

# Choking Under Pressure: Multiple Routes to Skill Failure

Marci S. DeCaro  
Miami University and Vanderbilt University

Robin D. Thomas  
Miami University

Neil B. Albert  
Spencer Foundation and The University of Chicago

Sian L. Beilock  
The University of Chicago

Poor performance in pressure-filled situations, or “choking under pressure,” has largely been explained by two different classes of theories. Distraction theories propose that choking occurs because attention needed to perform the task at hand is coopted by task-irrelevant thoughts and worries. Explicit monitoring theories claim essentially the opposite—that pressure prompts individuals to attend closely to skill processes in a manner that disrupts execution. Although both mechanisms have been shown to occur in certain contexts, it is unclear when distraction and/or explicit monitoring will ultimately impact performance. The authors propose that aspects of the pressure situation itself can lead to distraction and/or explicit monitoring, differentially harming skills that rely more or less on working memory and attentional control. In Experiments 1–2, it is shown that pressure that induces distraction (involving performance-contingent outcomes) hurts rule-based category learning heavily dependent on attentional control. In contrast, pressure that induces explicit monitoring of performance (monitoring by others) hurts information-integration category learning thought to run best without heavy demands on working memory and attentional control. In Experiment 3, the authors leverage knowledge about how specific types of pressure impact performance to design interventions to eliminate choking. Finally, in Experiment 4, the selective effects of monitoring-pressure are replicated in a different procedural-based task: the serial reaction time task. Skill failure (and success) depends in part on how the performance environment influences attention and the extent to which skill execution depends on explicit attentional control.

*Keywords:* pressure, category learning, working memory, attention, serial reaction time task

People often find themselves in high-stakes situations where performing their best carries implications for future opportunities and success. Whether it is a high school student taking the SAT, a golfer playing to make the cut for the PGA tour, or a violinist auditioning for an orchestra, high-level performance in important situations is crucial for advancement in most facets of life. In these types of high-stakes situations, the desire to perform as well as possible is thought to create *performance pressure* (Baumeister, 1984; Hardy, Mullen, & Jones, 1996; Beilock & Carr, 2001). Unfortunately, in both real-world (e.g., Dandy, Brewer, & Tottman, 2001; Davis & Harvey, 1992; Dohmen, 2008; Forgas, Brennan, Howe, Kane, & Sweet, 1980; Heaton & Sigall, 1989; Paulus,

Shannon, Wilson, & Boone, 1972) and laboratory situations (e.g., Beilock, 2008; Beilock & Gray, 2007), this pressure to attain performance success often causes people to perform below their actual abilities. The term *choking under pressure* describes this phenomenon. Choking is not just poor performance. Rather, it is performing more poorly than expected, given one’s skill level, in situations where performance pressure is at a maximum (Beilock & Gray, 2007).

Understanding why choking occurs is important for devising training regimens to alleviate it. Yet investigations into unwanted skill failure can do a lot more. Understanding skill failure and success under pressure may shed light on the similarities and differences in the cognitive control structures underlying a diverse set of skills, ranging from math problem solving to golf putting. Moreover, by uncovering the mechanisms governing pressure-induced failure, we can also further our understanding of how emotional and motivational factors combine with memory and attention processes to impact skill learning and performance. An understanding of how the performance environment alters cognitive processes not only advances our understanding of the “choking under pressure” phenomenon specifically but also provides insight into related situations in which performance inadvertently falters, ranging from test anxiety (Ashcraft, 2002; Wine, 1971) to the threat of conforming to a negative stereotype (i.e., stereotype threat; Steele, 1997).

---

This article was published Online First May 16, 2011.

Marci S. DeCaro is now in the Department of Psychology and Human Development at Vanderbilt University; Robin D. Thomas, Department of Psychology; Neil B. Albert, Department of Psychology; Sian L. Beilock, Department of Psychology.

This research was supported by National Science Foundation CAREER Grant DRL-0746970 to Sian L. Beilock and National Science Foundation Grant BCS-0544688 to Robin D. Thomas. The authors thank Krista Carlson for her assistance with response modeling.

Correspondence concerning this article should be addressed to Marci S. DeCaro, Department of Psychology and Human Development, 230 Appleton Place, Peabody #552, Vanderbilt University, Nashville, TN 37203. E-mail: marci.decaro@vanderbilt.edu

## Mechanisms of Skill Failure

Why does poor performance sometimes occur in high-pressure situations? Two different theories have been proposed to answer this question. *Distraction theories*, originating from work in academic testing situations, propose that high-pressure situations harm performance by diverting individuals' attention to task-irrelevant thoughts, such as worries about the situation and its consequences (Beilock & Carr, 2001; Lewis & Linder, 1997; Wine, 1971). Pressure essentially creates a dual-task environment in which situation-related worries compete with the attention needed to execute the task at hand.

Attention is a key component of *working memory* (Engle, 2002), a short-term memory system responsible for actively maintaining a limited amount of task-relevant information while inhibiting irrelevant information (Engle, 2002; Miyake & Shah, 1999). According to distraction theories, because high-pressure situations coopt attentional resources, tasks that rely heavily on working memory should be most negatively impacted under pressure. This is exactly what has been found (Beilock & DeCaro, 2007; Gimmig, Huguet, Caverni, & Cury, 2006; Markman, Maddox, & Worthy, 2006). For example, Beilock, Kulp, Holt, and Carr (2004) demonstrated that math problems heavily dependent on working memory (i.e., requiring the online maintenance and manipulation of intermediate problem steps) were solved less accurately in a high-pressure test compared with a low-pressure test. In contrast, math problems that were highly practiced and thus could be directly retrieved from long-term memory (Logan, 1988), circumventing demanding computations in working memory, were performed just as well in low- and high-pressure situations.

Although there is evidence that pressure induces failure by distracting attention away from skill execution, a very different class of theories has also been put forth to explain skill failure as well. *Explicit monitoring* or skill-focus theories suggest that pressure increases self-consciousness about performing correctly, which in turn leads performers to focus their attention on skill execution to ensure an optimal outcome (Beilock & Carr, 2001). Explicit attention to step-by-step processes is thought to disrupt the learning and execution of proceduralized processes that normally run outside of conscious awareness (Baumeister, 1984; Beilock, Bertenthal, McCoy, & Carr, 2004; Beilock & Carr, 2001; Kimble & Perlmutter, 1970; Langer & Imber, 1979; Masters, 1992).

Support for explicit monitoring theories is found primarily in proceduralized skills ranging from golf putting (Beilock & Carr, 2001; Lewis & Linder, 1997; Masters, 1992), to hockey dribbling (Jackson, Ashford, & Norsworthy, 2006), to baseball batting (R. Gray, 2004). For instance, R. Gray (2004) examined how expert baseball players batted in a baseball simulator in both low-pressure and high-pressure conditions. Gray found an increase in batting errors and movement variability under high pressure, relative to the low-pressure situation. It is notable that this pressure-induced batting failure was accompanied by improvement in the players' ability to judge the direction their bat was moving during skill execution. Specifically, when batters were asked to judge whether their bat was traveling upward or downward at the moment an intermittent tone sounded, these judgments were more accurate under pressure. These results suggest that, under pressure, the batters focused explicitly on the components of the batting skill. This focus of attention led to more accurate judgments about bat

direction but disrupted proceduralized batting processes and hurt hitting performance overall.

Thus, distraction and explicit monitoring theories of choking under pressure pose very different mechanisms of skill failure. Whereas distraction theories suggest that pressure harms performance by shifting attention and working memory resources *away from* execution, explicit monitoring theories suggest that pressure shifts too much attention *toward* skill processes and procedures. How can pressure do both? One possibility is that pressure coopts working memory when individuals are performing demanding cognitive tasks, whereas it induces attention to skill processes during proceduralized motor skill execution. But it seems odd to suggest that a high-pressure situation would exert different effects simply depending on whether one is holding a pencil or a baseball bat in one's hands.

We believe the answer to the above question lies in aspects of the performance environment itself. High-pressure situations may actually involve multiple components and thus exert multiple effects, leading to distracting thoughts, explicit monitoring, or even both, depending on specific elements of the stress imposed. Whether or not performance will fail, and how this failure will come about, depends on aspects of the pressure situation and the attentional demands of the task being performed. In the current work, we examine whether different types of pressures exert distinct effects and, moreover, whether these pressures carry different implications for skill success and failure based on the type of task being performed.

## The Pressure Situation

Although most investigations of performance under pressure have largely ignored the makeup of the pressure situation itself, real-world pressure situations (and the laboratory pressure manipulations that imitate them) have multiple elements. Individuals might be watched by a teacher, audience, or video camera; people may try to obtain a personally important title, scholarship, or monetary reward; or people may be out for a high test score. Even though these distinct elements of a high-pressure situation may elicit feelings of pressure and anxiety, a closer look reveals subtle differences among them.

For instance, the pressure of being watched by others—which we refer to as *monitoring pressure*—may increase attention to skill processes and procedures, particularly when one's performance is evaluated in some manner. This notion is supported by work in social psychology showing that the presence of an audience, video camera, or mirror increases self-consciousness or self-awareness (e.g., Carver & Scheier, 1978; Davis & Brock, 1975; Duval & Wicklund, 1972; Geller & Shaver, 1976). On the other hand, pressure induced by offering an incentive if a certain outcome is achieved—which we refer to as *outcome pressure*—may serve to shift performers' focus of attention to the situation and its consequences. A greater metacognitive awareness of the performance situation may lead people to simulate different outcome possibilities or think about how they are measuring up during performance (e.g., "I haven't even gotten one right so far"; Sarason, 1972, p. 411), diverting attention away from skill execution.

A given pressure situation may therefore differentially emphasize outcome pressures or monitoring pressures. In many high-pressure situations, aspects of both may be present (e.g., Beilock,

Kulp, et al., 2004; Beilock & Carr, 2005; Beilock & DeCaro, 2007; R. Gray, 2004; Hardy et al., 1996), simultaneously disrupting working memory availability and directing what attention that remains in ways that are counterproductive. Thus, pressure may lead to skill failure in multiple ways, depending in large part on features of the performance environment itself.

We are aware of no previous research that has attempted to tease apart the pressure situation to explore whether multiple pressure elements exert systematically different effects on performance. However, a close examination of one recent study offers some preliminary support for our ideas. Markman et al. (2006) investigated performance under pressure on category-learning tasks (i.e., rule-based and information-integration category tasks) that differ in their reliance on attentional control. For *rule-based* category-learning tasks, individuals must discover an easily verbalizable rule to categorize stimuli into two groups (e.g., all items with a particular feature belong to Category A). Doing so involves explicitly testing various hypotheses about category membership and therefore relies on working memory and executive attention, similar to many of the demanding cognitive tasks (e.g., math problem solving) explored in the pressure literature to date (DeCaro, Thomas, & Beilock, 2008; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). *Information-integration* category-learning tasks, on the other hand, rely on the integration of multiple stimulus dimensions at a predecisional stage. Categories are similarity-based, and categorization is thought to rely on procedural learning (i.e., accumulating stimulus–response associations) that occurs largely outside of attentional control. This task representation is said to be similar to that of the complex sensorimotor skills (e.g., golf putting) studied under pressure (Maddox & Ashby, 2004).

In their work, Markman et al. (2006) looked at both rule-based and information-integration category learning under pressure. The authors found that rule-based category learning was worse under high-pressure compared with low-pressure conditions. However, the opposite was found for information-integration category learning: Performance was actually better under high-pressure than low-pressure conditions. These findings can be readily explained with the distraction theory of choking: Pressure consumed the working memory and attention needed for hypothesis testing, which in turn harmed rule-based category learning. In contrast, information-integration category learning, relying instead on similarity-based responses elicited by the procedural learning system, benefited when attention was diverted from performance (for related work in complex sensorimotor skills, see Beilock, Carr, MacMahon, & Starkes, 2002).

In light of explicit monitoring theories, however, the latter findings are unexpected: Why doesn't information-integration category learning fail, via pressure-induced explicit monitoring? According to the above hypothesis concerning multiple types of pressure, the answer has to do with the pressure situation itself—Markman et al. (2006) only employed an outcome-based incentive scenario. Individuals were told that, by exceeding a performance criterion, they could earn money for both themselves and a partner. There was no aspect of monitoring pressure (e.g., social evaluation, video camera) that we propose would lead to explicit monitoring. Had monitoring pressure been present, we anticipate that information-integration category learning would falter. Such a result would provide evidence for multiple routes to skill failure under pressure.

## Current Experiments

Across four experiments, we examine whether different types of pressures exert distinct effects and, moreover, whether these pressures carry different implications for skill success and failure based on the control structure of the task being performed. Because the category-learning domain uniquely affords the opportunity to solicit attention-demanding learning processes in one task and less attending-demanding processes in a different task, while holding all other aspects of the learning situation invariant (e.g., same stimuli, same general learning paradigm), we start in Experiment 1 by showing that categorization tasks similar to the ones used by Markman et al. (2006) can differentially fall prey to distraction and explicit monitoring. In Experiment 2 we examine whether outcome pressure hurts rule-based categorization, in line with distraction theories of pressure (as did Markman et al., 2006), and monitoring pressure hurts information integration-integration categorization, in accordance with explicit monitoring theories of pressure (similar to work in proceduralized sensorimotor skills). In Experiment 3 we leverage findings of the first two experiments to examine how pressure-induced performance decrements can be alleviated in both types of categorization tasks. Finally, in Experiment 4 we extend these findings to proceduralized performance in another domain, examining whether monitoring pressure, as opposed to outcome pressure, harms skilled performance on the serial reaction time task (SRTT; Nissen & Bullmeier, 1987).

### Experiment 1

Individuals completed both rule-based and information-integration category-learning tasks in a baseline single-task condition followed by one of two types of secondary task conditions: (a) a distracting secondary task designed to divert attention away from category learning or (b) an explicit monitoring secondary task intended to prompt attention to the step-by-step components of learning.

If rule-based category learning relies on attention and working memory to discover and maintain different hypotheses about category membership (Waldron & Ashby, 2001), then it should take longer to learn this type of category structure when simultaneously performing a distracting versus explicit monitoring secondary task. The opposite should occur for information-integration category learning, where optimal performance is not thought to occur via attention-demanding hypothesis testing but by processes that make little demand on attention and working memory (Maddox & Ashby, 2004). Learning this type of category structure should be unaffected by a distracting secondary task but impaired when performing an explicit monitoring secondary task. Such findings would set the stage in Experiment 2 to test whether there are systematically different types of pressures and to explore how these pressures might exert their impact.

### Method

**Participants.** Undergraduate students ( $N = 103$ ) at a large U.S. Midwestern university served as participants ( $M$  age = 19.1 years,  $SD = 2.7$ ). Participants had no reported colorblindness. Individuals were randomly assigned to a distracting secondary task condition ( $n = 46$ ) or an explicit monitoring secondary task condition ( $n = 57$ ).

**Procedure.** After giving informed consent, participants completed the category-learning task individually on the computer (see below for details). Individuals were instructed to place each stimulus into either Category A or Category B by pressing one of two marked keys on the computer keyboard. Following each categorization selection, immediate feedback was displayed, with the word “correct” or “incorrect” appearing directly below the stimulus, until the individual pressed the spacebar to continue to the next trial. The screen then went blank for a 1,500-ms intertrial interval. Once participants reached a learning criterion of eight correct categorization trials in a row or a 200-trial maximum, they exited that particular category structure (Waldron & Ashby, 2001). They were then given a rest period during which they were informed that they would be given a new category structure. Only individuals who successfully reached this learning criterion prior to the 200-trial maximum in the single-task baseline block (see below) were included in the current work. This allowed us to examine the impact of the experimental manipulations on performance for only those individuals who demonstrated that they were able to perform these types of categorization tasks in the first place.

All participants performed the same four category structures, separated into two blocks. Within each block, individuals saw two different types of category structures: one rule-based category (R) and one information-integration category (I; see below). Thus, four global orders were possible (i.e., RI RI, IR IR, RI IR, IR RI) and counterbalanced across participants.<sup>1</sup> The specific category structure (e.g., which rule-based task came first) was randomly selected without replacement across participants, and the categorization stimuli within each category structure were randomly selected with replacement across participants.

The first block served as a single-task baseline. Before the second block, participants read instructions for either the distracting or explicit monitoring secondary tasks, depending on their assigned condition. Then the second block was performed concurrently with one of the two secondary tasks. After the second block was complete, individuals completed a series of questionnaires and were thanked and debriefed.

**Category learning task.** Stimuli were adapted from Waldron and Ashby (2001). Each was a square with one or two symbols embedded within it. Sixteen stimuli were constructed by taking the factorial combination of four dimensions, two levels each: square/background color (yellow or blue), embedded symbol shape (circle or square), symbol color (red or green), and number of embedded symbols (one or two). The four category structures each used all 16 stimuli but differed in the mapping from stimuli to responses.

Rule-based categories had one relevant dimension, affording an easily verbalizable rule (e.g., “If the embedded symbol is red, choose Category A; if the symbol is green, choose Category B”). Because previous studies (e.g., Waldron & Ashby, 2001) found no differences in performance depending on the dimension selected to be relevant, symbol color was randomly chosen as the relevant dimension for one rule-based structure and symbol shape for the other.

Information-integration categories involved three relevant dimensions and one irrelevant dimension (randomly determined as background color for one information-integration structure and number of embedded symbols for the other). The three relevant dimensions were labeled X, Y, and Z, and each binary value of these three dimensions was randomly assigned either a  $-1$  or a  $+1$

(e.g., a red symbol =  $-1$ , and a green symbol =  $+1$ ). Stimuli were then categorized according to the following rule: If  $\text{value}(X) + \text{value}(Y) + \text{value}(Z) > 0$ , classify as Category A; otherwise classify as Category B (Waldron & Ashby, 2001). Because this rule would be very difficult to derive verbally over a series of individual learning trials, it is more likely that these dimensional values are integrated at a predecisional stage, presumably without access to conscious awareness (Ashby & Maddox, 2005).

**Secondary tasks.** A letter-monitoring task served as the *distracting secondary task*. At the beginning of each trial, a 200-ms fixation point (a plus sign in the center of the screen) was replaced by a randomly selected alphanumeric letter appearing for 2,000 ms. Individuals were instructed to press the space bar if the displayed letter was an “S.” If the letter was not an “S,” they were instructed to do nothing (i.e., a go/no-go task). The letter “S” was shown more often than any other individual letter (37.5% of the time if all 200 trials were completed).

After either the spacebar was pressed or the 2,000-ms time interval had passed, the word “correct” or “incorrect” replaced the letter stimulus for 1,000 ms, providing immediate feedback for the secondary task. Feedback was used to emphasize the importance of the letter-monitoring task for the participant. The screen then went blank for 1,000 ms, and the categorization trial resumed as described above. Only participants who maintained 90% accuracy on the letter-monitoring task were included to ensure that participants were allocating attention to both the primary and secondary tasks.

A confidence-judgment task served as the *explicit monitoring secondary task*. This task was intended to induce individuals to explicitly monitor the component steps of the categorization process—how they went about deciding the stimuli should go in a particular category. At the beginning of each trial, participants were shown a categorization stimulus along with a confidence-rating prompt. They were asked to first think about how they were going to categorize the stimulus and then rate how confident they were in their category selection, by typing the number corresponding to their confidence rating on a scale ranging from 1 (*not at all confident*) to 7 (*extremely confident*). After selecting their confidence rating, the prompt disappeared, and individuals categorized the stimulus as usual.

**Questionnaires.** Following the category-learning task, participants rated how important they felt it was to perform at a high level during the last two sets of category-learning tasks (i.e., the secondary task block), on a scale ranging from 1 (*not at all important*) to 7 (*extremely important*). We only included participants who responded at the midpoint or higher in our analyses. Because performance pressure, by definition, only occurs when individuals feel it is important to perform their best (Baumeister, 1984), reporting at least moderate task importance is often used as a criterion for study participation in experiments exploring the choking phenomenon (Beilock & Gray, 2007). Because we implement this criterion in Experiments 2–4, where pressure is explicitly manipulated, for consistency we implement it in Experiment 1 as well. Last, individuals completed a brief demographics questionnaire and were thanked and debriefed.

<sup>1</sup> In both Experiments 1 and 2, preliminary analyses indicated that order did not moderate the key Condition  $\times$  Category Structure interaction.

## Results and Discussion

**Categorization performance.** Of primary interest was how many categorization trials individuals performed before reaching the learning criterion of eight correct categorization trials in a row. Trials to criterion were log transformed because of a positive skew in the distribution (Tabachnick & Fidel, 1996), which is common for category-learning tasks (e.g., Waldron & Ashby, 2001; DeCaro et al., 2008). These data are included in Appendix A for all three experiments. However, because we were interested in understanding how adding a secondary task changes category-learning performance, our dependent variable was a difference score created by subtracting Block 1 (single-task baseline) log-transformed trials to criterion from Block 2 (secondary task block) log-transformed trials to criterion. This was done separately for the rule-based and information-integration category structures to examine the impact of the secondary tasks on each type of category structure. Thus, higher scores on this difference measure indicate worse category-learning performance (i.e., taking more trials to learn the categories) during the secondary task block compared with the single-task baseline.

These difference scores were submitted to a 2 (Category Structure: rule-based, information-integration)  $\times$  2 (Secondary Task Type: distracting, explicit monitoring) analysis of variance (ANOVA), with the last factor between subjects. There was no main effect of category structure or secondary task condition ( $F_s < 1$ ). However, there was a significant Category Structure  $\times$  Secondary Task Condition interaction,  $F(1, 101) = 4.78, p = .03, MSE = .26$ .

To understand this interaction, we examined each category structure separately. As can be seen in Figure 1, for rule-based category learning, the difference in learning from the baseline to distracting secondary task block was significantly larger than zero ( $M = .17, SE = .07, t(45) = 2.21, p = .03$ ). Rule-based category learning, heavily reliant on attention and working memory for optimal performance, was worse when individuals were distracted. These results are consistent with previous research using an

attention-demanding dual task during rule-based category learning (i.e., Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). Simultaneously performing two attention-demanding tasks typically leads to worse performance than if either task was performed alone (Baddeley & Logie, 1999). In contrast, rule-based category learning was not impacted in the explicit monitoring secondary task—the difference between baseline and explicit monitoring rule-based learning blocks was essentially zero ( $M = -.03, SE = .07, t(56) = -.48$ ). Asking individuals to think explicitly about each rule-based categorization judgment and rate their confidence in their selection did not impact performance.

In terms of information-integration category learning, the difference from the single-task baseline to the explicit monitoring secondary task block was significantly greater than zero ( $M = .17, SE = .07, t(56) = 2.47, p < .02$ ). Information-integration category learning was worse when individuals were asked to make confidence judgments by explicitly attending to the steps of the categorization process. Such findings are consistent with those of Maddox, Love, Glass, and Filoteo (2008), who demonstrated that additional feedback that draws attention to explicit aspects of categorization can impair information-integration category learning. These results also resemble findings in complex sensorimotor skills (Wulf, 2007). Secondary tasks requiring individuals to monitor the processes of performance (e.g., attending to the hands or feet) disrupt skilled field-hockey dribbling (Jackson et al., 2006), soccer dribbling (Beilock et al., 2002), and baseball batting (R. Gray, 2004). Finally, the distracting letter-monitoring task had no impact on information-integration category learning ( $M = .06, SE = .08, t(45) = .78$ ). This finding is consistent with work with proceduralized sensorimotor skills as well. Performing a secondary tone-monitoring task during hockey or soccer dribbling does not negatively impact well-learned skill execution (Beilock et al., 2002; Jackson et al., 2006).

**Categorization strategies.** We also examined the strategies individuals used to learn the information-integration categories in the two secondary task conditions. We focused specifically on the information-integration task for two reasons. First, although individuals are generally thought to learn the information-integration categories to criterion by relying on the integration of multiple category dimensions at an implicit predecisional stage (Ashby & Maddox, 2005; Waldron & Ashby, 2001), previous work using the same category-learning task has shown that, in some circumstances, people can successfully learn these categories to a criterion of eight correct categorization trials in a row by employing simple, one-dimension rules (e.g., “Categorize all items with a green symbol as Category A”; DeCaro, Carlson, Thomas, & Beilock, 2009; Tharp & Pickering, 2009). Such simple rules, as well as more complex explicit rules involving two or three stimulus dimensions, can lead to a relatively high degree of accuracy (75%–87.5% of the stimuli), and therefore may provide alternative ways in which participants can attain the learning criterion. However, using explicit rules can also prevent responding based on the more optimal, procedural, system and therefore could also lead to poorer performance. Thus, by modeling strategies, we hoped to shed light on how individuals were solving the information-integration task.

Second, if individuals are more likely to pay explicit attention to performance in the explicit monitoring secondary task condition relative to a single-task baseline condition, then one might also

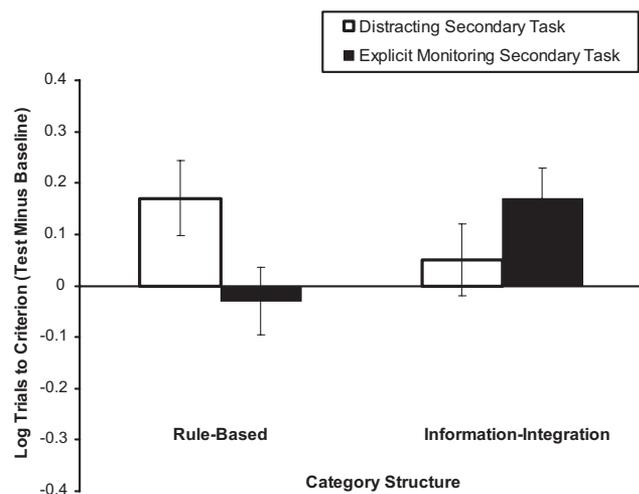


Figure 1. Mean trials to criterion (log transformed) difference score (test minus baseline) as a function of category structure and secondary task type. Error bars represent standard errors.

expect them to rely more heavily on explicit rules. In contrast, people given a distracting secondary task should not use explicit rules more (and may even use them less) than they do during a single-task baseline condition. Thus, modeling strategy data allowed us to specifically test our prediction that the secondary task conditions would alter how individuals allocated attention to performance (for response modeling details, see Appendix B).

Figure 2 displays the degree to which participants changed their strategies during information-integration category learning from the single-task baseline to each secondary task condition (secondary task minus baseline); a higher number represents a greater proportion of strategy use in the secondary task condition relative to the single-task condition. As shown in Figure 2, participants in the explicit monitoring task condition used a greater proportion of both one-dimension ( $M = .07$ ,  $SE = .04$ ),  $t(56) = 2.52$ ,  $p < .02$ , and two-dimension rules ( $M = .03$ ,  $SE = .01$ ),  $t(56) = 2.39$ ,  $p = .02$ . In addition, the more individuals implemented these rule-based strategies, the more trials they took to learn the information-integration task ( $r_{1\text{-Dim.}} = .89$ ,  $p < .001$ ,  $r_{2\text{-Dim.}} = .79$ ,  $p < .001$ ). Participants in the explicit monitoring condition were also less likely to implement the “optimal” strategy (i.e., categorizing the stimuli as set up by the experimenters, at 100% accuracy;  $M = -.11$ ,  $SE = .04$ ),  $t(56) = -2.51$ ,  $p < .02$ . Greater use of the optimal strategy was associated with taking fewer trials to learn the information-integration task ( $r = -.93$ ,  $p < .001$ ). No other changes in response strategies were found for the explicit monitoring condition. Moreover, participants in the distracting secondary task condition did not change any strategies relative to baseline.

Thus, explicit monitoring leads to shifts in the strategies used to perform the information-integration categorization task. Moreover, because participants in the explicit monitoring condition showed a greater decrement than participants in the distracting secondary

task condition, these strategy data are also consistent with the idea that using explicit verbal rules on the information-integration task hurts learning. Individuals used these rules more often in the explicit monitoring secondary task condition and performed more poorly.

The response-modeling data also support the idea that the optimal strategy is indeed less attention-demanding and likely procedural in nature. Use of the optimal strategy decreased in the skill-focus secondary task condition—the same condition in which other explicit strategies increased. In addition, the use of the optimal strategy did not decrease in the distracting secondary task condition—if the optimal strategy depended heavily on attention (e.g., a complex rule-plus-exception strategy), then one would expect the use of this strategy to decrease under distraction. Taken together, these results provide support for the less attention-demanding nature of the optimal information-integration strategy.

In conclusion, rule-based category learning was harmed when individuals were required to perform a distracting secondary task. In contrast, information-integration category learning was harmed when individuals performed a task that prompted attention to execution. These results set the stage to test whether different high-pressure conditions elicit effects analogous to the distracting and explicit monitoring secondary tasks.

In Experiment 2, we examine the effects of outcome pressure versus monitoring pressure on rule-based and information-integration category learning. If outcome pressure coopts attention and working memory, then we should find the same performance pattern as the distracting secondary task condition in Experiment 1. And to the extent that monitoring pressure increases attention to performance, category-learning results should parallel those seen in the explicit monitoring secondary task condition in Experiment 1.

## Experiment 2

Individuals learned rule-based and information-integration categories under a low-pressure baseline condition followed by either a low-pressure control condition or one of two high-pressure conditions. In the outcome-pressure condition, individuals were told that a 20% performance improvement from baseline would earn both themselves and a partner a monetary reward. As outlined in the introduction, we predicted that outcome-pressure would lead individuals to worry or ruminate about the consequences of their performance, distracting them from the task at hand. If so, rule-based category learning should suffer, but information-integration category learning should not. In the monitoring-pressure condition, participants were watched and videotaped during the category-learning task and told that the footage would be viewed by other students and researchers. We predicted that monitoring pressure would lead individuals to focus on what others are watching—the step-by-step task processes they are performing. If so, then information-integration category learning should suffer relative to baseline whereas rule-based category learning should not.

## Method

**Participants.** Undergraduate students ( $N = 130$ ) at the same university as in Experiment 1 participated in Experiment 2 ( $M$  age = 19.13 years,  $SD = 1.12$ ). Participants had no reported

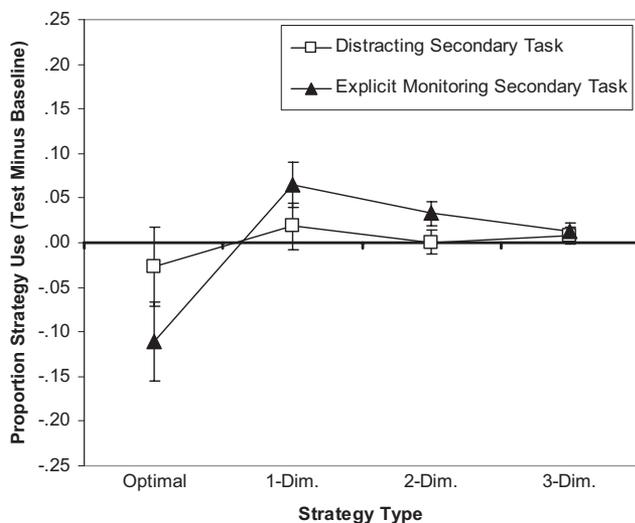


Figure 2. Differences in proportion of strategy use (test minus baseline) during information-integration category learning as a function of each strategy type (optimal, one-dimension rule, two-dimension rule, and three-dimension rule) and secondary task condition. Error bars represent standard errors.

colorblindness. Individuals were randomly assigned to a low-pressure control condition ( $n = 47$ ), an outcome-pressure condition ( $n = 43$ ), or a monitoring-pressure condition ( $n = 40$ ).

**Procedure.** The procedure was exactly the same as in Experiment 1, except the last two category structures (Block 2) were performed under either a low-pressure control condition or one of two types of pressure conditions, rather than with secondary tasks.

**Low-pressure control condition.** After Block 1 (single-task baseline) was completed, the experimenter returned to the testing room and explained to participants in the low-pressure control group that they would continue to perform a category-learning task and instructed them to try to do their best. The experimenter then left the room, and participants completed the last block of the category-learning task.

**Outcome-pressure condition.** Participants in the outcome-pressure condition were given a high-pressure scenario before continuing to Block 2. Specifically, the experimenter explained that the computer had been calculating a score based on the participant's categorization accuracy in the first block, and a 20% improvement in this score during the next sets of categories would earn the participant an additional \$10 at the end of the study. The experimenter explained, however, that the study was actually about teamwork, and both the participant and a "partner" must improve their scores to earn the money. The partner, participants were told, had already completed the experiment and improved his or her score, leaving it up to the present participant to do well for both individuals to be rewarded. Of course, if their categorization accuracy did not improve, they were told that neither the participant nor his or her partner would receive the bonus. After explaining the stakes, the experimenter left the room, and the participant completed Block 2. At the end of the experiment, participants were fully debriefed, including the fact that their partner was actually fictitious, and participants were given the money regardless of their performance.

**Monitoring-pressure condition.** Before beginning Block 2, the experimenter informed participants in the monitoring-pressure condition that their performance during the next sets of categories would be videotaped, for students and professors at the university to watch how people perform this skill. In addition, participants were told that the footage may also be used in a film about the basic skills of category learning funded for nationwide distribution to researchers and psychology classes. The experimenter set up the camera about 1 m to the left of the participant, so that the participant and the computer screen were in view, and stayed behind the camera watching individuals' performance during the category learning task. When Block 2 was completed, the experimenter turned off the camera and faced it away from the participant. After the experiment was completed, participants were fully debriefed concerning the purpose of the study and were reassured that, in fact, no one would be watching the tapes of their performance.

**Questionnaires.** As in Experiment 1, all individuals completed a question regarding how important it was to them to perform at a high level on the last two sets of category-learning tasks. Next, individuals rated how much pressure they felt to perform at a high level, ranging from 1 (*very little performance pressure*) to 7 (*extreme performance pressure*). Last, participants completed a brief demographics questionnaire.

## Results and Discussion

**Pressure ratings.** Ratings of performance pressure were significantly higher for individuals in the outcome-pressure ( $M = 4.95$ ,  $SE = .20$ ) and monitoring-pressure ( $M = 5.15$ ,  $SE = .15$ ) conditions, compared with the low-pressure control group ( $M = 4.29$ ,  $SE = .20$ ),  $t(89) = 2.31$ ,  $p = .02$ , and  $t(86) = 3.27$ ,  $p < .01$ , respectively. The outcome-pressure and monitoring-pressure groups did not differ in their pressure reports,  $t(81) = .76$ . Thus, both pressure conditions served to elevate feelings of performance pressure, to comparable levels, above that of the low-pressure control group.

**Categorization performance.** The dependent measure was again the number of trials taken to learn the categories to the criterion of eight correct trials in a row, log transformed. As in Experiment 1, a difference score was obtained by subtracting Block 1 (single-task baseline) trials to criterion from Block 2 (high-pressure or control block) trials to criterion. This computation was done separately for the rule-based and information-integration category structures to examine the relative impact of the pressure conditions on each type of category structure. Again, higher scores indicate worse category-learning performance (i.e., taking more trials to learn to criterion) during Block 2 relative to the single-task baseline.

This difference score was examined as a function of category structure and pressure condition. A 2 (Category Structure: rule-based, information-integration)  $\times$  3 (Pressure Condition: low-pressure control, outcome-pressure, monitoring-pressure) mixed ANOVA revealed no main effects of category structure ( $F < 1$ ) or pressure condition ( $F < 1$ ). However, a significant Category Structure  $\times$  Pressure Condition interaction was found,  $F(2, 127) = 4.75$ ,  $p = .01$ ,  $MSE = .21$ .

As shown in Figure 3, the low-pressure control condition did not alter either rule-based ( $M = .04$ ,  $SE = .05$ ),  $t(46) = 1.14$ , *ns*, or information-integration ( $M = .03$ ,  $SE = .08$ ),  $t(27) = .51$ , *ns*, category learning—the changes from baseline (i.e., the difference scores) were essentially zero.

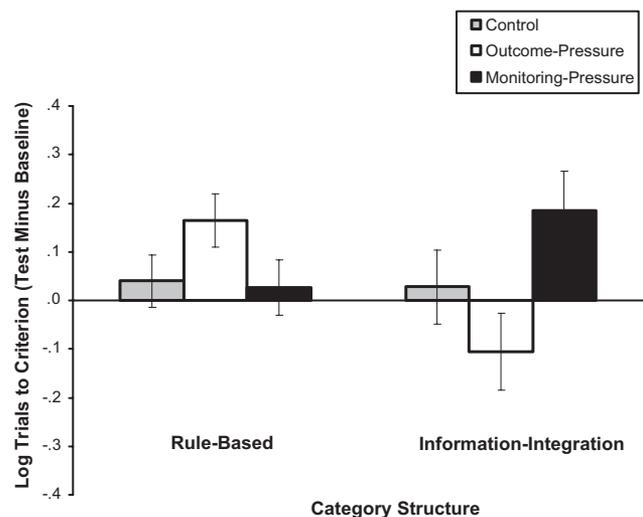


Figure 3. Mean trials to criterion (log transformed) difference score (test minus baseline) as a function of category structure and pressure type. Error bars represent standard errors.

Under high-pressure conditions, rule-based category learning was hurt by outcome pressure ( $M = .16$ ,  $SE = .06$ ),  $t(42) = 2.93$ ,  $p < .01$ , but not monitoring pressure ( $M = .03$ ,  $SE = .06$ ),  $t(39) = .40$ . Outcome pressure was created by informing individuals that a monetary reward, for both themselves and another person, was contingent on their performance. Given that rule-based category learning only suffered in this pressure condition, and only via a distracting secondary task in Experiment 1, these findings suggest that attention was diverted away from performance during the outcome-pressure condition. This result is consistent with the distraction theory of choking under pressure.

The opposite pattern was found for information-integration category learning, which was unaffected by outcome pressure ( $M = -.11$ ,  $SE = .08$ ),  $t(42) = -1.20$ , *ns*, but harmed by monitoring pressure ( $M = .19$ ,  $SE = .08$ ),  $t(39) = 2.51$ ,  $p < .02$ . In the monitoring-pressure condition, individuals were watched by an experimenter who videotaped their performance, nominally so that others could watch how people learn new categories. Information-integration category learning was only hurt by this type of pressure, and the findings were analogous with the explicit monitoring secondary task condition in Experiment 1. These findings are consistent with the explicit monitoring theory of choking under pressure. It appears that, under the watchful eye of others, individuals focus more explicitly on the steps of the skill being performed. This skill monitoring disrupts performance that operates best outside of explicit attentional control.

**Categorization strategies.** As in Experiment 1, we also examined the relative use of explicit rules during information-integration category learning between Block 1 (low-pressure baseline) and Block 2 (high-pressure or low-pressure control) by computing difference scores (Block 2 minus Block 1) for the proportion of optimal, and one-, two-, and three-dimension strategies (see Appendix B). As shown in Figure 4, participants used

more two-dimension ( $M = .04$ ,  $SE = .02$ ),  $t(39) = 2.14$ ,  $p < .04$ , and three-dimension ( $M = .03$ ,  $SE = .01$ ),  $t(39) = 2.45$ ,  $p < .02$ , strategies under monitoring pressure. In addition, the more individuals used these rule-based strategies, the more trials they took to reach the learning criterion ( $r_{2\text{-Dim.}} = .83$ ,  $p < .001$ ;  $r_{3\text{-Dim.}} = .69$ ,  $p < .001$ ). The optimal strategy was also used less often ( $M = -.11$ ,  $SE = .05$ ),  $t(39) = -2.25$ ,  $p = .03$ , and was associated with quicker learning ( $r = -.92$ ,  $p < .001$ ). Although somewhat higher, one-dimension rule use did not significantly increase under monitoring pressure ( $M = .05$ ,  $SE = .03$ ),  $t(39) = 1.69$ ,  $p = .10$ .

Outcome pressure had little impact on response strategies. One-dimension rules tended to be used less often ( $M = -.06$ ,  $SE = .03$ ),  $t(42) = -1.90$ ,  $p = .06$ , which, in conjunction with the unimpaired performance seen in this condition, aligns with the idea that performance on the information-integration task is better when individuals do not use these rules. No other strategies changed significantly under outcome pressure. Strategies remained unchanged in the low-pressure control condition as well. These results are consistent with the prediction that monitoring pressure increases explicit attention to the processes of category learning, leading to an increase in the use of explicit rule-based strategies. When people adopt explicit strategies to perform a task that operates best outside of explicit attentional control, performance suffers.

In conclusion, different types of tasks failed—and thrived—under different types of pressure situations. These findings underscore the idea that performance situations can be composed of various elements that impact working memory and attentional control in multiple ways. Revealing predictable relationships between pressure type and task type not only allows us to integrate disparate theories of skill failure but it also enables us to develop systematic ways to aid performance under pressure. We do this in Experiment 3.

### Experiment 3

In Experiment 3, we specifically focused our investigation on those pressure/task combinations shown to be most detrimental to performance in Experiment 2: Rule-based category learning under outcome pressure, and information-integration category learning under monitoring pressure. We added the distracting and explicit monitoring secondary tasks (from Experiment 1) to these pressure/task combinations as a means to demonstrate that knowledge of (a) how a particular type of pressure impacts performance and (b) the control structure of a task can be used to inoculate people against choking under pressure.

If both outcome pressure and a distracting secondary task serve to divert attention and working memory from performance, then rule-based category learning should be impaired when both of these conditions are combined. However, if an explicit monitoring secondary task that prompts people to attend closely to the steps of performance is added to rule-based categorization under outcome pressure, then performance decrements should be counteracted (at least in part).

On the other hand, if both monitoring pressure and an explicit monitoring secondary task act to enhance attention toward performance, then information-integration category learning should be harmed when these conditions occur simultaneously. However, a distracting secondary task that continually diverts attention away

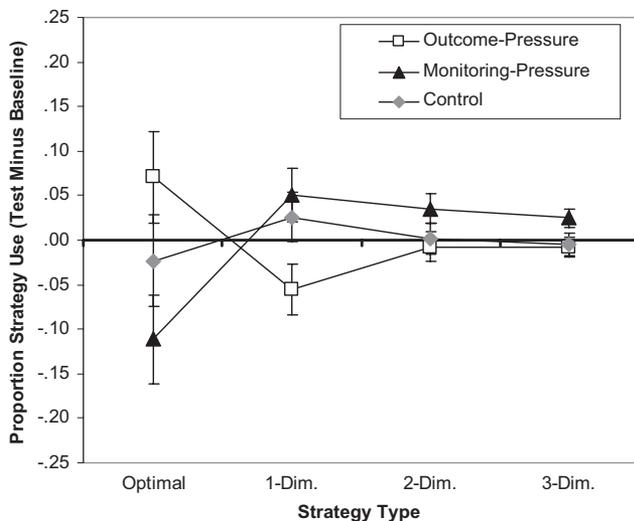


Figure 4. Differences in proportion of strategy use (test minus baseline) during information-integration category learning as a function of each strategy type (optimal, one-dimension rule, two-dimension rule, and three-dimension rule) and pressure condition. Error bars represent standard errors.

from task procedures may prevent monitoring pressure from taking such a toll. Demonstrating the validity of the above predictions would not only provide further support for multiple types of pressure but it would also provide useful clues for counteracting pressure's negative effects. Thus, we can demonstrate not only how certain triggers in the performance environment impact attention but also how we can redirect attention and reclaim performance.

## Method

**Participants.** Undergraduate students ( $N = 37$ ) at the same university as in Experiments 1 and 2 served as participants ( $M$  age = 18.97 years,  $SD = .73$ ). Participants had no reported colorblindness. Individuals were randomly assigned to either the outcome-pressure/rule-based category-learning condition ( $n = 15$ ) or the monitoring-pressure/information-integration category-learning condition ( $n = 22$ ).

**Procedure.** Participants were introduced to the same category-learning task as in Experiments 1 and 2, with a few exceptions. In the current experiment, each participant completed either three rule-based or three information-integration category structures. Recall that the stimuli consisted of four total dimensions: symbol color, symbol shape, background color, and number of embedded symbols. For each rule-based structure, the salient dimension for each set in Experiment 3 was randomly determined, without replacement, from these four possibilities. For each information-integration structure, the irrelevant dimension was randomly selected from the four total dimensions.

Participants first completed a low-pressure single-task baseline set, followed by two high-pressure sets. The two high-pressure manipulations were exactly the same as in Experiment 2. In the outcome-pressure condition, participants were informed by the experimenter that they could earn \$10 for themselves and a partner if they improved their categorization accuracy by 20%. Participants in the monitoring-pressure condition were monitored by the experimenter, who stood behind a video camera under the auspices that the film could be watched by students and professors at their university and across the country.

Each set performed under high pressure was also performed concurrently with one of the two secondary tasks used in Experiment 1: a distracting and an explicit monitoring secondary task, in counterbalanced order. As in Experiment 1, the distracting task was a letter-monitoring task, in which participants viewed a letter between every categorization trial and pressed the space bar on the computer keyboard if the letter was an "S." The explicit monitoring task was a confidence-rating task, in which individuals were asked to rate how confident they were that the category they were about to select was correct on a 7-point scale ranging from 1 (*not at all confident*) to 7 (*very confident*).

Following the categorization task, participants were given the same importance, pressure, and demographics questionnaires as in Experiment 2. Last, they were thanked and debriefed.

## Results and Discussion

**Pressure ratings.** As in Experiment 2, participants rated feelings of performance pressure as equivalent between the two

high-pressure groups: outcome pressure ( $M = 5.07$ ,  $SE = .41$ ), monitoring pressure ( $M = 4.96$ ,  $SE = .31$ ;  $F < 1$ ). These ratings were very similar to those reported by the pressure groups in Experiment 2.

**Categorization performance.** As in the previous two experiments, the number of trials taken to learn rule-based and information-integration categories to criterion (eight correct in a row) was measured for each set and log transformed, and difference scores were computed by subtracting the low-pressure set baseline score from each high-pressure set score. These scores were examined in a 2 (Category Structure/Pressure Condition: rule-based/outcome-pressure, information-integration/monitoring-pressure)  $\times$  2 (Secondary Task Condition: distracting, explicit monitoring) mixed ANOVA, with secondary task condition within-subjects. This analysis revealed no main effects of category structure/pressure condition ( $F < 1$ ) or secondary task condition ( $F < 1$ ). But a Category Structure/Pressure Condition  $\times$  Secondary Task interaction was found,  $F(1, 35) = 7.22$ ,  $p = .01$ ,  $MSE = .09$ .

As shown in Figure 5, rule-based categories were learned more slowly during a combination of outcome pressure and distracting secondary task conditions, relative to a low-pressure, single-task baseline—the difference score was significantly greater than zero ( $M = .22$ ,  $SE = .12$ ,  $t(14) = 2.35$ ,  $p < .04$ ). However, when rule-based category learning was performed with an explicit monitoring secondary task under outcome pressure, learning was no longer significantly impacted by pressure ( $M = .10$ ,  $SE = .11$ ,  $t(14) = 1.24$ , *ns*). Instructing individuals to think about how they were going to categorize each stimulus counteracted outcome pressure effects.

Information-integration category learning was slowed during simultaneous monitoring pressure and explicit monitoring secondary tasks relative to baseline performance ( $M = .24$ ,  $SE = .09$ ,  $t(21) = 2.20$ ,  $p < .04$ ). However, information-integration category learning was not harmed under monitoring pressure when learning was coupled with a distracting secondary task designed to redirect

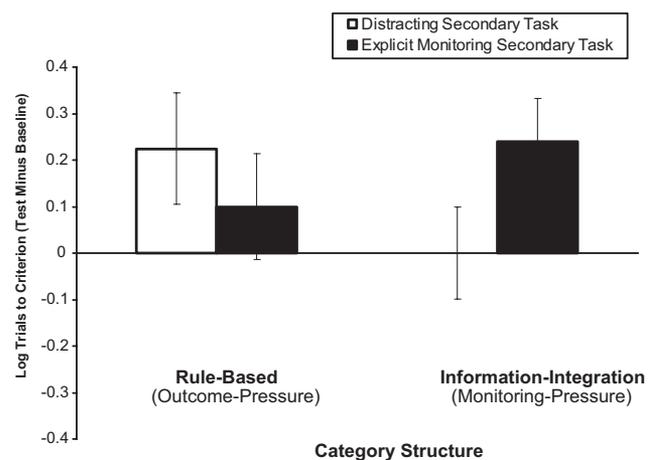


Figure 5. Mean trials to criterion (log transformed) difference score (test minus baseline) as a function of secondary task type and category structure/pressure type (rule-based tasks were performed under outcome pressure; information-integration tasks were performed under monitoring pressure). Error bars represent standard errors.

the focus of attention away from the online steps of performance ( $M = -.01$ ,  $SE = .10$ ),  $t(21) = -.09$ , *ns*).

**Categorization strategies.** Response strategies for information-integration category learning were consistent with these findings. As shown in Figure 6, in the explicit monitoring secondary task condition (under monitoring pressure), participants increased their use of explicit one-dimension rules ( $M = .08$ ,  $SE = .04$ ),  $t(21) = 2.39$ ,  $p < .03$ . Moreover, greater use of these rules was associated with worse category learning ( $r = .84$ ,  $p < .001$ ). Participants in this condition also used the optimal strategy less ( $M = -.14$ ,  $SE = .07$ ),  $t(21) = -2.06$ ,  $p = .05$ , and greater use of the optimal strategy was associated with better learning ( $r = -.91$ ,  $p < .001$ ). No other significant strategy changes were found for this or the distracting secondary task condition. It appears that combining monitoring pressure with a secondary task that increases attention to the task increases use of explicit, verbal strategies. However, adding a distracting secondary task helps prevent the increased use of these suboptimal strategies.

These findings enable us to dissociate further the attentional mechanisms at play in these secondary task and pressure types by demonstrating their opposing effects. Crossing situations that distract with those that enhance attention toward a skill seems to counteract the negative impact that would otherwise occur. These findings also offer promise for interventions aimed to alleviate performance pressure. When a performance situation tends to distract a performer from an attention-demanding task, perhaps a method to redirect attention back to the steps of performance will help (cf. DeCaro, Rotar, Kendra, & Beilock, 2010). On the other hand, when a performance situation leads individuals to focus explicitly on the component processes of a proceduralized skill, an intervention designed to mildly distract performers may prove beneficial.

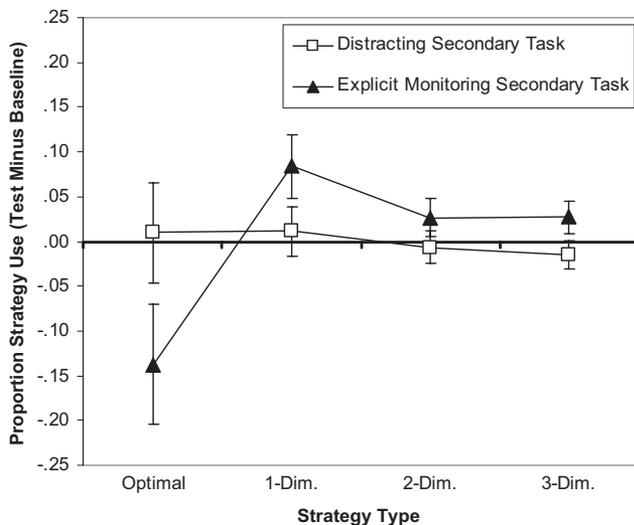


Figure 6. Differences in proportion of strategy use (test minus baseline) during information-integration category learning as a function of each strategy type (optimal, one-dimension rule, two-dimension rule, and three-dimension rule) and secondary task/monitoring-pressure condition. Error bars represent standard errors.

## Experiment 4

In Experiments 1–3, we established that skill failure depends on both aspects of the performance environment and the attentional demands of the task being performed. In Experiment 4, we conducted a final experiment to replicate and extend our findings to another skill domain. In particular, some people may be concerned about the comparability of the information-integration task used in Experiments 1–3 with sensorimotor skills that have formed the basis of much of the choking literature to date. Specifically, performance of the information-integration task relies on the procedural learning system (i.e., is largely procedural at the outset of learning this task; Ashby & Maddox, 2005), whereas sensorimotor skills typically become less reliant on explicit attentional control and more proceduralized with increasing practice (Anderson, 1982). Thus, in Experiment 4 we investigate whether monitoring pressure selectively disrupts a skill that has become largely proceduralized with practice, much like the well-learned skills that are often examined under performance pressure. Because of their similar reliance on the procedural system, we expect to see the same pattern of performance as with the information-integration category-learning task.

Participants were trained on the SRTT, in which they learned to press four keys on the computer keyboard in response to shifting probes on the computer screen. Unknown to participants, oftentimes these probes repeated a regular sequence (not unlike, for example, a simple sequence on a piano keyboard). Like other sensorimotor skills, the SRTT is believed to rely initially on both attention-demanding (declarative) and nondemanding (procedural) processes to learn the associations between subsequent key-presses (Robertson, 2007). However, over time (even within the course of a single learning session), the demand on attention decreases, and pressing the keys in sequence relies more on procedural processes, largely outside of explicit attentional control (Nissen & Bullemer, 1987).

Because the SRTT is a well-studied sensorimotor task that, like other sensorimotor skills, increases in reliance on the procedural system with increasing skill (Anderson, 1982), we expected performance to be affected in much the same way as the information-integration task used in Experiments 1–3. Specifically, after training on the SRTT, we examined performance in either a low-pressure control condition or one of the two high-pressure conditions used in Experiments 2 and 3 (i.e., outcome pressure and monitoring pressure). We expected worse SRTT performance under monitoring pressure relative to both outcome pressure and the low-pressure control condition.

## Method

**Participants.** Participants were right-handed undergraduate students ( $N = 65$ ) at the same university as in Experiments 1–3 ( $M$  age = 19.15 years,  $SD = .87$ ). Individuals were randomly assigned to either the outcome-pressure condition ( $n = 24$ ), monitoring-pressure condition ( $n = 21$ ), or low-pressure control condition ( $n = 20$ ).

### Procedure.

**SRTT training.** Participants were asked to complete the SRTT by placing their dominant right hand on four horizontally adjacent marked keys on the computer keyboard. Participants were

told that they should press one of the four keys every time they see a square appear on the screen and that the location of the square on the screen would indicate which key to press. Individuals were told to press the keys as quickly as possible and that the square would remain on the screen until the correct key was pressed. Following instructions, participants completed two 50-trial practice blocks in random order, to acquaint themselves with the key-pressing procedure. Next, participants completed eight training blocks, during which a 12-item second-order conditional sequence (1-4-3-1-2-4-2-3-2-1-3-4) was repeated 12 times (eight times in the introductory block). Before and after each training block were 24 random trials (48 in the introductory block). Thus, during training, participants completed a total of 1,536 trials (1,104 sequence trials and 432 random trials). As is commonly done with the SRTT, participants were never told that they were learning a sequence, to support the procedural nature of this skill.

**SRTT test.** Following SRTT training, all participants completed a test block including eight repetitions of the sequence, preceded and followed by 48 random trials (for a total of 192 trials). The test block was completed under either a low-pressure control condition or one of two high-pressure conditions, all of which were nearly identical to those in Experiment 2. Participants in the *low-pressure control condition* were informed that they would be completing the same task they had been doing and asked to try to do their best. Participants in the *outcome-pressure condition* were informed that they needed to improve both their speed and accuracy by 20% in order to earn \$10 for themselves and a partner. Participants in the *monitoring-pressure condition* were videotaped and told that the footage may be watched by students and professors interested in research on basic skill learning.

**Data analysis.** All incorrect responses were removed from analyses, and five participants were excluded from analyses because they had less than 80% accuracy during training (one in the control condition, two in the outcome-pressure condition, and two in the monitoring-pressure condition). Any reaction time (RT) greater than three standard deviations from a participant's median for the trial type (sequence or random) was also removed. Because we were interested in comparing differences in performance across conditions on the test block, we computed a measure of relative skill (rSRT; Galea, Albert, Ditye, & Miall, 2010): the difference between medians for the random and sequence trials, divided by the median of the random trials ( $[\text{random} - \text{sequence}] / \text{random}$ ). This skill score (rSRT) accounts for individual differences in time to respond to random stimuli and can range from 0 (indicating that random and sequence trials were performed at the same speed; i.e., no learning of the sequence) to 1 (indicating the highest possible amount of sequence learning, which is approached as time to respond to the sequence approaches zero). To ensure that carryover effects from the preceding trial type did not influence results (e.g., taking inordinately longer to respond to random trials immediately following a series of sequence trials could inflate the differences between these trial types), only the second half of the trials from the random and sequence blocks were included in this skill score (Brown & Robertson, 2007).

**Questionnaires.** Following the SRTT, participants were given the same importance and pressure questionnaires as in Experiments 2–3. They also completed the State version of the State-Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, & Lushene, 1970) as an additional measure of felt anxiety during the

test block. Finally, they completed a brief demographics questionnaire and were thanked and debriefed.

## Results and Discussion

**Perceptions of pressure.** Reports of state anxiety (STAI) differed between low- and high-pressure conditions. Compared with the low-pressure control condition ( $M = 36.56, SE = 2.55$ ), participants reported feeling more anxiety in the outcome-pressure condition ( $M = 41.71, SE = 2.36, 95\%$  confidence interval [CI] = 36.99–46.44,  $d = 0.45$ ) and the monitoring-pressure condition ( $M = 41.83, SE = 2.55, CI = 38.73\text{--}46.94, d = 0.56$ ), which did not differ. Although they were in the expected direction, ratings of performance pressure did not significantly differ between the high-pressure conditions (outcome-pressure  $M = 4.96, SE = .28$ ; monitoring-pressure  $M = 5.26, SE = .27$ ) and the low-pressure control condition ( $M = 4.68, SE = .27, CI = 4.09\text{--}5.28$ ).

**SRTT performance.** We first verified that participants learned the sequence by examining their median RT on test-block sequence trials compared with random trials. As expected, participants overall were faster to respond to sequence trials during the test block ( $M = 418.40, SE = 7.43$ ) than to random trials ( $M = 458.22, SE = 7.45$ ),  $t(59) = -9.08, p < .001$ , indicating that they were indeed skilled at this task.

We next compared performance on the test block across the three conditions, using the skill score described above (rSRT). Overall, condition had a marginal effect on skill,  $F(2, 59) = 2.98, p < .06$ . As shown in Figure 7, planned contrasts revealed that participants in the monitoring-pressure condition ( $M = .05, SE = .02$ ) performed worse on the test block than participants in the control condition ( $M = .10, SE = .02$ ),  $t(57) = -2.00, p = .05$ , and the outcome-pressure condition ( $M = .10, SE = .02$ ),  $t(57) = -2.24, p < .03$ . The latter two conditions did not differ,  $t(57) = .17, ns$ . Again, higher skill scores as measured by the rSRT indicate better SRTT performance during the test block. Thus, we see that different types of pressure had selective effects on the trained SRTT—a sensorimotor skill that operates largely outside of conscious awareness (Nissen & Bullemeyer, 1987). The pres-

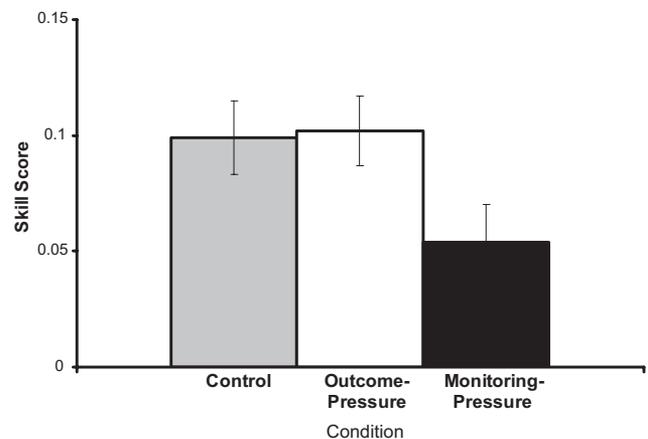


Figure 7. Serial Reaction Time Task skill score as a function of pressure type. Higher skill scores represent better task performance during the test block. Error bars represent standard errors.

sure associated with being monitored by others significantly harmed performance relative to a low-pressure control condition, whereas pressure associated with a monetary outcome for oneself and another person did not affect performance.

These findings are consistent with those of information-integration category learning in Experiments 2–3, a task also believed to rely on procedural control processes. Thus, the selective effects of monitoring pressure appear to extend beyond the category learning domain to performance of a complex sensorimotor skill that is thought to become proceduralized with practice. Again, we see that, by understanding the processes supporting different skills, we can not only better understand how these skills might fail but also in what circumstances.

### General Discussion

Situation-induced performance decrements occur across a variety of skill domains under a diverse set of situations. However, not everyone fails under what might typically be considered a high-pressure situation. In the current work, we examined when and for what reasons people fail versus succeed under pressure. By examining tasks that differ in their reliance on working memory and attentional control under two different types of high-pressure situations, we found evidence that pressure can impair performance in multiple ways.

Pressure to attain a particular performance-based outcome harmed a skill that relies on working memory and attention (i.e., rule-based category learning) but not skills less dependent on executive control (i.e., information-integration category learning and skilled SRTT performance). These findings suggest that outcome pressure coopts attention and working memory resources, perhaps with distracting worries or ruminations about performance. Thus, the distraction mechanism of choking under pressure may be most likely to occur when attention-demanding tasks are performed under pressure that emphasizes an important score, grade, or monetary reward.

Pressure involving performance monitoring (e.g., by another person and a video camera) disrupted information-integration category learning and SRTT performance but not rule-based category learning. Online evaluation of performance by others may lead individuals to attend to step-by-step skill execution because they expect others are doing so as well. Thus, explicit monitoring theories of choking under pressure are also supported by our findings. Under the pressure of being watched and evaluated by others, skills that operate best outside of close attentional control (e.g., procedural skills) suffer from a performer's enhanced attention to performance processes.

Given these findings, *both* distraction and explicit monitoring theories of choking under pressure seem to be correct. Whether attention is diverted from and/or enhanced toward the task at hand depends in large part on characteristics of the performance situation one is facing. Moreover, whether performance fails because of this situation depends also on the attentional demands of the task being performed. Our findings offer an important step toward reconciling these seemingly disparate theories of skill failure. With the multiple mechanisms of performance failure more clearly delineated, we can not only better understand when and how choking under pressure may occur, but we are also in a better

position to design interventions to alleviate the negative effects of performance pressure (Baumeister & Showers, 1986).

In Experiment 3 we revealed one way to mitigate the negative impact of pressure, by setting up secondary tasks that counteract pressure's harmful impact. We asked individuals under outcome pressure to complete rule-based category learning concurrently with a secondary task involving explicit monitoring. Whereas outcome pressure seems to divert attention and working memory away from performance, frequently prompting people to think explicitly about the processes they are using to perform the skill inoculated them against the ill effects of stress (cf. DeCaro et al., 2010). The opposite intervention helped in a monitoring-pressure situation. Mildly distracting people during performance made information-integration category learning immune to pressure's negative effects. This type of distraction may have kept individuals from overthinking a skill that operates best with little explicit control.

### Classifying Performance Situations

The current work demonstrates that the performance environment can influence how a skill is performed. Moreover, by better understanding when distraction and/or explicit monitoring may come about under stress, we are in a better position to link performance under pressure to other high-stress situations where performance may go awry.

**Multifaceted high-pressure situations.** The two high-pressure situations studied in the current work were designed to isolate pressures related to concerns over performance outcomes versus performance process. However, high-pressure situations are often composed of both of these elements, such as when an athlete vies for an important title (and possible monetary award) in front of a large audience. Under these type of combined pressure situations, participants perform more poorly on attention-demanding math skills (Beilock & Carr, 2005; Beilock & DeCaro, 2007; Beilock, Kulp, et al., 2004) and perform worse at proceduralized golf putting and baseball batting skills (Beilock & Carr, 2001; R. Gray, 2004). One might wonder, however, how this can be the case if these two types of pressure exert seemingly opposite effects.

One possibility is that both types of pressure actually impact performance at the same time. For example, working memory capacity could be reduced in these situations, but what attention is still available may be explicitly devoted to step-by-step control. Thus, attention-demanding skills (e.g., difficult math problem solving) would suffer from the reduction in working memory capacity, whereas proceduralized skills (e.g., well-learned sensorimotor skills) would be impaired by the increase in explicit monitoring.

An alternative possibility is that the impact of one type of stress may actually serve to lessen the impact of the other. Much like the results of Experiment 3, in which a secondary task that impacts attention in the opposite way as pressure reduces pressure's impact (e.g., an explicit monitoring secondary task performed under outcome pressure), people could become less distracted under outcome pressure if they are also facing the pressure of being explicitly monitored (and vice versa).

A third possibility is that people are most affected by the aspect of the pressure situation that is most salient to them. If an individual is most concerned about earning high marks, for example,

then he or she may suffer more from distraction than explicit monitoring, even if someone is watching. Although the current work has dissociated aspects of the pressure situation that differentially impact attentional control, whether and how these attentional mechanisms interact in multifaceted pressure situations remains an important empirical question.

**Stereotype threat.** As with the high-pressure situations investigated in the current work, other situations in which a personally important ability in a domain is questioned, such as conditions of stereotype threat, also often lead to performance decrements (Steele, 1997). Introducing a negative stereotype about one's social group impairs performance across a variety of domains, including proceduralized skills (e.g., Beilock, Jellison, Rydell, McConnell, & Carr, 2006; Chalabaev, Sarrazin, Stone, & Cury, 2008) and attention-demanding tasks (e.g., Beilock, Rydell, & McConnell, 2007; Schmader & Johns, 2003). There is evidence for both distraction and explicit monitoring mechanisms of skill failure under stereotype threat, yet the literature is inconclusive regarding when each impacts performance (see Schmader, Johns, & Forbes, 2008; Steele, Spencer, & Aronson, 2002, for a review).

Drawing from the current findings, one might speculate that aspects of the stereotype-threat situation may influence whether attention is diverted from and/or enhanced toward skill performance. Stereotype threat is typically brought about in the laboratory by highlighting the diagnosticity of a test regarding one's aptitude in a domain or intelligence in general (Steele et al., 2002). On some occasions, performance is also measured and watched by another person (e.g., Beilock et al., 2006; Chalabaev et al., 2008). The emphasis on a performance outcome may lead attention and working memory to be disrupted during performance, whereas explicit monitoring may play a bigger role when a stereotype is made salient and then monitored by others. Thus, as in the current work, the mechanisms of skill failure in stereotype-threat situations may also depend on aspects of the performance situation itself, in addition to the type of task being performed. Viewing stereotype threat in terms of the situational impact on working memory and attentional control may offer new insight into how (and when) stereotype threat will impact performance and shed light on new interventions for counteracting its impact (e.g., setting up secondary tasks to refocus attention optimally).

**Test anxiety.** Related work on test anxiety has suggested that, like distraction theories of choking under pressure, performance anxiety interferes with working memory processes critical for successful test performance (e.g., Ashcraft, 2002; Ashcraft & Kirk, 2001; Eysenck & Calvo, 1992; Hayes, Hirsch, & Mathews, 2008; Rapee, 1993; Wine, 1971). Test-anxious individuals have more worries and intrusive thoughts, particularly in situations with important contingencies, such as reward, test, or ego-involving situations (e.g., "This is an intelligence test"; Eysenck & Calvo, 1992, p. 421; Ikeda, Iwanaga, & Seiwa, 1996). The precise nature of these distractions is typically thought to be of worry over evaluation and concern over the level of performance relative to that required (Eysenck & Calvo, 1992; Sarason, 1972). Such self-imposed pressure is much like the outcome pressure in the current work—test-anxious individuals become distracted by the potential outcomes of performance and are most negatively impacted on tests relying heavily on attention and working memory (e.g., inductive reasoning; Calvo, 1985). This body of literature can thus be easily connected with the present work, adding the idea that

certain individual differences (i.e., test anxiety) may lead some people to be more sensitive to performance-based outcomes than others.

**Regulatory focus.** Other work has linked the pressure and regulatory focus literatures, suggesting that individuals in high-pressure situations become sensitive to the potential for losses in the environment (i.e., prevention focus), as opposed to being in low-pressure situations where people are simply compensated for participating in the research study (i.e., promotion focus). Worthy, Markman, and Maddox (2009) posited that this regulatory focus would interact with the rewards structure of the task, where a gains rewards structure entails accumulating points for correct responses, and a losses rewards structure involves losing fewer points for correct compared to incorrect responses.

Worthy et al. (2009) found that rule-based category learning was best in regulatory fit conditions (i.e., high-pressure with a losses reward structure, low-pressure with a gains reward structure) and worse in regulatory mismatch conditions (i.e., high-pressure with a gains rewards structure, low-pressure with a losses reward structure). Information-integration category learning showed the opposite pattern (i.e., worse performance in fit than mismatch conditions). Worthy et al. suggested that regulatory fit may lead to a sense of "feeling right," which increases confidence in performance, whereas regulatory mismatch decreases such confidence, potentially leading to anxiety that reduces working memory resources.

Worthy et al.'s (2009) findings coincide with some aspects of the current work, namely by showing that pressure can lead to failure or success with different types of tasks because of the availability of attentional resources during performance. Specifically, using Worthy et al.'s framework, one might presume that our outcome-pressure condition (the same pressure manipulation as Worthy et al.) induced a prevention focus, interacting with our "gains" reward structure to invoke a regulatory mismatch. This regulatory mismatch coopts working memory and attention, consistent with the distraction theory of choking under pressure.

However, our monitoring-pressure condition produced the opposite effect, possibly because it led to a promotion focus instead. According to Worthy et al. (2009), this promotion focus, coupled with the gains reward structure of our task, should not impair rule-based categorization but, instead, harm information-integration performance, which is exactly what we found. We admit that the above classification seems a bit arbitrary, as someone given a monitoring-pressure condition could just as easily wish to avoid a negative evaluation as they could desire to earn high regard by those watching. Similarly, outcome pressure could prompt individuals to either attend to the potential monetary and social gains or to the potential losses. It is notable that the stereotype threat literature is also inconsistent regarding whether stereotype threat leads to promotion or prevention focus (e.g., Chalabaev et al., 2008; Grimm, Markman, Maddox, & Baldwin, 2009; Keller & Dauenheimer, 2003; Seibt & Förster, 2004). Nonetheless, if prevention focus is more likely to come about in situations that highlight a performance-contingent outcome, whereas monitoring pressure is more likely to elicit promotion focus, then the findings of the current experiments fit nicely with the work on regulatory fit under stress. Our work also extends this previous research by demonstrating that explicit monitoring may occur in regulatory fit conditions, whereas distraction may occur in regulatory mismatch

situations, and these differences in attentional control may, in turn, be what drives performance success versus failure. Future research addressing these issues is needed.

## Conclusion

The current work demonstrates that attention can be diverted or enhanced because of such factors as whether a personally important performance-based incentive is at stake or whether performance is monitored by other people. Moreover, such changes in attentional control have different effects on skills that rely more or less on this important cognitive resource. By denoting certain aspects of the pressure environment that may lead individuals to focus on the process of performance versus the outcome of performance, we are not only in a position to predict when performance will fail or succeed but we can also provide interventions to help mitigate the possibility of failure. Thus, this work enables us to better understand performance failure—and ways to prevent it—across a variety of skill types and situations, from a student taking a math test to an expert on the playing field.

This work joins a recent body of literature seeking to understand the interplay between cognition, motivation, and emotion (e.g., Beilock, 2008; J. R. Gray, 2004; J. R. Gray, Braver, & Raichle, 2002; Grimm et al., 2009). The present findings align with, and potentially inform, research across several related areas (e.g., stereotype threat, test anxiety, regulatory focus). By focusing on the specific cognitive mechanisms by which pressure can exert its impact, we can begin to cut across such domains, working toward an overarching theory of when performance will fail versus succeed under stressful situations.

## References

- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89, 369–406. doi:10.1037/0033-295X.89.4.369
- Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, 56, 149–178. doi:10.1146/annurev.psych.56.091103.070217
- Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, 11, 181–185. doi:10.1111/1467-8721.00196
- Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, 130, 224–237. doi:10.1037/0096-3445.130.2.224
- Baddeley, A., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 28–61). New York, NY: University Press.
- 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. doi:10.1037/0022-3514.46.3.610
- Baumeister, R. F., & Showers, C. J. (1986). A review of paradoxical performance effects: Choking under pressure in sports and mental tests. *European Journal of Social Psychology*, 16, 361–383. doi:10.1002/ejsp.2420160405
- Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science*, 17, 339–343. doi:10.1111/j.1467-8721.2008.00602.x
- Beilock, S. L., Bertenthal, B. I., McCoy, A. M., & Carr, T. H. (2004). Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychonomic Bulletin & Review*, 11, 373–379. doi:10.3758/BF03196585
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, 130, 701–725. doi:10.1037/0096-3445.130.4.701
- Beilock, S. L., & Carr, T. H. (2005). When high-powered people fail: Working memory and “choking under pressure” in math. *Psychological Science*, 16, 101–105. doi:10.1111/j.0956-7976.2005.00789.x
- Beilock, S. L., Carr, T. H., MacMahon, C., & Starkes, J. L. (2002). When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *Journal of Experimental Psychology: Applied*, 8, 6–16. doi:10.1037/1076-898X.8.1.6
- Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 983–998. doi:10.1037/0278-7393.33.6.983
- Beilock, S. L., & Gray, R. (2007). Why do athletes “choke” under pressure? In G. Tenenbaum & R. C. Ecklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 425–444). Hoboken, NH: Wiley.
- Beilock, S. L., Jellison, W. A., Rydell, R. J., McConnell, A. R., & Carr, T. H. (2006). On the causal mechanisms of stereotype threat: Can skills that don’t rely heavily on working memory still be threatened? *Personality and Social Psychology Bulletin*, 32, 1059–1071. doi:10.1177/0146167206288489
- Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology: General*, 133, 584–600. doi:10.1037/0096-3445.133.4.584
- Beilock, S. L., Rydell, R. J., & McConnell, A. R. (2007). Stereotype threat and working memory: Mechanisms, alleviation, and spill over. *Journal of Experimental Psychology: General*, 136, 256–276. doi:10.1037/0096-3445.136.2.256
- Brown, R. M., & Robertson, E. M. (2007). Off-line processing: Reciprocal interactions between declarative and procedural memories. *The Journal of Neuroscience*, 27, 10468–10475. doi:10.1523/JNEUROSCI.2799-07.2007
- Calvo, M. G. (1985). Effort, aversive representations and performance in test anxiety. *Personality and Individual Differences*, 6, 563–571. doi:10.1016/0191-8869(85)90005-4
- Carver, C. S., & Scheier, M. F. (1978). Self-focusing effects of dispositional self-consciousness, mirror, presence, and audience presence. *Journal of Personality and Social Psychology*, 36, 324–332. doi:10.1037/0022-3514.36.3.324
- Chalabaev, A., Sarrazin, P., Stone, J., & Cury, F. (2008). Do achievement goals mediate stereotype threat? An investigation on females’ soccer performance. *Journal of Sport & Exercise Psychology*, 30, 143–158.
- Dandy, J., Brewer, N., & Tottman, R. (2001). Self-consciousness and performance decrements within a sporting context. *The Journal of Social Psychology*, 141, 150–152. doi:10.1080/00224540109600540
- Davis, D., & Brock, T. C. (1975). Use of first person pronouns as a function of increased objective self-awareness and performance feedback. *Journal of Experimental Social Psychology*, 11, 381–388. doi:10.1016/0022-1031(75)90017-7
- Davis, M. H., & Harvey, J. C. (1992). Declines in major league batting performance as a function of game pressure: A drive theory analysis. *Journal of Applied Social Psychology*, 22, 714–735. doi:10.1111/j.1559-1816.1992.tb00999.x
- DeCaro, M. S., Carlson, K. D., Thomas, R. D., & Beilock, S. L. (2009). When and how less is more: Reply to Tharp and Pickering. *Cognition*, 111, 397–403. doi:10.1016/j.cognition.2009.03.001
- DeCaro, M. S., Rotar, K. E., Kendra, M. S., & Beilock, S. L. (2010). Diagnosing and alleviating the impact of performance pressure on math-

- emational problem solving. *Quarterly Journal of Experimental Psychology*, 63, 1619–1630. doi:10.1080/17470210903474286
- DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition*, 107, 284–294. doi:10.1016/j.cognition.2007.07.001
- Dohmen, T. J. (2008). Do professionals choke under pressure? *Journal of Economic Behavior & Organization*, 65, 636–653. doi:10.1016/j.jebo.2005.12.004
- Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York, NY: Academic Press.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11, 19–23. doi:10.1111/1467-8721.00160
- Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, 6, 409–434. doi:10.1080/02699939208409696
- Forgas, J. P., Brennan, G., Howe, S., Kane, J. F., & Sweet, S. (1980). Audience effects on squash players' performance. *Journal of Social Psychology*, 111, 41–47. doi:10.1080/00224545.1980.9924271
- Galea, J. M., Albert, N. B., Ditye, T., & Miall, R. C. (2010). Theta-burst TMS to the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills. *Journal of Cognitive Neuroscience*, 22, 1158–1164. doi:10.1162/jocn.2009.21259
- Geller, V., & Shaver, P. (1976). Cognitive consequences of self-awareness. *Journal of Experimental Social Psychology*, 12, 99–108. doi:10.1016/0022-1031(76)90089-5
- Gimmig, D., Huguet, P., Caverni, J. P., & Cury, F. (2006). Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence. *Psychonomic Bulletin & Review*, 13, 1005–1010.
- Gray, J. R. (2004). Integration of emotion and cognitive control. *Current Directions in Psychological Science*, 13, 46–48.
- Gray, J. R., Braver, T. S., & Raichle, M. E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Sciences, USA*, 99, 4115–4120.
- Gray, R. (2004). Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied*, 10, 42–54. doi:10.1037/1076-898X.10.1.42
- Grimm, L. R., Markman, A. B., Maddox, W. T., & Baldwin, G. C. (2009). Stereotype threat reinterpreted as a regulatory mismatch. *Journal of Personality and Social Psychology*, 96, 288–304. doi:10.1037/a0013463
- Hardy, L., Mullen, R., & Jones, G. (1996). Knowledge and conscious control of motor actions under stress. *British Journal of Psychology*, 87, 621–636.
- Hayes, S., Hirsch, C., & Mathews, A. (2008). Restriction of working memory capacity during worry. *Journal of Abnormal Psychology*, 117, 712–717. doi:10.1037/a0012908
- Heaton, A. W., & Sigall, H. (1989). The “championship choke” revisited: The role of fear and acquiring a negative identity. *Journal of Applied Social Psychology*, 19, 1019–1033. doi:10.1111/j.1559-1816.1989.tb01236.x
- Ikeda, M., Iwanaga, M., & Seiwa, H. (1996). Test anxiety and working memory system. *Perceptual and Motor Skills*, 82, 1223–1231.
- Jackson, R. C., Ashford, K. J., & Norsworthy, G. (2006). Attentional focus, dispositional reinvestment, and skilled motor performance under pressure. *Journal of Sport & Exercise Psychology*, 28, 49–68.
- Keller, J., & Dauenheimer, D. (2003). Stereotype threat in the classroom: Dejection mediates the disrupting effect on women's math performance. *Personality and Social Psychology Bulletin*, 29, 371–381. doi:10.1177/0146167202250218
- Kimble, G. A., & Perlmutter, L. C. (1970). The problem of volition. *Psychological Review*, 77, 361–384. doi:10.1037/h0029782
- Langer, E. J., & Imber, L. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology*, 37, 2014–2024. doi:10.1037/0022-3514.37.11.2014
- Lewis, B., & Linder, D. (1997). Thinking about choking? Attentional processes and paradoxical performance. *Personality and Social Psychology Bulletin*, 23, 937–944. doi:10.1177/0146167297239003
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492–527. doi:10.1037/0033-295X.95.4.492
- Maddox, W. T., & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, 66, 309–332. doi:10.1016/j.beproc.2004.03.011
- Maddox, W. T., Love, B. C., Glass, B. D., & Filoteo, J. V. (2008). When more is less: Feedback effects in perceptual category learning. *Cognition*, 108, 578–589. doi:10.1016/j.cognition.2008.03.010
- Markman, A. B., Maddox, W. T., & Worthy, D. A. (2006). Choking and excelling under pressure. *Psychological Science*, 17, 944–948. doi:10.1111/j.1467-9280.2006.01809.x
- Masters, R. S. W. (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.
- Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York, NY: University Press.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32. doi:10.1016/0010-0285(87)90002-8
- Paulus, P. B., Shannon, V. C., Wilson, D. L., & Boone, T. D. (1972). The effects of operator presence on gymnastic performance in a field situation. *Psychonomic Science*, 29, 88–90.
- Rapee, R. M. (1993). The utilisation of working memory by worry. *Behaviour Research and Therapy*, 31, 617–620. doi:10.1016/0005-7967(93)90114-A
- Robertson, E. M. (2007). The serial reaction time task: Implicit motor skill learning? *Journal of Neuroscience*, 27, 10073–10075. doi:10.1523/JNEUROSCI.2747-07.2007
- Sarason, I. G. (1972). Test anxiety and the model who fails. *Journal of Personality and Social Psychology*, 22, 410–413. doi:10.1037/h0032893
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85, 440–452. doi:10.1037/0022-3514.85.3.440
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115, 336–356. doi:10.1037/0033-295X.115.2.336
- Seibt, B., & Förster, J. (2004). Stereotype threat and performance: How self-stereotypes influence processing by inducing regulatory foci. *Journal of Personality and Social Psychology*, 87, 38–56. doi:10.1037/0022-3514.87.1.38
- Spielberger, C. C., Gorsuch, R. L., & Lushene, R. (1970). *State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychology Press.
- Steele, C. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52, 613–629.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, 34, 379–440. doi:10.1016/S0065-2601(02)80009-0
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics*. New York, NY: HarperCollins.
- Tharp, I. J., & Pickering, A. D. (2009). A note on DeCaro, Thomas, and Beilock (2008): Further data demonstrate complexities in the assessment of information-integration category learning. *Cognition*, 111, 411–415. doi:10.1016/j.cognition.2008.10.003
- Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learn-

- ing systems. *Psychonomic Bulletin & Review*, 8, 168–176. doi:10.3758/BF03196154
- Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin*, 76, 92–104. doi:10.1037/h0031332
- Worthy, D. A., Markman, A. B., & Maddox, W. T. (2009). What is pressure? Evidence for social pressure as a type of regulatory focus. *Psychonomic Bulletin & Review*, 16, 344–349. doi:10.3758/PBR.16.2.344
- Wulf, G. (2007). Attentional focus and motor learning: A review of 10 years of research. *E-Journal Bewegung und Training*, 1, 4–14.
- Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition*, 34, 387–398.

## Appendix A

Table A1  
*Mean Trials to Criterion (Log-Transformed) in Experiments 1 and 2*

Experiment	Condition	Category structure	Block 1: Baseline <i>M (SD)</i>	Block 2: Manipulation <i>M (SD)</i>
1	Distracting secondary task	RB	1.23 (.32)	1.40 (.38)
		II	1.70 (.38)	1.76 (.35)
2	Explicit monitoring secondary task	RB	1.28 (.34)	1.25 (.31)
		II	1.54 (.35)	1.71 (.36)
	Control	RB	1.21 (.24)	1.24 (.24)
		II	1.65 (.35)	1.68 (.36)
Outcome pressure	RB	1.23 (.31)	1.39 (.33)	
	II	1.66 (.41)	1.55 (.37)	
Monitoring pressure	RB	1.30 (.37)	1.32 (.27)	
	II	1.59 (.36)	1.77 (.34)	

*Note.* RB = rule-based; II = information-integration.

Table A2  
*Mean Trials to Criterion (Log-Transformed) in Experiment 3*

Condition	Category structure	Baseline <i>M (SD)</i>	Distracting secondary task <i>M (SD)</i>	Explicit monitoring secondary task <i>M (SD)</i>
Outcome pressure	RB	1.22 (.19)	1.45 (.34)	1.32 (.32)
Monitoring pressure	II	1.58 (.38)	1.57 (.36)	1.82 (.34)

*Note.* RB = rule-based; II = information-integration.

(Appendices continue)

## Appendix B

### Response Strategies and Information-Integration Category Learning

We examined the extent to which participants' responses matched four different strategy types during information-integration category learning. Participants could use the "optimal" strategy, categorizing the stimuli as established by the experimenters (at 100% accuracy), most likely by using a procedural learning strategy. Or participants could use explicit, rule-based strategies to try to learn the information-integration categories. Such explicit strategies could involve either one dimension (e.g., items with a green symbol belong to Category A), two dimensions (e.g., items with a blue background or a square symbol belong to Category A), or three dimensions (e.g., items with a blue background color, or items with green squares, belong to Category A). One- and two-dimension strategies can result in 75% accuracy, and three-dimension strategies can lead to 87.5% accuracy. We found three possible strategies for each dimension, for a total of 10 possible strategies (including the optimal strategy; see also DeCaro, Carlson, Thomas, & Beilock, 2009). One could also postulate a number of very complex rule-plus-exception strategies that lead to 100% accuracy, but these cannot be dissociated from the optimal strategy. As we discuss in Experiments 1–3, the evidence suggests that the optimal strategy indeed reflects a less attention-demanding strategy.

We attempted to ascertain the strategies individuals used to learn the information-integration categories by examining the re-

sponses given on each learning trial and comparing these with the predicted responses from each possible strategy. To model these strategies, we divided the 200 possible trials into 20 blocks of 10 trials each and determined the strategy that matched the most responses given within each block. If two (or more) strategies accounted for the maximum number of responses within a given trial, then each was given an equal weight for that block (e.g., .5 for two strategies). If participants exited the task after reaching the learning criterion, the remaining blocks were scored as matching the optimal strategy. The number of response agreements across all blocks was summed for each strategy and converted to proportions for analysis.

To examine whether individuals relied more on explicit strategies in secondary task and/or high-pressure conditions than at baseline, we created a difference score. The proportion of each strategy type during the single-task and/or low-pressure baseline (depending on the experiment) was subtracted from the proportion of each strategy type during the secondary task and/or pressure condition. Results are presented in Figures 2, 4, and 6 and discussed in the text.

Received June 12, 2009

Revision received February 4, 2011

Accepted February 4, 2011 ■