed significantly worse than both the single-task and self-consciousness groups (who did not differ) early in the training condition ($p < .035$ and $p < .017$, respectively) and significantly worse than the self-consciousness group ($p < .05$) and nonsignificantly worse than the single-task group ($p < .227$) late in the training condition. This pattern did not change in either the low-pressure or high-pressure posttests as shown by a 3 (single-task, distraction, self-consciousness) × 2 (low-pressure posttest, high-pressure posttest) ANOVA that revealed main effects of both training group, $F(2, 51) = 4.41$, $p < .017$, $MSE = 2.4 \times 10^5$, and posttest, $F(1, 51) = 43.42$, $p < .001$, $MSE = 1.3 \times 10^4$, with no interaction, $F(2, 51) = 2.27$, $p > .114$, $MSE = 1.3 \times 10^4$. Tukey's HSD tests on the main effect of training group further revealed that in the low-pressure posttest, the distraction group produced significantly slower reaction times than the self-consciousness group ($p < .04$) and nonsignificantly slower reaction times than the single-task group ($p < .293$). In the high-pressure posttest, Tukey's HSD tests revealed that the distraction group produced significantly slower reaction times than both the self-consciousness ($p < .008$) and single-task ($p < .04$) groups. Thus, similar to golf putting accuracy, all three groups in the alphabet arithmetic task significantly improved their reaction times with practice. However, in contrast to the golf putting task, the dual-task alphabet arithmetic group performed substantially worse than the single-task and self-consciousness groups both early and late in the training condition, as well as in the posttest situations. The main effect of low-pressure versus high-pressure posttests observed in alphabet arithmetic also contrasts with putting. All three training groups improved somewhat in the high-pressure posttest, showing no signs of choking under pressure in the alphabet arithmetic task.

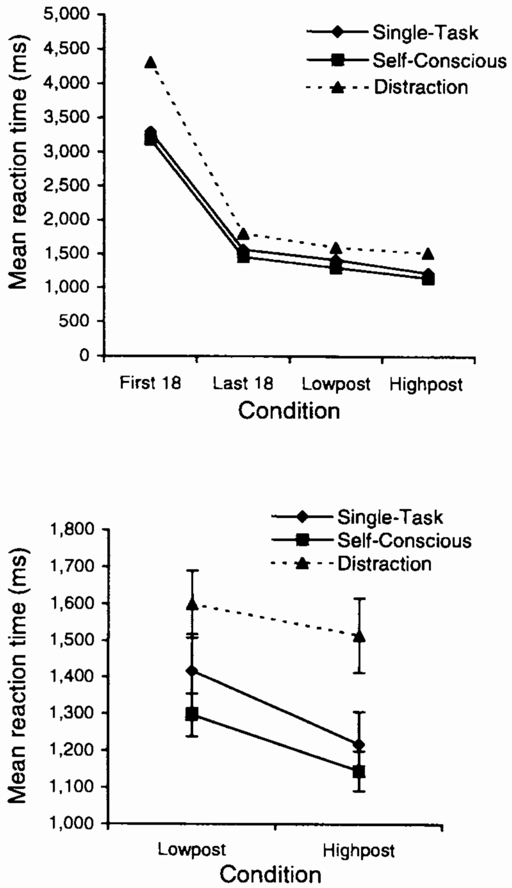

Figure 9. Mean reaction times (± *SE*) for arithmetic equations in the training and posttest conditions for each group in the alphabet arithmetic task (top) and posttest performance only (bottom). Lowpost = low-pressure posttest; Highpost = high-pressure posttest.

As an addendum, we should note that we believe the absence of choking in alphabet arithmetic to be a result of the fact that this type of skill does not become proceduralized with practice but moves from a declaratively accessible algorithm to retrieval of declaratively accessible facts into working memory. However, it may be that at higher levels of practice than we have examined, pressure-induced decrements in alphabet arithmetic performance could appear. Significant differences in performance between the dual-task training group and the other two groups in the alphabet arithmetic task at later stages of practice indicate that alphabet arithmetic performance was not yet fully automatized during the high-pressure situation (see Klapp et al., 1991). Thus, it may be that once differences in reaction times between these groups disappear, indicating full automatization, alphabet arithmetic will more closely resemble golf putting performance under pressure.[2]

To pursue this possibility, we conducted an analysis of alphabet arithmetic reaction times as a function of digit addend—the number of counting steps up the alphabet required by each equation. An effect of this variable has been used to diagnose the extent to which the control structure of alphabet arithmetic has shifted from the counting algorithm (which produces a significant effect of addend) to memory retrieval (which is independent of addend). This analysis produced a significant interaction of digit addend by pretest versus high-pressure posttest, $F(2, 104) = 7.06$, $p < .01$. Reaction time averaged across training groups increased markedly as a function of digit addend in the pretest ($M = 3{,}240$ ms for two-digit addend, $M = 3{,}690$ ms for three-digit addend, $M = 3{,}880$ ms for four-digit addend). Reaction time in the high-pressure posttest flattened considerably ($M = 1{,}185$ ms for two-digit addend, $M = 1{,}417$ ms for three-digit addend, $M = 1{,}310$ ms for four-digit addend), though there was still a significant effect of digit addend averaged across training groups. Further analysis of individual training groups showed that the significant effect of digit addend during the high-pressure posttest was a result of the single-task and distraction groups ($M = 1{,}084$ ms for two-digit addend, $M = 1{,}286$ ms for three-digit addend, $M = 1{,}268$ ms for four-digit addend and $M = 1{,}324$ ms for two-digit addend, $M = 1{,}679$ ms for three-digit addend, $M = 1{,}533$ ms for four-digit addend, respectively). Reaction time did not differ significantly as a function of addend for the self-consciousness group, $F(2, 34) = 2.73$, $p > .1$ ($M = 1{,}098$ ms for two-digit addend, $M = 1{,}211$ ms for three-digit addend, $M = 1{,}125$ ms for four-digit addend).

These data indicate that the self-consciousness training group achieved the most automated alphabet arithmetic performance, diagnosed by the relative independence they showed between alphabet arithmetic reaction time and digit addend. This is consistent with the prediction that increased monitoring of task components enhances skill acquisition among novices undergoing training (Anderson, 1987, 1993). If, similar to golf putting performance, those individuals trained under self-consciousness-raising conditions are immune to the detrimental effects of performance pressure, whereas those trained in single-task or distraction conditions are not, and the likelihood of choking increases as performance becomes more automated, then differences in high-pressure posttest reaction times should be apparent between the single-task and distraction groups on the one hand, and the self-consciousness group on the other. However, as can be seen in Figure 9, reaction time shows the same pattern across all three training groups. Thus it would appear that neither degree of automatization as measured by the effect of digit addend nor the differential impact of training condition bears on whether choking is observed in alphabet arithmetic. Increasing the amount of practice so that susceptibility to choking could be assessed in a completely automatized alphabet arithmetic skill would serve to further clarify this issue.

Putting Versus Alphabet Arithmetic

In order to further verify the differences in performance across posttests in the golf putting and alphabet arithmetic tasks, measurements taken in the putting and alphabet arithmetic tasks were converted into z scores. A 3 (single-task, distraction, self-consciousness) × 2 (low-pressure posttest, high-pressure posttest) × 2 (putting task, alphabet arithmetic) ANOVA was then performed, revealing a significant three-way interaction, $F(2, 102) = 6.5$, $p < .002$, $MSE = .08$. This confirms the pattern of data obtained above demonstrating that performance across the golf putting and alphabet arithmetic posttest conditions is different.

Discussion

Experiment 3 yielded three main results. First, following single-task practice, choking under pressure occurred in golf putting but not in alphabet arithmetic. Second, practice under dual-task conditions reduced performance in both tasks and altered practice benefits in alphabet arithmetic but did not alter either task's susceptibility to choking. Finally, practice under conditions intended to raise self-consciousness and execution-oriented achievement anxiety did not harm performance or change practice benefits relative to single-task practice in either skill but did inoculate putters against choking. Thus, at least at the levels of practice examined in the present study, choking arises in a task whose underlying knowledge base is thought to be procedural, but not one in which the underlying knowledge base is assumed to be more explicitly accessible. Furthermore, in terms of the effects of the two training regimens in the proceduralized task, it appears that when choking occurs, it results from explicit monitoring in response to self-consciousness and achievement anxiety. Performance pressure appears to elicit maladaptive efforts to impose step-by-step monitoring and control on complex, procedural knowledge that would have run off more automatically and efficiently had such monitoring not intervened. Practice at dealing with self-consciousness-raising situations counteracts this tendency.

We now turn to Experiment 4 in which we sought to replicate and extend Experiment 3's findings concerning the choking under pressure phenomenon. Because we were interested in the mechanisms governing choking, and alphabet arithmetic did not appear to show decrements in performance under pressure, in Experiment 4 we only examined the sensorimotor task of golf putting.

Experiment 4

In Experiment 4, the two possible sources of choking were examined at different stages of practice. It has been proposed that

---

[2] We thank Stuart Klapp for suggesting this possibility.

early in skill acquisition performance is supported by unintegrated control structures that are held in working memory and attended to in a step-by-step fashion (Anderson, 1987, 1993; Fitts & Posner, 1967). With practice, however, control evolves toward the type of integrated procedures assumed by explicit monitoring theory. According to explicit monitoring, tasks that follow this developmental trajectory should benefit from performance pressure early in learning yet be susceptible to choking at later stages of practice. Attention to task components is thought to be an integral part of novel skill performance (Proctor & Dutta, 1995). The explicit monitoring theory predicts that performance pressure prompts individuals to attend to skill execution processes. Thus, at low levels of practice, performance pressure should facilitate skill execution by prompting novice performers to allocate more attention to the task at hand.

According to distraction theory, however, performance pressure serves to create a dual-task environment. If individuals are attending to step-by-step execution processes in the early stages of skill learning, a distracting environment that draws attention away from the task at hand may harm performance. One could infer from the distraction hypothesis, then, that novice performers with little or no practice under divided attention conditions would be negatively affected by performance pressure, whereas those trained to a high skill level in a divided attention environment would not.

In Experiment 4 participants learned a golf putting task to a high level of skill under dual-task or self-consciousness-raising training conditions and were subjected to identical single-task low- and high-pressure situations both early and late in the training phase. If distraction is the reason for suboptimal performance under pressure, then individuals trained in either a dual-task or self-consciousness-raising environment should show performance decrements in pressure situations early in skill learning because, at this point, individuals in either training condition have not adapted to performing under divided attention conditions and do not possess a proceduralized skill response. Later in learning, however, those individuals trained in a dual-task environment will presumably be accustomed to performing under divided attention conditions and thus will not be affected by pressure, whereas the performance of those trained under conditions designed to increase anxiety and self-consciousness should decline. In contrast, if explicit monitoring is the reason for skill decrements under pressure, than at low levels of practice individuals trained under either distraction or self-consciousness-raising conditions should improve under pressure. If, as the explicit monitoring hypothesis predicts, pressure induces attention and control to skill performance, then novice performers may benefit from performance pressure in the initial stages of task learning. However, once the golf putting skill has become proceduralized later in practice, only those individuals who have adapted to performance anxiety and the demands to explicitly monitor skill performance (i.e., those trained under self-consciousness-raising conditions) should improve under pressure.

Method

Participants

Undergraduate students ($N = 32$) with little or no golf experience who were enrolled in an introductory psychology class at Michigan State University served as participants. Participants were randomly assigned to either a self-consciousness ($n = 16$) or dual-task distraction training group ($n = 16$).

Procedure

Participants completed the same golf putting task as in Experiment 3. Individuals took part in a 27-putt training condition followed by the first 18-putt single-task low-pressure posttest and 18-putt single-task high-pressure posttest. Participants then took part in a 225-putt training condition followed by a second 18-putt single-task low-pressure posttest and a second 18-putt single-task high-pressure posttest.

Distraction group. Participants completed the putting task in the same distraction training environment used in Experiment 3. Participants completed the first training condition of 27 putts, after which the tape recorder was turned off. Participants were then given a short break during which the experimenter computed the mean distance from the target of their last 18 putts.

Following the first training condition, participants completed the first 18-putt single-task low-pressure posttest, though they were not made aware of the test situation. Participants were then informed of their mean putting performance for the last 18 putts of the first training condition and given a scenario designed to create a high-pressure situation. Specifically, participants were told there would be two test situations in the present experiment. Participants were informed that they were about to take part in the first test situation and that the second test situation would take place toward the end of the experiment. Participants were given the same high-pressure scenario used in the golf putting task in Experiment 3, with the exception that they were told that they needed to improve their putting accuracy in two test situations to receive the monetary award. Participants then took the 18 putts constituting the first single-task high-pressure posttest. Following these putts, the experimenter informed participants that their putting average would be computed at the end of the experiment after both test situations had been completed.

Participants were then told that they would be taking another series of practice putts under the same dual-task conditions. The experimenter turned on the tape recorder, and participants completed the second training condition consisting of a total of 225 putts broken down into one training block of 72 putts, a training block of 81 putts, and another training block of 72 putts, with a short break after each block in which the tape recorder was turned off. When participants completed the second training condition, the tape recorder was turned off. Participants were then given a short break during which the experimenter computed the mean distance from the target of the last 18 putts in the training condition.

Participants then took part in the second 18-putt single-task low-pressure posttest. As in the first low-pressure posttest, participants were not made aware of the test situation. The experimenter then informed participants that they were about to take part in the second test situation and repeated the high-pressure scenario. Participants completed the second 18-putt single-task high-pressure posttest, were fully debriefed, and were given the monetary award regardless of their performance.

Self-consciousness group. Participants were set up at the first putting location and given instructions similar to those given to the distraction group regarding putting from the nine different locations on the green. Participants completed the putting task in the same self-consciousness-training environment used in Experiment 3. Participants completed the first training condition of 27 putts, after which the video camera was turned off and faced away from participants. Following the first training condition, participants were given a short break during which the experimenter computed the mean distance from the target of their last 18 putts. Participants then completed the first 18-putt single-task low-pressure posttest and high-pressure posttest identical to that of the distraction group.

Participants were then informed that they were going to complete another series of practice putts, again while being filmed by the video camera. The experimenter turned on the video camera, and participants completed the second training condition consisting of 225 total putts broken down into one training block of 72 putts, a training block of 81 putts, and another training block of 72 putts, with a short break after each block during which the video camera was turned off. When participants completed the second training condition, the camera was turned off and faced away from participants. Participants next completed the second 18-putt single-task low-pressure posttest and high-pressure posttest identical to that of the distraction group. Following the second high-pressure posttest, participants were fully debriefed and were given the monetary award regardless of their performance.

Results

Accuracy of putting was measured by the distance (in centimeters) away from the center of the target at which the ball stopped after each putt. Both groups improved significantly with practice as demonstrated by a 2 (distraction, self-consciousness) × 2 (mean distance from target of first 18 putts in first training condition, mean distance from target of last 18 putts in second training condition) ANOVA revealing a main effect of practice, $F(1, 30) = 58.63, p < .001, MSE = 31.73$; a nonsignificant main effect of training group ($F < 1$); and no interaction ($F < 1$; see Figure 10).

In the first low-pressure posttest (measured by the mean distance from the target of the 18 putts in the first low-pressure test), putting accuracy was similar across groups, $F(1, 30) = 1.40, p > .245, MSE = 26.15$. This homogeneity continued in the first high-pressure posttest (measured by the mean distance from the target of the 18 putts in the first high-pressure test) as confirmed by a one-way ANOVA again revealing no significant difference between groups ($F < 1$). These results are further supported by a 2 (distraction, self-consciousness) × 2 (first low-pressure posttest, first high-pressure posttest) ANOVA revealing a significant effect of test, $F(1, 30) = 17.73, p < .001, MSE = 12.38$; no significant effect of group, $F(1, 30) = 1.00, p > .323, MSE = 34.96$; and no interaction ($F < 1$). Direct comparisons of putting performance within each group showed that both the distraction and the self-consciousness groups significantly improved in putting accuracy from the first low-pressure posttest to the first high-pressure posttest, $t(15) = 3.76, p < .002$, and $t(15) = 2.30, p < .036$, respectively.

As can be seen in Figure 10, following the first high-pressure posttest, both training groups' performance accuracy decreased. This was confirmed by a 2 (distraction, self-consciousness) × 2 (first high-pressure posttest, first 18 putts of second training condition) ANOVA revealing a significant effect of condition, $F(1, 30) = 4.92, p < .034, MSE = 10.83$; no main effect of training group ($F < 1$); and no interaction ($F < 1$). Thus, whereas the first high-pressure posttest led to an increase in golf putting accuracy in comparison to the first low-pressure posttest in both the distraction and self-consciousness training groups, both groups showed performance decrements in the initial putts following the high-pressure situation. The explicit monitoring hypothesis suggests that performance pressure prompts individuals to explicitly monitor skill execution. Under this hypothesis, one would expect individuals in the initial stages of skill learning to improve under pressure as a result of increased attention to the novel demands of skill execution. However, once performance pressure and increased monitoring of performance are alleviated, a reduction in accuracy should occur.

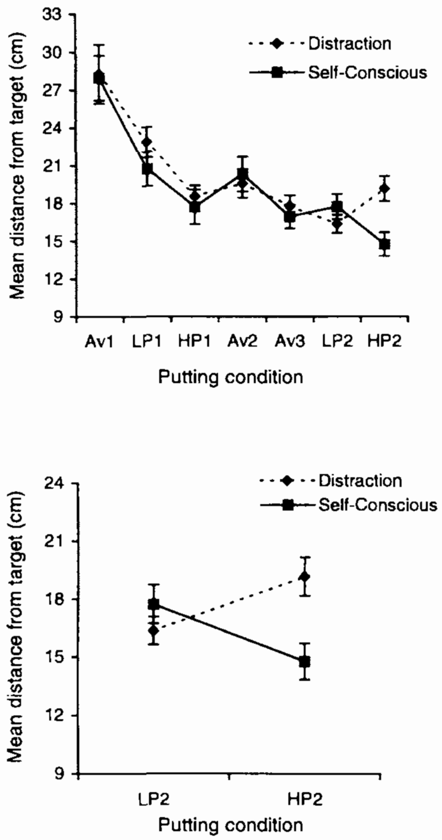

Figure 10. Top: Mean (± *SE*) distance from the target at which the ball stopped after each putt for the first 18 putts in the first training condition (Av1), the first low-pressure posttest (LP1), the first high-pressure posttest (HP1), the first 18 putts of the second training condition (Av2), the last 18 putts of the second training condition (Av3), the second low-pressure posttest (LP2), and the second high-pressure posttest (HP2). Bottom: Second posttest performance only.

In the second low-pressure posttest (measured by the mean distance from the target of the 18 putts in the second low-pressure test), putting accuracy was similar across groups, $F(1, 30) = 1.27, p > .269, MSE = 12.11$. This homogeneity disappeared in the second high-pressure posttest (measured by the mean distance from the target of the 18 putts in the second high-pressure test) as confirmed by a one-way ANOVA revealing a significant difference between groups, $F(1, 30) = 10.43, p < .003, MSE = 14.87$. The above results are further supported by a 2 (distraction, self-consciousness) × 2 (second low-pressure posttest, second high-pressure posttest) ANOVA revealing a significant interaction of Training Group × Second Posttest, $F(1, 30) = 24.16, p < .001, MSE = 5.55$. Direct comparisons of putting performance within each group showed that the distraction group significantly declined in putting accuracy from the second low-pressure posttest to the second high-pressure posttest, $t(15) = -2.79, p < .014$. In contrast, the self-consciousness group improved in putting accuracy from the second low-pressure posttest to the second high-pressure posttest, $t(15) = 4.84, p < .001$. Thus, as can be seen in Figure 10, both the distraction and self-consciousness groups improved from the first low- to the first high-pressure posttest. However, later in learning, those individuals trained in a dual-task environment showed decrements in performance under pressure, whereas those who learned the golf putting task under conditions designed to foster adaptation to a self-consciousness-raising environment that would increase achievement anxiety and explicit monitoring actually improved.

Discussion

The results of Experiment 4 once again are consistent with the predictions of the explicit monitoring theory of choking under pressure. Early in practice, regardless of training environment, performance pressure facilitated skill acquisition. However, as the golf putting skill became more proceduralized at later stages of practice, only those individuals who were accustomed to performing under conditions that heightened performance anxiety and the explicit monitoring of task processes and procedures were inoculated against the detrimental effects of performance pressure. These findings lend support to the notion that increased attention to the execution of a well-learned, complex skill may disrupt skill execution.

General Discussion

The purpose of the present study was to explore the cognitive mechanisms responsible for the disruption in the execution of a well-learned skill under pressure. Experiments 1 and 2 assessed the declarative accessibility of the knowledge representations governing real-time performance of golf putting at various levels of expertise. Results conformed in remarkable detail to predictions derived from current theories of automaticity and proceduralization of task performance as a function of practice. In Experiments 3 and 4 we looked at the phenomenon of choking under pressure in two very different tasks: the complex, sensorimotor skill of golf putting and a simpler, declaratively based alphabet arithmetic task. Results showed that choking under pressure occurred in golf putting but not in alphabet arithmetic, which demonstrates that sequential complexity, proceduralization, or both determine susceptibility to choking at the levels of practice we have examined. Furthermore, it was found that a particular training environment can eliminate choking when it does occur. Whereas single-task and dual-task practice left individuals in the golf putting task susceptible to performance decrements under pressure, self-consciousness training eliminated choking completely. Indeed, performers who experienced self-consciousness training actually improved under pressure—a highly desirable result. In addition to supporting the explicit monitoring hypothesis about why choking occurs, these experiments lead immediately to very practical ideas about training for real-world tasks in which serious consequences depend on good or poor performance in relatively public or consequential circumstances.

Properties of Tasks That are Susceptible to Choking

The pattern of results found in the present study speaks to the kinds of task properties that should be considered in the investigation of the pressure–performance relationship. Evidence of choking in the complex, proceduralized sensorimotor skill of golf putting but not in the simpler, declaratively based alphabet arithmetic task suggests at least three task properties that may be involved in choking. The first is task complexity. Masters (1992) argued that performance pressure prompts attention to skill execution, which results in the breakdown of task components. Well-learned complex skills may possess on-line control structures that run off as uninterrupted units. When attended to, these units may be broken down into a sequence of smaller, independent units, each of which must be run separately. As a result, performance slows and the transition between units creates an opportunity for error that was not present in the integrated control structure. Although skill breakdown may occur in complex, multistep tasks, this may not be the case in simple, one-step retrieval tasks. One-step retrieval tasks are not thought to consist of multiple integrated units and thus may not be susceptible to dismantling in the event of performance pressure. According to current theory, the automated form of alphabet arithmetic is such a one-step task (Klapp et al., 1991; Logan, 1988; Logan & Klapp, 1991).

However, our data indicate that alphabet arithmetic was not fully automated among our participants and hence was supported by some mixture of one-step fact retrieval and the multistep algorithm based on counting through the alphabet. Nevertheless, no hint of choking was observed in alphabet arithmetic. If task complexity is involved in choking, then performance decrements might have been expected at least on those trials supported by the multistep algorithm. In contrast, if task complexity does not affect susceptibility to choking, and instead choking occurs for alphabet arithmetic equations based completely on fact retrieval, then one might expect to see some indication of performance decrements for the portion of alphabet arithmetic equations that have switched to a fact retrieval mechanism. Either the former or latter of these possibilities should affect overall reaction time. However, as can be seen in Figure 9, alphabet arithmetic reaction time shows the same pattern both across training groups and between the low- and high-pressure conditions.

This leads to the second task property that may be involved in mediating the pressure–performance relationship. This is the degree to which task components become proceduralized with practice. Attention to the explicit processes involved in skill execution is thought to decrease as a function of skill level (Anderson, 1987, 1993; Fitts & Posner, 1967). As a result, skilled performances (e.g., complex sensorimotor tasks) are thought to operate largely outside of working memory. However, there may be certain skill types that rely on working memory for storage of control-relevant information during all stages of skill acquisition. Alphabet arithmetic is one such task. There is a substantial body of evidence demonstrating that performance on practiced alphabet arithmetic problems is not based on the establishment of a proceduralized version of the algorithm that controls action directly with no involvement from working memory. Instead, practice results in a shift from running through the steps of the algorithm in working memory to retrieving the answer into working memory from episodic memory (Klapp et al., 1991; Logan, 1988; Logan & Klapp, 1991). In either case the answer enters working memory, from where it controls the choice of an overt response. If choking is due to explicit monitoring, such skills should not be susceptible to decrements in performance because the practiced version of this task does not rely on the right kind of complex but proceduralized control structure. First, as already discussed, the control structure is too simple, and, second, control-relevant information always enters working memory and hence is always declaratively accessible. The alphabet arithmetic equation as a perceptual stimulus retrieves a single piece of information from episodic memory as the answer, and the elements of the control structure, the perceived equation and the retrieved answer, enter working memory rather than remaining outside the scope of attention, as does the relatively encapsulated procedure or motor program.

Finally, cognitive and motor tasks may differ in their susceptibility to breakdowns under pressure. The present study demonstrated that the sensorimotor task of golf putting, but not the cognitive task of alphabet arithmetic, was negatively affected by performance pressure. From these results it is tempting to conclude that choking may be confined to sensorimotor skills. However, such a conclusion is problematic in that it does not speak to the specific task characteristics that make a skill vulnerable to breakdowns under pressure. As mentioned in the introduction to this study, sensorimotor skills in both real-world and experimental settings are often associated with the choking phenomenon. Yet, the apparent prevalence of choking in sensorimotor domains may be not a function of sensorimotor skills per se but instead a result of specific task characteristics embedded in sensorimotor tasks that are susceptible to performance pressure (e.g., complexity and/or proceduralization). Furthermore, the notion that choking is limited to sensorimotor skills contrasts with research in the educational psychology literature demonstrating decrements in academic test performance under pressure in highly anxious individuals (Eysenck, 1979; Kahneman, 1973; Wine, 1971). Clearly such academic test performances do not have a large sensorimotor component, yet evidence of choking under pressure still emerges. Distraction theorists have suggested that suboptimal academic test performance results from the creation of a dual-task, distracting environment in which attention is divided between the task at hand and worries about the situation and its consequences. Thus, it remains a possibility that distraction as a mechanism for choking does hold for certain task types. It may be that pressure-induced distraction is detrimental to performance in tasks in which a large amount of information must be held in working memory and is susceptible to interference when attention is allocated to secondary sources (see, e.g., Tohill & Holyoak, 2000). This is a notion that is open to exploration in future work. However, it may also be the case that the types of problems encountered in these cognitive-based academic test situations have characteristics in common with many sensorimotor skills (e.g., complexity and/or proceduralizability) and are thus vulnerable to the same type of negative performance effects associated with the explicit monitoring of task execution processes. For example, Anderson (1993) suggested that complex cognitive skills such as algebra, geometry, and computer programming may become largely proceduralized in experts. We are pursuing this possibility in our laboratory.

Choking Research in Social Psychology

The present findings accord with research in the social psychology literature concerning the relationship between arousal, attention, and performance. It has been demonstrated that heightened anxiety and/or arousal levels induce self-focused attention (Fenigstein & Carver, 1978; Wegner & Giuliano, 1980). Wegner and Giuliano postulated that increments in arousal prompt individuals to turn their attention inward on themselves and current task performance in an attempt to seek out an explanation for their aroused state. Similar results have been found for skill execution in the presence of an audience. Butler and Baumeister (1998) recently demonstrated that supportive audiences were associated with unexpected performance decrements in the execution of complex, procedurally based tasks. The authors proposed that supportive audiences may increase attention to the processes involved in well-learned task performance, thus disrupting performance processes. And, finally, in a recent study investigating the effects of pressure on golf putting performance, Lewis and Linder (1997) found that pressure caused choking when participants had not adapted to performing in self-awareness-heightened environments—results similar to the present study. Furthermore, Lewis and Linder also found that decrements in performance could be alleviated through the use of a distractor (in this case, counting backward from 100) during real-time performance. Lewis and Linder suggested that attending to the distractor during on-line performance under pressure prevented participants from focusing attention inward on skill execution processes, thus alleviating the possibility of choking due to the "self-focus mediated misregulation" (p. 937) of performance. As can be seen from the literature just described, the notion that performance pressure induces self-focused attention, which in turn may lead to decrements in skill execution, is now a reasonably well-supported concept for proceduralized skills.

In conclusion, the findings of the four experiments in the present study lend support to the notion that pressure-induced attention to the well-learned components of a complex, proceduralized skill disrupts execution. Future research in this area is needed in order to illustrate the precise nature of the control structures that lead to decrements in performance under pressure. In addition, the generalizability of the present results to other task types and to different levels of practice must also be assessed. Further exploration of both the task and learning environments that mediate the pressure–performance relationship will serve to enhance our understanding of the choking under pressure phenomenon, which stands as an intriguing exception to the general rule that well-learned skills are robust and resistant to deterioration across a wide range of conditions.

References

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance and other domains. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 126–152). Cambridge, England: Cambridge University Press.

Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review, 89,* 369–406.

Anderson, J. R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94,* 192–210.

Anderson, J. R. (1993). *Rules of mind.* Hillsdale, NJ: Erlbaum.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46,* 610–620.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 686–700.

Butler, J. L., & Baumeister, R. F. (1998). The trouble with friendly faces: Skilled performance with a supportive audience. *Journal of Personality and Social Psychology, 75,* 1213–1230.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4,* 55–81.

Chi, M. T., Feltovitch, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5,* 121–152.

Craik, F. M., Govini, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General, 125,* 159–180.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

De Groot, A. (1978). *Thought and choice in chess.* The Hague, the Netherlands: Mouton. (Original work published 1946)

Dooling, D. J., & Christiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 428–436.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100,* 363–406.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 1–38). Cambridge, England: Cambridge University Press.

Eysenck, M. W. (1979). Anxiety learning and memory: A reconceptualization. *Journal of Research in Personality, 13,* 363–385.

Fenigstein, A., & Carver, C. S. (1978). Self-focusing effects of heartbeat feedback. *Journal of Personality and Social Psychology, 36,* 1241–1250.

Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 181–197.

Fitts, P. M., & Posner, M. I. (1967). *Human performance.* Belmont, CA: Brooks/Cole.

Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology, 87,* 621–636.

Hirst, W., Neisser, U., & Spelke, E. (1978). Divided attention. *Human Nature, 1,* 54–61.

Humphreys, M. S., & Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review, 91,* 153–184.

Jones, B. T., Davis, M., Crenshaw, B., Behar, T., & Davis, M. (1998). *Classic instruction in golf.* New York: Broadway.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Keele, S. W. (1986). Motor control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 2, pp. 1–60). New York: Wiley.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

Kimble, G. A., & Perlmuter, L. C. (1970). The problem of volition. *Psychological Review, 77,* 361–384.

Klapp, S. T., Boches, C. A., Trabert, M. L., & Logan, G. D. (1991). Automatizing alphabet arithmetic: II. Are there practice effects after automaticity is achieved? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 196–209.

Langer, E., & Imber, G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37,* 2014–2024.

Lesgold, A., Robinson, H., Feltovitch, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing X-ray pictures. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 311–342). Hillsdale, NJ: Erlbaum.

Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin, 23,* 937–944.

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

Logan, G. D., & Klapp, S. T. (1991). Automatizing alphabet arithmetic: I. Is extended practice necessary to produce automaticity? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 179–195.

Masters, R. S. (1992). Knowledge, knerves and know-how: The role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *British Journal of Psychology, 83,* 343–358.

McCandliss, B., & Carr, T. H. (1994, November). *Washing clothes twice: Do repetition and referential interpretation influence the same processes during reading?* Paper presented at the annual meeting of the Psychonomic Society, St. Louis, MO.

McCandliss, B., & Carr, T. H. (1996, November). *Repetition, comprehension, and intention to learn as sources of memory for text.* Paper presented at the annual meeting of the Psychonomic Society, Chicago.

Muter, P. (1980). Very rapid forgetting. *Memory & Cognition, 8,* 174–179.

Naveh-Benjamin, M., Craik, F. I., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1091–1104.

Norman, D. A., & Bobrow, D. J. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7,* 44–64.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Posner, M. I., & Rossman, E. (1965). Effect of size and location of informational transforms upon short-term retention. *Journal of Experimental Psychology, 70,* 496–505.

Priest, A. G., & Lindsay, R. O. (1992). New light on novice–expert differences in physics problem solving. *British Journal of Psychology, 83,* 389–405.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance.* Thousand Oaks, CA: Sage.

Raichle, M. E., Fiez, J. A., Videen, T. O., Macleod, A. M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional-anatomy during nonmotor learning. *Cerebral Cortex, 4,* 8–26.

Reimann, P., & Chi, M. T. (1989). Human expertise in complex problem solving. In K. J. Gilhooly (Ed.), *Human machine and problem-solving* (pp. 161–189). New York: Plenum.

Soloway, E., & Ehrlich, K. (1984). Empirical studies of programming knowledge. *IEEE Transactions on Software Engineering, SE-10,* 595–609.

Spelke, E., Hirst, W., & Neisser, U. (1976). Skills of divided attention. *Cognition, 4,* 215–230.

Squire, L. R., & Knowlton, B. J. (1994). The organization of memory. In H. Morowitz & J. L. Singer (Eds.), *The mind, the brain, and complex adaptive systems* (Vol. 22, pp. 63–97). Reading, MA: Addison-Wesley.

Tohill, J. M., & Holyoak, K. (2000). The impact of anxiety on analogical reasoning. *Thinking and Reasoning, 6,* 27–40.

Van Lehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527–580). Cambridge, MA: MIT Press.

Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 261–285). Hillsdale, NJ: Erlbaum.

Wegner, D. M., & Giuliano, T. (1980). Arousal-induced attention to self. *Journal of Personality and Social Psychology, 38,* 719–726.

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin, 76,* 92–104.

Woolfolk, R. L., Murphy, S. M., Gottesfeld, D., & Aitken, D. (1985). Effects of mental rehearsal of task motor activity and mental depiction of task outcome on motor skill performance. *Journal of Sport Psychology, 7,* 191–197.

Zbrodoff, N. J., & Logan, G. D. (1986). On the autonomy of mental processes: A case study of arithmetic. *Journal of Experimental Psychology: General, 115,* 118–130.

Appendix A

Questionnaires

Generic Questionnaire—Experiments 1 and 2

Certain steps are involved in executing a golf putt. Please list as many steps that you can think of, in the right order, which are involved in a typical golf putt.

Episodic Questionnaire—Experiment 1[a]

Pretend that your friend just walked into the room. Describe the last putt you took, in enough detail so that your friend could perform the same putt you just took.

Episodic Questionnaire—Experiment 2[b]

Pretend that your friend just walked into the room. Describe the last putt you took, in enough detail so that your friend could duplicate that last putt you just took in detail, doing it just like you did.

---

[a] Additional explanation was given in order to make it clear that what was being asked for was a "recipe" or "set of instructions" that would allow the putt to be duplicated in all its details by someone who had not seen it. Golf team members were told that the friend was not another golf team member but someone with an ordinary knowledge of the game. This was done to prevent excessive use of jargon or "in-group" shorthand, in an attempt to equate the need for knowledge that would be assumed by the describers across groups.

[b] This episodic questionnaire was changed slightly from Experiment 1 in an attempt to elicit the most detailed episodic descriptions possible from participants.

Appendix B

Steps Involved in a Typical Golf Putt

1. Judge the line of the ball.
2. Judge the grain of the turf.
3. Judge the distance and angle to the hole.
4. Image the ball going into the hole.
5. Position the ball somewhere between the center of your feet. You should be able to look straight down on top of the ball.
6. Align shoulders, hips, knees, and feet parallel and to the left of the target (e.g., image railroad tracks from the ball to the cup—feet outside the tracks, the ball in the middle).
7. Grip—thumbs should be pointed straight down, palms facing each other, a light grip.
8. Posture—stand tall enough so that if you were to practice putting for 30 minutes you would not experience a stiff or sore back.
9. Arms—should hang naturally and be relaxed.
10. Hands—should be relative to ball position. Hands should be slightly in front of the ball.
11. Head position—eyes should be positioned directly over the ball.
12. Weight—distribute weight evenly, about 50-50, or with a little more weight on the left foot.
13. Backswing—swing the club straight back. The distance back that the club goes must equal the through stroke distance.
14. Stroke—the club must accelerate through the ball. Finish with the "face" of the club head pointing directly at the target.
15. Length of the stroke—it is better to err to a shorter more compact stroke rather than a longer stroke.
16. Stroke direction—straight back and straight through.
17. Stroke rhythm—not too fast and not too slow.
18. Keep head and lower body stationary throughout stroke and swing with the arms.
19. Wrists—should not break during the stroke.
20. Arms and shoulders—should do most of the work.
21. Head/trunk/hips/legs—should remain still during the stroke.
22. Watch the ball go into the hole.

Received February 11, 2000
Revision received June 28, 2000
Accepted January 29, 2001 ■