2 The Benefits and Perils of Attentional Control

Marci S. DeCaro and Sian L. Beilock

Executive attention is involved in the learning and performance of an array of complex cognitive and motor skills, ranging from reading comprehension (Turner and Engle 1989) to mathematical problem solving (Beilock, Kulp, Holt, and Carr 2004) to learning a new sports skill (Beilock, Carr, MacMahon, and Starkes 2002). Although investigations of the link between executive attention and behavior have spanned diverse areas of psychological science, most of this work has yielded surprisingly similar conclusions regarding the role of this cognitive construct in high-level performance—the more attentional resources one is able to devote to performance at a given time, the higher one's success rate will be on the types of learning, problem solving, and comprehension tasks encountered in both the confines of the laboratory and the complexity of the real world (Engle 2002).

Executive attention allows memory representations to be maintained in a highly active state in the face of distraction (Conway et al. 2005) and is a key component of the working-memory system. By pairing domain-general executive attention resources with domain-specific (e.g., verbal and visual) short-term storage and processing resources, *working memory* functions to control, regulate, and actively maintain a limited amount of information with immediate relevance to the task at hand (Miyake and Shah 1999).

Working memory is thought to be "so central to human cognition that it is hard to find activities where it is not involved" (Ericsson and Delaney 1999, 259). In support of this idea, numerous studies have shown a positive relation between an individual's working-memory capacity and performance on an array of complex cognitive activities (Conway et al. 2005). And one's executive attention ability—the ability to attend to the most important information, while inhibiting irrelevant information—has been shown to drive this relation between individual differences in working memory and performance (Conway et al. 2005; Engle 2002; Kane, Bleckley, Conway, and Engle 2001; Kane and Engle 2000, 2003). For this reason, working-memory capacity is often conceptualized as executive attention (Engle 2002), and we do so in this chapter as well.

As mentioned above, working-memory capacity is positively related to higher level cognitive functions such as general intellectual ability, reasoning, and analytic skill and is touted as one of the most powerful predictive constructs in psychology (Conway et al. 2005). Despite its well-established utility, however, recent work suggests that increased attentional control can sometimes have a downside. In this chapter, we discuss research across a variety of tasks—problem solving, category learning, language learning, and correlation perception—to contrast the renowned benefits of attentional control with its potential pitfalls. In doing so, we demonstrate that *less* executive attention devoted to the planning and unfolding of performance is sometimes better than more.

Problem Solving

"A problem exists when a living organism has a goal but does not know how this goal is to be reached" (Duncker 1945, 2). Problem solving involves creating new knowledge in order to achieve a specific goal, not just extracting existing knowledge. As such, successful problem solving builds on other aspects of cognition, including perception, language, and working memory. When solving problems under normal conditions, individuals with higher working-memory capacity have an increased ability to maintain complex problem information in a transient store, while inhibiting ancillary information that might compete for attention. In contrast, individuals with less working-memory capacity are more apt to spread their attention superficially across multiple aspects of the performance environment rather than focusing intently on a subset of task information.

Support for the idea that individual differences in working memory capture variation in attentional control ability comes from an investigation of dichotic listening by Conway, Cowan, and Bunting (2001). These researchers asked individuals lower and higher in working memory to listen to a message in one ear and ignore a message in the other ear. In the irrelevant, to-be-ignored message, the participant's name was sometimes mentioned. Of interest was whether an individual noticed his or her name, despite being instructed to ignore the message in which his or her name was played. Conway et al. found that individuals lower in working-memory capacity were more likely to detect their name in the irrelevant message than were those higher in working memory.

This ability of higher working-memory individuals to selectively control attention, so that ancillary information is blocked out, is typically viewed as an aid to problem solving—facilitating a planned, deliberate memory search for problem solutions and supporting the online execution of a series of problem steps. In contrast, simultaneously attending to information both focal and disparate to the task at hand typically leads to suboptimal performance. However, this is not always the case. We begin by describing situations in which higher working memory is useful for problem solving and how performance suffers when this cognitive control capability is compromised. We then go on to demonstrate that performance on some types of problems actually benefits when one has less opportunity or less ability to exert attentional control.

In many problem-solving situations, the more working-memory capacity individuals bring to the table, the better they perform. As an example, Beilock and Carr (2005; see also Beilock and DeCaro 2007) asked individuals to complete a demanding mental arithmetic task called modular arithmetic and looked at their performance as a function of individual differences in working memory. Modular arithmetic involves judging the truth-value of equations such as "34 = 18 (mod 4)." Although there are several ways to solve modular arithmetic equations, Beilock and Carr taught their participants a problem-solving method that involves two key problem steps. First, the problem's middle number is subtracted from the first number (i.e., 34 – 18), and then this difference is divided by the last number (i.e., 16 ÷ 4). If the result is a whole number (here, 4), the statement is true. If not, the statement is false. As one can see, successful performance on this task requires the ability to allocate attentional resources to multiple problem steps and the ability to work with and manipulate this information in memory (e.g., holding 16 in mind while dividing it by 4).

Individual differences in working memory were measured using two common assessment tools: Operation Span (OSPAN; Turner and Engle 1989) and a modified Reading Span (RSPAN; Daneman and Carpenter 1980). In the OSPAN, individuals are asked to solve a series of arithmetic equations while remembering a list of unrelated words. Equation–word combinations are presented one at a time on the computer screen (e.g., "(3 × 4) – 2 = 8? CAT"), and individuals are asked to read the equation aloud and verify whether it is correct. Individuals then read the word aloud. At the end of a series of two to five of these strings, participants are asked to write down the series of words, in the correct order. The RSPAN follows the same general procedure, except instead of verifying equation accuracy and reading a word, individuals verify whether a sentence makes sense and then read a letter aloud for later recall (e.g., "On warm sunny afternoons, I like to walk in the park.? G"). Working-memory scores on these tasks consist of the total number of words/letters recalled from all series in which recall was 100% accurate. The ability to maintain this type of information (e.g., the words/letters) in the face of distraction (e.g., equation or sentence verification) is said to reflect executive attention, or working-memory capacity (Engle 2002).

What Beilock and Carr (2005) found was quite consistent with the idea that more working memory is better than less. The higher individuals' working memory, the more accurately they solved the modular arithmetic problems. Attention benefits performance on this type of multistep mental arithmetic task. Beilock and DeCaro (2007, experiment 1) have recently replicated this effect (see figure 2.1, top line) and also shed light on why these working-memory differences might occur. To do this, we

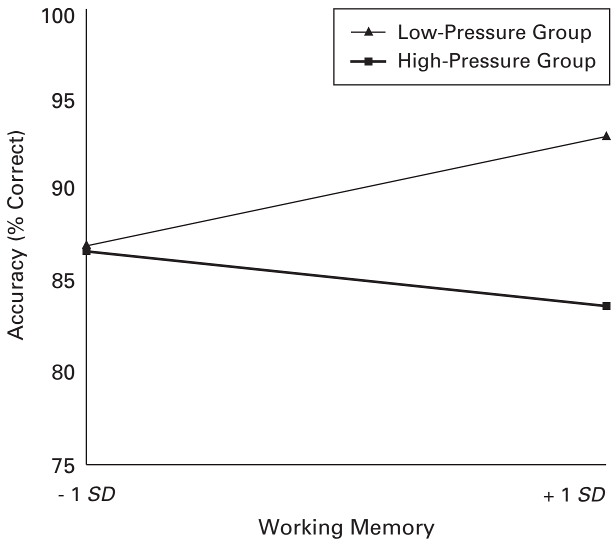

Figure 2.1
Mean modular arithmetic problem accuracy (percentage correct) as a function of individual differences in working memory and pressure condition. Nonstandardized coefficients are plotted at ±1 *SD*.
Adapted from Beilock and DeCaro 2007, experiment 1.

prompted individuals to describe the steps and processes they used to solve a selection of the modular arithmetic problems. Despite the fact that modular arithmetic is based on common subtraction and division procedures, there are shortcut strategies that can be employed to derive the correct answer, some of the time, without requiring a multistep problem-solving algorithm. For example, if one automatically responded to problems with all even numbers as "true," this strategy would result in a correct answer some of the time (as in the problem above), but not always (e.g., 52 = 16 (mod 8)). Successfully computing a multistep algorithm (i.e., subtract, then divide) would result in a correct answer every time.

We hypothesized that individuals with lower working-memory capacity, and therefore with less capacity to maintain and execute the complex procedures the algorithm required, would rely on shortcut strategies to circumvent this demand on attentional control (cf. Siegler 1988). On the other hand, individuals who can execute the algorithm with ease, those higher in working-memory capacity, would be more likely to do so in order to attain the highest accuracy possible. Consistent with this idea, we found that individuals lower in working-memory capacity were less likely to report using complex multistep strategies to solve the math problems than were their higher

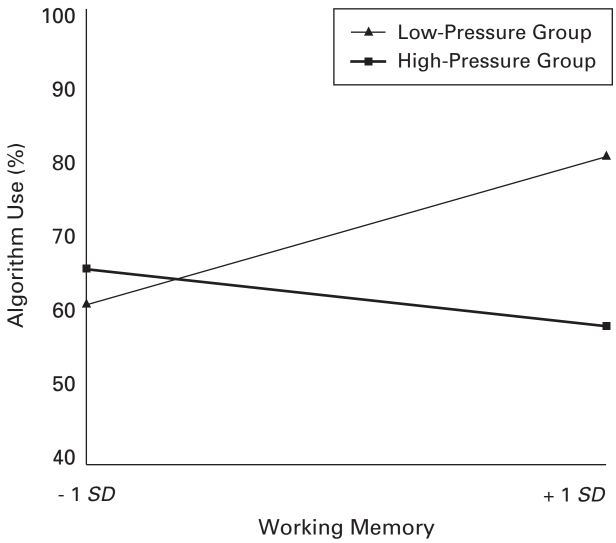

Figure 2.2
Proportion of rule-based algorithm use as a function of individual differences in working memory and pressure condition. Nonstandardized coefficients are plotted at ±1 *SD*.
Adapted from Beilock and DeCaro 2007, experiment 1.

capacity counterparts (see figure 2.2, top line). When these individuals were not using the complex strategies, they were using shortcuts. Use of shortcuts resulted in less accurate performance overall.

Given these findings, one might think that individuals higher in working memory should always outperform their low-capacity counterparts when solving difficult problems. What happens, however, if a particular performance situation compromises one's attentional resources? As an example, testing situations that elicit pressure to perform at a high level oftentimes lead to worries about performing poorly. These worries have been shown to consume attention and working-memory resources needed to successfully solve difficult math problems (Ashcraft and Kirk 2001; Beilock 2008; Beilock, Kulp, et al. 2004). One possibility is that all individuals, regardless of working-memory capacity, are equally impacted by pressure. Everyone's performance might drop by the same amount when the pressure is on. If so, then higher working-memory individuals will still outperform those with less capacity. A second possibility, however, is that because individuals higher in working-memory capacity rely heavily on this important resource for their typical success at demanding tasks like math, they might have more to lose in a pressured testing situation. That is, under pressure, individuals higher in working memory may perform as if they were lower in working memory in the first place, precisely because pressure-induced worries co-opt the very working-memory resources that higher capacity individuals normally use to showcase superior performance.

We have tested these ideas using the same modular arithmetic problems described above (Beilock and DeCaro 2007). After performing a set of practice problems during which individuals were merely instructed to perform as quickly and accurately as possible, participants were given a scenario intended to elicit commonly experienced pressures such as social evaluation, peer pressure, and a potential outcome-dependent reward. Specifically, individuals were told that if they could improve their problem-solving speed and accuracy by 20% relative to the first set of problems, they could earn a monetary reward. This reward, however, was said to be part of a "team effort," and both the participant and a "partner" needed to improve in order for both parties to receive the reward. The partner, however, was said to have already participated in the study and improved by the required amount, leaving the rewards for both participants dependent on the present individuals' performance. Individuals were also videotaped by an experimenter and informed that the footage would be examined by math teachers and students in order to examine how individuals learn this type of math skill. After hearing these stakes, participants completed the second set of math problems.

In line with the idea that our type of pressure situation compromises the attentional resources of those who typically rely on this capacity the most, individuals higher in working memory performed the modular arithmetic problems significantly worse under high-pressure compared to low-pressure tests. As shown in figure 2.1 (bottom line), under pressure the performance of higher working-memory individuals (right side of the graph) was at the same level as individuals lower in this capacity. The performance of those lower in working-memory capacity (left side of the graph) was not affected by pressure—their performance was equivalent in both high- and low-pressure testing environments.

Why might the performance of low working-memory individuals be so resilient to pressure's negative effects? And why might the performance of high-working-memory individuals fall under pressure? As mentioned previously, in normal situations individuals lower in working memory are less likely to solve the math problems with a complex algorithm. And when individuals were not using complex strategies, they used shortcuts that circumvent the heavy demand on attentional control. Under pressure, lower working-memory individuals were still able to use these shortcut strategies (see figure 2.2, bottom line), given that they are not attention-demanding in the first place. This simpler problem-solving approach allows individuals to maintain adequate, above-chance (but less-than-perfect) problem-solving accuracy (see figure 2.1). As shown in figure 2.2, higher working-memory individuals under high pressure also adopted the problem-solving shortcuts used by their lower capacity counterparts. Pressure limited high-working-memory individuals' ability to use the intensive problem-solving approach. When working memory was compromised by environmental demands, those who typically perform at the top (i.e., higher working-memory individuals) showed the largest performance decline (see also Kane et al. 2001; Kane and Engle 2002; Rosen and Engle 1997). Here again, we see the necessity of executive attention resources for problem solving—when these resources are taken away by environmental distractions, performance falters relative to where one was under normal, low-stakes conditions.

As we saw in Conway et al.'s (2001) dichotic listening study, where lower working-memory individuals were more likely to notice their name in the message they were supposed to be ignoring than their higher working-memory counterparts, instead of focusing intently on a subset of task information, individuals with lower working-memory capacity are more apt to spread their attention superficially across multiple aspects of the performance environment (Conway et al. 2001). For these individuals, learning and skill execution may be more associative in nature, less dependent on controlled effort, and rely more on shortcuts or heuristics. Of course, attending to information both focal and disparate to the task at hand typically leads to suboptimal performance, such as when performing modular arithmetic problems requiring attention to multiple task steps. However, a diffuse attentional focus may not always prove harmful. Having less ability to maintain complex information in the focus of attention may, in some situations, lead to more inventive problem-solving approaches than would be discovered if attention were more stringently controlled.

Beilock and DeCaro (2007, experiment 2) examined this idea by asking individuals to complete a series of water jug problems (Luchins 1942). In this task, three jugs are shown on a computer screen, each able to hold a different maximum capacity and labeled as jugs A, B, and C (see figure 2.3). Individuals must use the capacity of these three jugs to derive a goal quantity of water. A mathematical formula is used to denote a solution, and importantly, individuals are instructed to use the simplest strategy possible, without the aid of pencil and paper. Six problems were used in total. The first three can only be solved with a complex algorithm (i.e., B – A – 2C). These complex problems require multiple problem steps (e.g., computing different subtraction operations while also maintaining the results of prior calculations in transient memory) and therefore rely heavily on attentional resources. Each of the last three problems, however, can be solved in two different ways: with the same complex algorithm as the first three problems or with a much simpler formula (i.e., A – C or A + C). The latter solution is more optimal in this case, because it is the simplest solution in terms of the number of steps involved. Notably, the formula given as a problem solution is directly reflective of one's problem-solving strategy. Of interest is whether these problem-solving strategies vary as a function of working-memory capacity—specifically whether individuals continue to use the more complex problem solution or whether they switch to the simpler, shortcut strategy when it is available.

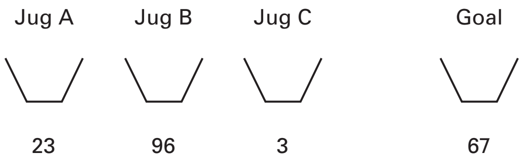

Figure 2.3
Water jug display. Participants derived a formula to obtain a "goal" quantity of water using jugs of various capacities. The first three problems were only solvable by the formula B – A – 2C (i.e., Fill jug B, pour out enough to fill jug A, then pour the remaining into jug C twice, leaving the goal quantity in jug B). The last three problems were solvable by this same difficult formula in addition to a much simpler formula (e.g., A – C). Individuals were informed that the water supply was unlimited and not all jugs had to be used.

We found that lower working-memory individuals were *more* likely to switch to the simpler solution when it became available. In contrast, individuals higher in working memory were more likely to persist in using the complex problem solution. Such persistence is known as *mental set* and, here, represents a negative artifact of previous experience in which individuals who are used to performing a task in a particular way tend to repeat this behavior in lieu of a more efficient strategy (Wiley 1998). Having a greater ability to execute multiple problem steps in memory seems to lead higher working-memory individuals to set in on a narrower problem-solving approach in line with their high capabilities. This is true even though, at the outset of the water jug task, we asked all subjects to solve the problems using the *simplest* strategy possible.

Such mental set effects can be especially pronounced when one is not only high in working memory but also has a lot of experience in a given domain. Ricks, Turley-Ames, and Wiley (2007) nicely demonstrated this phenomenon in the domain of baseball. They asked baseball experts and novices (as determined by a baseball knowledge test) to perform a creative problem-solving task called the Remote Associates Task (RAT; Mednick 1962). In this task, individuals view three words (e.g., "cadet, crawl, ship") and are asked to discover a fourth word (i.e., "space") that can be combined into a meaningful phrase with each of the three other words (i.e., "space cadet," "crawlspace," "spaceship"). The test words were either baseball neutral, having no obvious association with any aspect of baseball (as in the previous example), or baseball misleading. Baseball-misleading stimuli have one word that can be associated with baseball, but not in a way that would likely lead to a correct solution. For example, given the words "plate, broken, shot," a baseball expert might quickly retrieve the word "home" as associated with "plate," when the correct answer (i.e., glass) actually has no association with baseball at all.

To the extent that greater attentional control enables efficient retrieval and testing of multiple problem solutions, while inhibiting previously tested or ineffective solutions (Rosen and Engle 1997), one would expect higher working memory to be related to more successful performance on this problem-solving task. Indeed, for the neutral stimuli, the higher individuals' working memory, the better their solution accuracy (regardless of baseball expertise). A different pattern of results was seen for the baseball-misleading problems, however. First, expertise played a detrimental role. Baseball experts were outperformed by novices on the baseball-misleading problems. Experts have been shown to fixate on problem solutions that are activated by their extensive prior knowledge, leading to a negative mental set on this type of task (Wiley 1998). Moreover, the higher baseball experts' working memory, the worse they performed on the baseball-misleading problems. Working memory appears to have exacerbated the strategy rigidity commonly associated with expertise, by allowing hyperfocus on the incorrectly selected problem solution.

However it is triggered, whether from prior facility with a solution path or extensive knowledge of a particular domain, working memory supports a persistent approach in ways that are sometimes too selective. Such reliance on cognitive control not only may limit the discovery of new problem-solving approaches but may also lead to an attention-dependent learning strategy that overrides a more optimal associative strategy. We now turn to an example of the latter case in the category learning domain.

Category Learning

Similar to most problem-solving tasks, there are various ways one can go about learning the many categories that exist in our world. For example, individuals encountering new information, objects, or even people can explicitly test various hypotheses about the categories to which these belong. In order to learn to categorize objects in this way, individuals must form and test hypotheses about the potentially relevant features of the stimulus, move on to new hypotheses if current ones prove incorrect, and refrain from reexamining the hypotheses that have already been tested. This kind of complex process relies heavily on executive attention (Dougherty and Hunter 2003). However, there are other category learning strategies that are less attention-demanding, and in such cases, trying to devote executive attention resources to performance can actually result in a less-than-optimal learning situation.

When definitive rules can be applied to determine category membership, the best strategy is typically to hypothesize about the features that determine category membership. Tasks used to resemble this process in the lab are called *rule-based category learning* tasks (Ashby and Maddox 2005). Individuals usually see a series of categorization stimuli one at a time and are instructed to categorize each into category "A" or "B." Following each categorization choice, individuals usually receive feedback. The idea is that, over a series of categorization trials, individuals will learn to correctly categorize the stimuli to some criterion (e.g., eight correct categorization responses in a row; Waldron and Ashby 2001). A variety of categorization stimuli have been used for these tasks. For example, Waldron and Ashby (2001) created 16 stimuli, each a square with an embedded symbol in it. Each stimulus had four dimensions, with one of two levels of each dimension: square–background color (yellow or blue), embedded symbol shape (circle or square), symbol color (red or green), and number of embedded symbols (1 or 2). For a rule-based task, stimuli are correctly categorized based on an easily verbalizable rule regarding one of these features (e.g., "If the embedded symbol is a circle, choose category A; if the symbol is a square, choose category B"). The specific rule is established beforehand by the experimenter, and the individual discovers it over a series of learning trials.

Because generating and selecting different rules about category membership, while inhibiting previously selected features, relies extensively on working-memory resources (Ashby and O'Brien 2005), it is not surprising that individuals with more of this capacity outperform lower working-memory individuals on this type of rule-based learning task (see figure 2.4, left side; DeCaro, Thomas, and Beilock 2008). Moreover, when

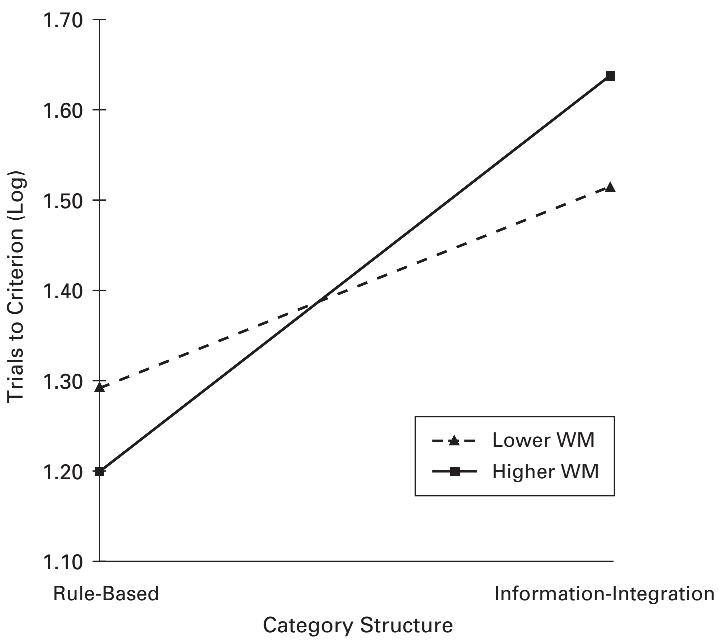

Figure 2.4
Mean number of trials taken to learn categories to a criterion of eight correct categorization responses in a row (log transformed), as a function of category structure and individual differences in working memory (WM). WM was measured as a continuous variable—nonstandardized regression coefficients are plotted at ±1 *SD*.
Adapted from DeCaro, Thomas, and Beilock 2008.

working-memory capacity is limited by a requirement to perform another demanding task simultaneously (Waldron and Ashby 2001; Zeithamova and Maddox 2006), or by a distracting high-pressure situation (Markman, Maddox, and Worthy 2006), the ability to learn rule-based categories is diminished.

Other categories are better learned without such reliance on attentional control. Indeed, when learning categories based on stimulus–response combinations too complex to occur within the bounds of explicit awareness, attentional control can simply get in the way. *Information-integration category learning* tasks are used to investigate this type of learning (Maddox and Ashby 2004). For example, the same 16 stimuli used in the rule-based task mentioned above can be grouped into similarity-based information-integration categories. To do so, one of the four stimulus dimensions is selected to be irrelevant, and each level of the remaining dimensions is randomly assigned a +1 or –1 value (e.g., a blue background could be assigned a +1 and a yellow background a –1). Then the dimension values for each stimulus are added together. If the sum of the three numbers is greater than one, that stimulus belongs to Category A; otherwise it belongs to Category B (Waldron and Ashby 2001). As can be seen, information-integration categories are not easily verbalized but instead rely on similarities between items and their respective categories that are associated over a series of learning events. This type of learning is believed to rely extensively on the procedural learning system (Maddox and Ashby 2004).

When new categories of any type are learned, it is thought that individuals employ both of the learning processes mentioned above—explicitly testing hypotheses about category membership while also accruing procedural-based associations between items and their respective categories. Whichever strategy accomplishes learning the fastest wins out (Zeithamova and Maddox 2006). As long as explicit hypothesis testing is occurring, however, this strategy will dominate responding. Therefore, in a rule-based task, individuals will typically successfully test different hypotheses about category membership until an explicit rule is discovered. However, in an information-integration task, individuals are actually slower to learn the categories the longer they persist in testing different rules—they are better off abandoning rule-based testing and responding only as guided by the procedural learning system (Markman et al. 2006; Zeithamova and Maddox 2006).

To the extent that individuals higher in working-memory capacity are better able to carry out complex hypothesis testing (Dougherty and Hunter 2003), they may be more likely to persist in using these multistep rules even when such a strategy is not ideal. In support of this idea, DeCaro, Thomas, and Beilock (2008) demonstrated that individuals higher in working-memory capacity took significantly longer to learn information-integration category structures than individuals lower in working memory (see figure 2.4). Similarly, Markman, Maddox, and Worthy (2006) found that in a high-pressure situation in which working-memory capacity is consumed, information-integration learning performance actually improved relative to a low-pressure testing condition. Distracting attentional resources away from category learning appears to have reduced the ability to hypothesis test, leading individuals to abandon this strategy sooner and allowing the procedural learning system to dominate categorization responses. Recently, Maddox and colleagues (2008) found that more detailed feedback after each categorization trial (i.e., the correct category assignment is displayed in addition to the minimal feedback labels "correct" or "incorrect") hurt information-integration category learning but helped rule-based learning. The additional information seems to have led individuals to rely on rule-based processing, to the detriment of a learning task that operates more optimally outside explicit attentional control.

Much like problem-solving tasks for which the optimal approach involves dissipating attention and allowing simpler strategies to become apparent, learning how to categorize new information or objects can sometimes be best accomplished by *not* thinking too much. Individual cognitive capabilities or situational factors that lead one to attend more explicitly to the factors determining category membership can serve to impair this type of learning.

Language Learning

Information-integration category learning is similar in nature to other tasks requiring the gradual accrual of environmental regularities, such as language learning or the perception of correlation. It is widely known that adults have more difficulty adeptly learning a new language than do children (Cochran, McDonald, and Parault 1999). One hypothesis of language learning (Newport 1990) posits that the limited cognitive resources of children may facilitate the learning of new language. In order to learn language, one must be able to correctly select from a stream of conversation not only words and their combinations but also the simple morphemes that change the meanings of words (e.g., adding an "s" to a word to denote plurality). Analyzing the errors made by adults versus children learning a second language, Newport discovered that adults are more likely to rely on unanalyzed wholes—words or phrases that often appear together in a particular context but may not always be appropriate in a new context. Children are more adept at the componential analysis that eventually results in better grasp of the language—they pick up the pieces of the complex linguistic input to which they have been exposed and flexibly learn to use them correctly.

Several studies have supported the idea that "less is more" when learning language (Newport 1990). Kersten and Earles (2001) found that adults learn a miniature artificial language better when initially presented with small segments of language rather than the full complexity of language, purportedly allowing them to process language as if their cognitive resources were more limited in the first place. Other work has found that language learning improves if an adult concurrently performs another task designed to consume working-memory resources (Cochran et al. 1999).

It should be noted that the exact role of executive attention in language learning has not yet been fully unpacked. Some use the term “working memory” (e.g., Kersten and Earles 2001), and others use terms like “maturational state” (Newport 1990). Moreover, Newport and others primarily describe the potential benefits of working-memory limitations in language learning in terms of the limited storage capacity to perceive and remember small segments of language, highlighting the short-term storage aspects of working memory more than the attentional control capabilities central to this construct. Yet, although the specific role of the executive attention component of working memory has not been central to this theory of language learning, this initial research does point to the potential negative impact of greater attentional control abilities and is consistent with research in similar domains such as information-integration category learning and, as will be seen below, correlation perception.

Correlation Perception

Research on the perception of correlation, or statistical regularities between two events, has also found an advantage of limited processing capacity. In one demonstration of this effect, Kareev, Lieberman, and Lev (1997) presented participants with a large bag containing 128 red and green envelopes and asked them to select one envelope at a time. Inside each envelope was a coin, marked with either an “X” or an “O.” When selecting each envelope, individuals were asked to predict which marking would appear on the coin, based on the color of the envelope. If the prediction was correct, participants earned the coin in the envelope. Counterintuitively, individuals performing worse on a digit span task, a measure of short-term memory, rated the correlations between envelope color and coin marking more accurately than those performing well at the memory task. Kareev and colleagues explained that individuals with less cognitive capacity are more likely to perceive narrow “windows” of events out of an expansive experience with co-occurring events—that is, lower capacity individuals will perceive and remember only a small chunk of these trials. Smaller subsets of trials are more likely to be highly skewed, and therefore lower capacity individuals will perceive correlations as more extreme, facilitating performance on this type of task (for a debate of these findings, see Anderson et al. 2005; Cahan and Mor 2007; Juslin and Olsson 2005; Kareev 2005).

Gaissmaier, Schooler, and Rieskamp (2006) replicated Kareev and colleagues’ key findings but offered an interpretation based on strategy differences between individuals lower and higher in cognitive resources. Specifically, high-span individuals are said to employ complex hypothesis testing such as *probability matching*, in which the next event to be predicted in a series is judged from the overall probability that the event has been shown to occur in the past. For example, if event “A” has occurred about

70% of the time up to the current trial, an individual using a probability matching strategy will choose this event about 70% of the time on the following trials. To follow this strategy, one must constantly mentally update the probabilities of past event occurrences and nonoccurrences and hypothesize about the likelihood of subsequent event occurrences based on this information. Although an impressive capability of higher capacity individuals, this strategy, on average, leads to lower accuracy (e.g., 58% in the previous example) than much less intensive strategies (Gaissmaier et al. 2006).

One less intensive strategy, *maximizing*, involves simply remembering which event has occurred most frequently in the past and always predicting that this dominant event will happen next. This simple strategy will generally produce greater accuracy than the more complicated probability matching approach. For example, if an event that has occurred 70% of the time is always predicted to occur next, one would be correct about 70% of the time. Of course, an individual using this strategy would not calculate this 70% probability, as he or she only gauges that one event seems to be happening more than the other. As noted by Gaissmaier and colleagues (2006), prior research has demonstrated that this simpler strategy is more often adopted by less intelligent individuals (Singer 1967), children (Derks and Paclisanu 1967), and even monkeys (Wilson and Rollin 1959) and pigeons (Herrnstein and Loveland 1975; Hinson and Staddon 1983). Consistent with the idea that maximizing is a simpler alternative to probability matching, Gaissmeier and colleagues found that individuals lower in short-term memory capacity were more likely to adopt this strategy.

Implementing a dual-task methodology often used to disrupt attentional control, Wolford and colleagues (2004) found a similar relation between decreased attentional control and the adoption of a simpler, but more effective, strategy. Specifically, in a probability-guessing paradigm similar to the correlation detection tasks mentioned above (Gaissmaier et al. 2006), individuals given a concurrent verbal-based task increased their use of maximizing relative to those in a single-task condition. Use of this less-demanding strategy improved probability-guessing performance. The previously mentioned studies (i.e., Kareev et al. 1997; Gaissmaier et al. 2006) linked strategy selection tendencies to individual difference measures rather than experimentally reduced attentional control. Thus, the finding from Wolford et al. of improved performance across all individuals when a secondary task is imposed provides a nice piece of converging evidence that decreased attentional capabilities elicit shortcut strategies that are sometimes better suited to the task at hand (Beilock and DeCaro 2007).

Toward a Comprehensive Understanding of the Benefits and Perils of Attentional Control

We have seen that individuals with greater attentional capacity available to them are inclined to use it, even when a task might benefit from less attentional control. For example, individuals with greater attentional control ability are sometimes more likely to focus selectively on less efficient problem solutions (Beilock and DeCaro 2007; Ricks et al. 2007), override the veritable responses of associative-based category learning (DeCaro et al. 2008; Markman et al. 2006), encode phrases of a new language holistically rather than analyzing important components (Newport 1990), and unnecessarily search for complicated patterns in a series of events (Wolford et al. 2004). Because the positive aspects of attentional control are so commonly seen, attention's accompanying pitfalls often receive little acknowledgment.

One question such work brings to the surface, however, is how to understand when "less is less" and "less is more." Fortunately, there is a literature one can look at to develop a more comprehensive understanding of attentional control and performance. The dual-process literature describes two types of cognitive processes used for performance across domains, differing specifically in their reliance on attentional control. By conceptualizing the tasks presented in this chapter in light of these overarching dual processes, we may begin to abstract a more comprehensive understanding of when explicit attention devoted to performance will hurt and when it will help.

Dual-process theories have become common across many domains, such as social cognition (Smith and DeCoster 2000), judgment and reasoning (De Neys 2006; Evans 2003; Sloman 1996; Stanovich and West 2000), attention (Barrett et al. 2004; Schneider and Shiffrin 1977; Shiffrin and Schneider 1977), and categorization (Ashby et al. 1998; Maddox and Ashby 2004), to name a few. The details and terminology vary from one particular theory to another, but recently there has been a drive to abstract generalities across the domain-specific dual-process theories (Kahneman 2003; Sloman 1996; Smith and DeCoster 2000). Most dual-process theories posit that optimal skill execution can differentially rely on one of two types of cognitive processes, generally believed to be functionally divided by separate neural pathways (Maddox and Ashby 2004; Smith and DeCoster 2000; see also Poldrack and Packard 2003). What we will refer to as *associative processing* (also referred to as implicit, automatic, intuitive, heuristic, procedural, or System 1) is said to operate automatically, without heavy use of working-memory resources (if any), largely outside of conscious awareness, and based on domain-specific stimulus–response pattern recognition and retrieval. In contrast, *rule-based processing* (also called explicit, controlled, analytic, algorithmic, declarative, or System 2) is thought to operate effortfully and sequentially, require attention and working-memory resources, and be available to conscious awareness, and it can (but does not necessarily always) utilize more domain-general symbolic processing (Sloman 1996; Smith and DeCoster 2000).

Associative and rule-based processes are said to operate separately, with both systems generating their own computational products such as problem solutions, category selections, or task approaches. Many times these systems act in concert, deriving the same output (De Neys 2006). For example, when viewing a robin for the first time, associative processing, using experiences with other animals with similar physical characteristics, would likely lead the robin to be classified as a bird. And the rule-based system, using sequential deliberation or hypothesis testing of the various features that constitute relevant categories, would likely arrive at the same conclusion. As another example, recall the math problem-solving work described earlier in this chapter (Beilock and DeCaro 2007). Individuals were asked to derive a solution to a math problem involving multiple steps, say (42 – 6) ÷ 6, and determine whether there is a remainder. A rule-based processing approach might entail performing the subtraction step first, using a borrow operation, and then dividing the result by 6, concluding that there is no remainder. An associative heuristic might instead easily derive an answer without heavy use of working-memory resources. Based on experiences where problems with all even numbers do not have remainders, an associatively derived answer of "no remainder" in this case would be consistent with that derived by rule-based processes.

Associative and rule-based processes may at other times conflict, deriving different responses to the same stimuli. For example, when asked to classify a dolphin, the associative system may readily derive the "fish" category, given the similarity of this animal to commonly encountered fish. The rule-based response, following a sequence of rules leading instead to a "mammal" classification, would therefore conflict with the associative-based response. Similarly, when asked whether (42 – 6) ÷ 8 has a remainder, the rule-based response would be "yes." However, the aforementioned associatively driven heuristic concerning all even numbers would incorrectly determine that "no," there is no remainder.

Situations in which rule-based and associative processes lead to conflicting solutions are of special interest, because they allow one to discern how these separate systems interact (Beilock and DeCaro 2007). Given that only one response can win out, in these conflicting situations we can ascertain whether the outcome is consistent with a rule-based or an associative processing strategy. If a solution is consistent with rule-based processing, then one might say that an approach involving explicit attentional control has been favored over more associatively driven and less attention-demanding processes. And the opposite can be said when results consistent with associative processes are seen.

Across the domains discussed in this chapter, we have seen countless examples of situations in which rule-based and associatively derived responses differentially lead to "correct" performance on a task. Attention-demanding rule-based processes prove successful for multistep computations in math problem solving (Beilock and DeCaro 2007) and the complex hypothesis testing required for rule-based categorization (DeCaro, et al. 2008; Markman et al. 2006). On the other hand, associatively driven processes lead to more efficient math problem solving on the water jug task (Beilock and DeCaro 2007; Gasper 2003) and quicker information-integration category learning (Ashby and Maddox 2005).

When conceptualizing rule-based processes as attention-demanding and associative processes as nondemanding, one might speculate as to when attention will hurt or harm performance on a particular task. Less-than-optimal skill performance may occur in those situations in which a mismatch between optimal and actual processing approaches occurs. For example, an individual solving multistep math problems with associative-based heuristics will generally perform more poorly than if he or she followed through with the complex rules required for the highest level of performance (Siegler 1988). Or an individual learning information-integration categories too complex to integrate within the bounds of working memory will learn even more slowly if trying to push a rule-based process on this more associative-based task (Zeithamova and Maddox 2006). The latter situation is an example of when less attention is more optimal for a task than an attention-demanding approach. Construing the findings reviewed in this chapter in this way, then, we may begin to establish a general rule about when attention will be beneficial versus harmful—attentional control will benefit tasks relying on attention-demanding rule-based processes but will hamper performance that more optimally relies on associative processing.

Of course the question remains—how do we know whether a task relies on rule-based or associative processes? There is little consensus regarding the defining characteristics of tasks that demand one or the other type of processing. Certain stimuli may evoke more associative processes by their physical similarity to objects stored in memory, whereas tasks requiring convoluted computations may demand rule-based processing (Kahneman 2003). If a task is believed to rely on rule-based processes, then concurrently performing an attention-demanding secondary task should disrupt performance (Kahneman 2003; Sloman 1996).

A further question centers on how we can determine whether a task is *best* performed by relying on rule-based or associative processes (e.g., Gaissmaier et al. 2006). "Optimal" performance is necessarily defined somewhat subjectively, with characteristics such as accuracy and speed factored into the equation. Cognitive economy can also play a role, in that a process used is described as the most optimal if the least amount of cognitive effort (e.g., attentional control) is exerted for the most adequate (e.g., quickest and most accurate) outcome. Although standards of optimality are widely debated (e.g., Gigerenzer and Todd 1999; Stanovich and West 2000), denoting the characteristics of rule-based versus associative processes can allow us to at least generally determine whether performance has been driven more so by one or the other processing strategy.

Considering tasks in terms of the processes required for successful performance also carries implications for skill types beyond those reviewed in this chapter. For example, high-level sensorimotor skill performance such as golf putting or soccer dribbling can be hurt or helped by attentional control depending on the level of practice an individual has with the skill. Novice performers rely on attention to execute the steps of a skill—knowledge about that skill is held in working memory and attended in a step-by-step fashion (Anderson 1982; Fitts and Posner 1967; Proctor and Dutta 1995). Thus, if attention cannot be sufficiently devoted to a novice skill, performance suffers. However, as a skill develops, its execution becomes more automatic, or procedural, in nature. The unintegrated control structures of the novice performer become integrated, running largely outside of conscious control (Anderson 1982; Fitts and Posner 1967). Thus, if an expert explicitly attends to the step-by-step performance of the skill itself, performance suffers (Beilock and Carr 2001; Beilock, Bertenthal, McCoy, and Carr 2004; Gray 2004; Jackson, Ashford, and Norsworthy 2006; Lewis and Linder 1997). Attending to the components of a procedural skill essentially reverts execution back to the unintegrated control processes of novices.

Thus, a dual-process perspective offers one way skill success and failure across disparate domains such as problem solving and sensorimotor skills may be understood by the common thread of attentional control. Notably, when considering skill execution in terms of rule-based versus associative processes, the traditional distinction between cognitive and motor tasks becomes blurred. Sports tasks such as expert soccer dribbling become classified with language learning and information-integration category learning, as all rely on associative processing and respond to attentional control in much the same way. With studies such as those reviewed here, we may also better inform existing dual-process theories by exposing the individual difference and situational factors that may impact the type of processing required and utilized for a given task. Whether associative processing is optimal, and whether rule-based processing wins out instead, will be determined by factors such as individual differences in attentional control, expertise, aspects of the performance environment, and the particular task itself. Such work allows us to begin to cut across research domains, not only speaking to a possible overarching theory of attentional control and skill performance but also providing a framework by which to inform future research endeavors.

Conclusion

It is commonly believed that the more extensively information is processed and attended to, the more optimal performance will be (Hertwig and Todd 2003). Such assertions are supported by the plethora of research demonstrating that working memory and attention are vital to performance across skill domains (Conway et al. 2005). Cognitive control abilities are held in such high esteem that the performance of those with more of these capabilities (i.e., individuals higher in intellectual or working-memory capacity) has been deemed the standard by which performance should be measured: "…whatever the 'smart' people do can be assumed to be right" (De Neys 2006, 432; Stanovich and West 2000). Even individuals who do not have the capacity to successfully perform working-memory-demanding processes are thought to adhere to the same norm as those higher in working memory, but they simply fall short in the capability to do so (De Neys 2006). As shown in this chapter, however, greater attentional control capabilities can impede performance, and individuals with less cognitive control can excel beyond their higher capacity counterparts by effectively utilizing simpler strategies. Such findings call into question the validity of characterizing attention-demanding processing strategies as the standard for rational or optimal behavior. They instead speak to the importance of considering not only cognitive capacity but also task demands and aspects of the performance environment when delineating the most "optimal" use of attention in any given performance scenario.

Acknowledgment

This research was supported by Institute of Education Sciences grant R305H050004 and National Science Foundation grant DRL-0746970 to Sian Beilock.

References

Anderson, J. R. 1982. Acquisition of cognitive skill. *Psychol. Rev.* 89:369–406.

Anderson, R. B., M. E. Doherty, N. D. Berg, and J. C. Friedrich. 2005. Sample size and the detection of correlation—A signal detection account: Comment on Kareev (2000) and Juslin and Olsson (2005). *Psychol. Rev.* 112:268–279.

Ashby, F. G., and W. T. Maddox. 2005. Human category learning. *Annu. Rev. Psychol.* 56:149–178.

Ashby, F. G., and J. B. O'Brien. 2005. Category learning and multiple memory systems. *Trends Cogn. Sci.* 9:83–89.

Ashby, F. G., L. A. Alfonso-Reese, A. U. Turken, and E. M. Waldron. 1998. A neuropsychological theory of multiple systems in category learning. *Psychol. Rev.* 105:442–481.

Ashcraft, M. H., and E. P. Kirk. 2001. The relationships among working memory, math anxiety, and performance. *J. Exp. Psychol. Gen.* 130:224–237.

Barrett, L. F., M. M. Tugade, and R. W. Engle. 2004. Individual differences in working memory capacity and dual-process theories of the mind. *Psychol. Bull.* 130:553–573.

Beilock, S. L. 2008. Math performance in stressful situations. *Curr. Dir. Psychol. Sci.* 17:339–343.

Beilock, S. L., and T. H. Carr. 2001. On the fragility of skilled performance: What governs choking under pressure? *J. Exp. Psychol. Gen.* 130:701–725.

Beilock, S. L., and T. H. Carr. 2005. When high-powered people fail: Working memory and "choking under pressure" in math. *Psychol. Sci.* 16:101–105.

Beilock, S. L., and M. S. DeCaro. 2007. From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *J. Exp. Psychol. Learn. Mem. Cogn.* 33:983–998.

Beilock, S. L., T. H. Carr, C. MacMahon, and J. L. Starkes. 2002. When paying attention becomes counterproductive: Impact of divided versus skill-focused attention on novice and experienced performance of sensorimotor skills. *J. Exp. Psychol. Appl.* 8:6–16.

Beilock, S. L., B. I. Bertenthal, A. M. McCoy, and T. H. Carr. 2004. Haste does not always make waste: Expertise, direction of attention, and speed versus accuracy in performing sensorimotor skills. *Psychon. Bull. Rev.* 11:373–379.

Beilock, S. L., C. A. Kulp, L. E. Holt, and T. H. Carr. 2004. More on the fragility of performance: Choking under pressure in mathematical problem solving. *J. Exp. Psychol. Gen.* 133:584–600.

Cahan, S., and Y. Mor. 2007. The effect of working memory capacity limitations on the intuitive assessment of correlation: Amplification, attenuation, or both? *J. Exp. Psychol. Learn. Mem. Cogn.* 33:438–442.

Cochran, B. P., J. L. McDonald, and S. J. Parault. 1999. Too smart for their own good: The disadvantage of a superior processing capacity for adult language learners. *J. Mem. Lang.* 41:30–58.

Conway, A. R. A., N. Cowan, and M. F. Bunting. 2001. The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychon. Bull. Rev.* 8:331–335.

Conway, A. R. A., M. J. Kane, M. F. Bunting, D. Z. Hambrick, O. Wilhelm, and R. W. Engle. 2005. Working memory span tasks: A methodological review and user's guide. *Psychon. Bull. Rev.* 12:769–786.

Daneman, M., and P. A. Carpenter. 1980. Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* 19:450–466.

De Neys, W. 2006. Dual processing in reasoning: Two systems but one reasoner. *Psychol. Sci.* 17:428–433.

DeCaro, M. S., R. D. Thomas, and S. L. Beilock. 2008. Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition* 107:284–294.

Derks, P. L., and M. I. Paclisanu. 1967. Simple strategies in binary prediction by children and adults. *J. Exp. Psychol.* 73:278–285.

Dougherty, M. R. P., and J. E. Hunter. 2003. Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychol. (Amst.)* 113:263–282.

Duncker, K. 1945. On problem-solving. *Psychol. Monogr.* 58:1–113.

Engle, R. W. 2002. Working memory capacity as executive attention. *Curr. Dir. Psychol. Sci.* 11:19–23.

Ericsson, K. A., and P. R. Delaney. 1999. Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In *Models of working memory: Mecha-*

nisms of active maintenance and executive control, ed. A. Miyake and P. Shah. New York: Cambridge University Press, 257–297.

Evans, J. St. B. T. 2003. In two minds: Dual-process accounts of reasoning. *Trends Cogn. Sci.* 7:454–459.

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

Gaissmaier, W., L. J. Schooler, and J. Rieskamp. 2006. Simple predictions fueled by capacity limitations: When are they successful? *J. Exp. Psychol. Learn. Mem. Cogn.* 32:966–982.

Gasper, K. 2003. When necessity is the mother of invention: Mood and problem solving. *J. Exp. Soc. Psychol.* 39:248–262.

Gigerenzer, G., and P. M. Todd. 1999. Fast and frugal heuristics: The adaptive toolbox. In *Simple heuristics that make us smart*, ed. G. Gigerenzer and P. M. Todd and the The ABC Research Group. New York: Oxford University Press, 3–34.

Gray, R. 2004. Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *J. Exp. Psychol. Appl.* 10:42–54.

Herrnstein, R. J., and D. H. Loveland. 1975. Maximizing and matching on concurrent ratio schedules. *J. Exp. Anal. Behav.* 24:107–116.

Hertwig, R., and P. M. Todd. 2003. More is not always better: The benefits of cognitive limits. In *Thinking: Psychological perspectives on reasoning, judgment and decision making*, ed. D. Hardman and L. Macchi. Chichester, UK: Wiley, 213–231.

Hinson, J. M., and J. E. R. Staddon. 1983. Matching, maximizing and hillclimbing. *J. Exp. Anal. Behav.* 40:321–331.

Jackson, R. C., K. J. Ashford, and G. Norsworthy. 2006. Attentional focus, dispositional reinvestment, and skilled motor performance under pressure. *J. Sport Exerc. Psychol.* 28:49–68.

Juslin, P., and H. Olsson. 2005. Capacity limitations and the detection of correlations: Comment on Kareev (2000). *Psychol. Rev.* 112:256–267.

Kahneman, D. 2003. A perspective on judgment and choice: Mapping bounded rationality. *Am. Psychol.* 58:697–720.

Kane, M. J., and R. W. Engle. 2000. Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26:336–358.

Kane, M. J., and R. W. Engle. 2002. The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychon. Bull. Rev.* 9:637–671.

Kane, M. J., and R. W. Engle. 2003. Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *J. Exp. Psychol. Gen.* 132:47–70.

Kane, M. J., K. M. Bleckley, A. R. A. Conway, and R. W. Engle. 2001. A controlled-attention view of working-memory capacity. *J. Exp. Psychol. Gen.* 130:169–183.

Kareev, Y. 2005. And yet the small-sample effect does hold: Reply to Juslin and Olsson (2005) and Anderson, Doherty, Berg, and Friedrich (2005). *Psychol. Rev.* 112:280–285.

Kareev, Y., I. Lieberman, and M. Lev. 1997. Through a narrow window: Sample size and the perception of correlation. *J. Exp. Psychol. Gen.* 126:278–287.

Kersten, A. W., and J. L. Earles. 2001. Less really is more for adults learning a miniature artificial language. *J. Mem. Lang.* 44:250–273.

Lewis, B., and D. E. Linder. 1997. Thinking about choking? Attentional processes and paradoxical performance. *Pers. Soc. Psychol. Bull.* 23:937–944.

Luchins, A. S. 1942. Mechanization in problem solving: The effect of Einstellung. *Psychol. Monogr.* 54:1–95.

Maddox, W. T., and F. G. Ashby. 2004. Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behav. Processes* 66:309–332.

Maddox, W. T., B. C. Love, B. D. Glass, and J. V. Filoteo. 2008. When more is less: Feedback effects in perceptual category learning. *Cognition* 108:578–589.

Markman, A. B., W. T. Maddox, and D. A. Worthy. 2006. Choking and excelling under pressure. *Psychol. Sci.* 17:944–948.

Mednick, S. 1962. The associative basis of the creative process. *Psychol. Rev.* 69:220–232.

Miyake, A., and P. Shah. 1999. *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.

Newport, E. L. 1990. Maturational constraints on language learning. *Cogn. Sci.* 14:11–28.

Poldrack, R. A., and M. G. Packard. 2003. Competition among memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia* 41:245–251.

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

Ricks, T. R., K. J. Turley-Ames, and J. Wiley. 2007. Effects of working memory capacity on mental set due to domain knowledge. *Mem. Cognit.* 35:1456–1462.

Rosen, V. M., and R. W. Engle. 1997. The role of working memory capacity in retrieval. *J. Exp. Psychol. Gen.* 126:211–227.

Schneider, W., and R. M. Shiffrin. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* 84:1–66.

Shiffrin, R. M., and W. Schneider. 1977. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychol. Rev.* 84:127–190.

Siegler, R. S. 1988. Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Dev.* 59:833–851.

Singer, E. 1967. Ability and the use of optimal strategy in decisions. *Am. J. Psychol.* 80: 243–249.

Sloman, S. A. 1996. The empirical case for two systems of reasoning. *Psychol. Bull.* 119:3–22.

Smith, E. R., and J. DeCoster. 2000. Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Pers. Soc. Psychol. Rev.* 4:108–131.

Stanovich, K. E., and R. F. West. 2000. Individual differences in reasoning: Implications for the rationality debate? *Behav. Brain Sci.* 23:645–665.

Turner, M. L., and R. W. Engle. 1989. Is working memory capacity task dependent? *J. Mem. Lang.* 28:127–154.

Waldron, E. M., and F. G. Ashby. 2001. The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychon. Bull. Rev.* 8:168–176.

Wiley, J. 1998. Expertise as mental set: The effects of domain knowledge in creative problem solving. *Mem. Cognit.* 26:716–730.

Wilson, W. A., Jr., and A. R. Rollin. 1959. Two-choice behavior of rhesus monkeys in a noncontingent situation. *J. Exp. Psychol.* 58:174–180.

Wolford, G., S. E. Newman, M. B. Miller, and G. S. Wig. 2004. Searching for patterns in random sequences. *Can. J. Exp. Psychol.* 58:221–228.

Zeithamova, D., and W. T. Maddox. 2006. Dual-task interference in perceptual category learning. *Mem. Cognit.* 34:387–398.