

Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence

DAVID GIMMIG, PASCAL HUGUET, and JEAN-PAUL CAVERNI
CNRS and University of Aix-Marseille 1, Marseille, France

and

FRANÇOIS CURY
CNRS and University of Aix-Marseille 1, Marseille, France
and University of Nice Sophia Antipolis, Nice, France

Recent findings (Beilock & Carr, 2005) have demonstrated that only individuals with a high working memory capacity (WMC) “choke under pressure” on math problems with high working memory demands. This suggests that performance pressure hinders those who are the most qualified to succeed, because it consumes the WMC they usually rely on to achieve superior performance. This puts into question the use of performance in high-pressure situations as a means of distinguishing individuals with lesser or greater WMC potentials. While addressing several limitations of past research, we offer evidence that such choking (1) occurs only in individuals with high WMC, because of their anxiety-ridden perceptions of high-stakes situations, and (2) is not confined to tasks involving acquired skills and knowledge, but encompasses fluid reasoning abilities or intelligence (*Gf*). These findings have strong implications for assessments of people’s intellectual capacities in academic, clinical, work, and research settings.

“Choking under pressure” is an expression used to refer to performing more poorly than expected—given one’s skill—in situations where the pressure to perform is high (i.e., where the desire to excel is maximal; see Baumeister, 1984; Beilock, Kulp, Holt, & Carr, 2004). As noted by Beilock and Carr (2005), the consequences of suboptimal performance, especially on examinations, include poor evaluations by mentors, teachers, and peers; lost scholarships; and relinquished educational and employment opportunities. Examining the relationship between choking under pressure and individual differences in working memory capacity (WMC), these authors concluded that the individuals who are the most likely to fail under performance pressure are those who, in the absence of pressure, have the highest potential for success (i.e., high-WMC individuals). Performance pressure, they reasoned, hinders the most qualified people by consuming the WMC they use in low-pressure circumstances to devise complex (resource-demanding) strategies and produce superior performance. In their study,

however, Beilock and Carr (2005) focused only on arithmetic problems (involving both subtraction and division).

In the present study, we examined the generalizability of Beilock and Carr’s (2005) conclusion using a task (Raven’s standard progressive matrices, or SPMs) known to measure *fluid intelligence* (*Gf*). Unlike *crystallized intelligence*, which quantifies aptitudes that are linked to acquired skills and knowledge, *Gf* is defined as the ability to reason on largely unfamiliar materials and problems (Cattell, 1943; Raven, Raven, & Court, 1998). The results of our experiment reveal that the occurrence of choking in high-powered people is not confined to tasks involving acquired skills and knowledge, but encompasses those that call for fluid reasoning abilities or intelligence as well. Our results also help to clarify how individuals who are high in WMC perceive high-performance-pressure situations, with strong implications for the assessment of people’s intellectual capacity.

Relating WMC and Choking Under Pressure

WMC is typically considered to reflect individual differences in the amount of goal-directed attention that is available for temporarily keeping information of interest activated while inhibiting irrelevant thoughts and thus preventing distraction (Barrett, Tugade, & Engle, 2004, p. 560). In exploring the potential impact of individual WMC differences on susceptibility to choking, Beilock and Carr (2005) brought three preliminary findings to bear: (1) Performance pressure increases reported state

This research was supported in part by a graduate fellowship from the Conseil Regional PACA to the first author and by CNRS Grant JC 6082 to P.H. The article is based on a doctoral dissertation by D.G. under the supervision of P.H. and J.-P.C. at the University of Aix-Marseille 1. The authors thank Sian L. Beilock, Mark Ashcraft, and Randall W. Engle for their helpful comments on an earlier version of this article. Correspondence should be addressed to P. Huguet, Université Aix-Marseille 1, 3 Place Victor Hugo, Case D, 13331 Marseille, France (e-mail: huguet@up.univ-mrs.fr).

anxiety (Beilock & Carr, 2001; Beilock et al., 2004); (2) intrusive thoughts, such as worries, reduce the storage and processing capacity of working memory (Eysenck & Calvo, 1992); and (3) anxiety-provoking situations reduce attentional resources, resulting in performance decrements (Ashcraft & Kirk, 2001; Beilock et al., 2004; Schmader & Johns, 2003). On this basis, they suggested that if pressure-induced consumption of working memory denies high-WMC individuals (HWMs) the resources they normally rely on to achieve superior performance, then their usual working memory advantage may be exactly what makes them prone to fail under pressure. In their experiment, the HWM advantage that was found on the most memory-demanding problems disappeared when pressure (manipulated within subjects) was applied.

The Present Experiment

With this experiment, we tested whether Beilock and Carr's (2005) fascinating findings can be generalized to *Gf*. WMC measures are indeed strongly related to various psychometric measures of intelligence, especially *Gf* (see, e.g., Kane et al., 2004). If pressure-induced consumption of working memory denies HWMs the resources needed to mobilize their advantageous strategies, then their *Gf* level should be suboptimal when pressure is applied. We chose Raven's SPMs—one of the best measures of *Gf*—which have the added advantage of being progressively more difficult, starting from problems low in working memory demands (involving gestalt continuation and visuospatial abilities; see Appendix A) and ending with high-demand problems requiring verbal-analytic reasoning (see Appendix B). If pressure consumes WMC, choking should be specific to verbal-analytic/high-demand problems.

Our experiment also addressed four limitations of Beilock and Carr's (2005) study. First, their manipulation of pressure implicated a combination of factors (e.g., a camera, money incentives) that prevented them from identifying the main component of performance pressure. Here, we induced pressure simply by characterizing the SPMs as diagnostic or nondiagnostic of analytic reasoning (a core faculty in the participants' scientific curriculum). Second, although Beilock and Carr (2005) associated pressure with increased state anxiety, they did not test this specific point. We measured state anxiety. In addition, when pressure is applied, HWMs may display a tendency to excuse their performance. We assessed such a tendency through a series of self-handicapping items. Third, Beilock and Carr (2005) measured WMC after testing participants on arithmetic problems. Although they used filler tasks between test problems, there is still a possibility that their WMC measures captured some pressure effects. In the present study, WMC was measured before the pressure induction. Finally, because WMC predicts success on a broad range of cognitive tasks (Engle & Kane, 2004), it could be that Beilock and Carr's (2005) findings were not due to the participants' WMC per se, but to their feelings of superiority or inferiority in the focal task domain. In the present experiment, effects on performance were calculated while this potential confounding factor was controlled.

METHOD

Participants

The participants were 67 undergraduates (35 females and 32 males; mean age = 22 years, $SD = 4.61$) from the University of Provence who had a concentration in math and science during secondary school. They agreed to take part in the study, which was presented as part of a larger project focusing on students' memory, attention, and reasoning abilities.

Procedure

First, participants (tested individually) were asked to use 5-point scales, ranging from 1 (*very inferior*) to 5 (*very superior*), to rate their standing in comparison with their classmates on mathematical ability, capacity to succeed in school, general intelligence, attention, reasoning, verbal ability, and general knowledge. (Hereafter, these ratings will be referred to as *comparative evaluations*).

Next, their WMCs were measured using Kane et al.'s (2004) computer-based version of the classic Reading Span Task (RSPAN; Daneman & Carpenter, 1980). Each display included a meaningful or meaningless sentence that the participants read aloud while verifying whether or not it made sense, and a to-be-remembered letter, which they also read aloud (e.g., "We were fifty lawns out at sea before we lost sight of land. ? X"). The sentences (each composed of 12–17 words, $M = 14.4$ words, $SD = 1.2$) were taken from the French version of the RSPAN (Desmette, Hupet, Schelstraete, & Van der Linden, 1995). Following the procedure of Kane et al., the meaningless sentences were created by changing only one word (*miles to lawns* in the previous example). The set size ranged from two to five sentence-letter problems per trial, with 3 trials per set size, for a total of 12 trials. At the end of each set, participants had to write down the sequence of letters in the correct order. An item was scored as correct when it was recalled in the correct serial position. RSPAN scores (range, 0%–100%) were equal to the total number of correctly recalled letters divided by the total number of letters to be recalled (42).

Participants were then introduced to the SPMs. For half the participants, the task was described as measuring analytic reasoning, which was explicitly associated with overall success in mathematics and science (high-pressure condition). For the other half, it was described as simply measuring attentional and perceptual capacities (low-pressure condition). All participants were then given a general description of the task via two sample exercises, an easy one (the first SPMs; see Appendix A) and a difficult one (from the advanced progressive matrices; see Appendix B). The experimenter did not give the correct answers or any feedback. Next, the participants had to use a 7-point scale, ranging from 1 (*none*) to 7 (*all*), to say how many exercises they expected to answer correctly. Then, they were given both the low- and the high-demand SPM exercises (see Lynn, Allik, & Irwing, 2004; van der Ven & Ellis, 2000, for this categorization). Both accuracy (percentage of correct answers) and inspection times (the time it took a participant to answer a problem) were measured. The inspection time ran from the start of a trial or exercise to the validation of the participant's choice. In the present experiment, only inspection times on correct exercises were considered. When they had finished the 60 exercises or when the allotted time (15 min) was up, the participants rated how frequently they felt anxious, comfortable, jittery, worried, at ease, nervous, relaxed, and calm while doing the SPMs, using a 5-point scale ranging from 1 (*never*) to 5 (*always*) on a questionnaire adapted from Spielberger, Gorsuch, and Lushene's (1970) state anxiety scale (see Schmader & Johns, 2003). Finally, participants rated statements about how handicapped they had felt during the task performance on a 5-point scale ranging from 1 (*totally disagree*) to 5 (*totally agree*) (e.g., "I slept well last night"; adapted from Rhodewalt, 1990). The items on each scale were averaged to obtain a state anxiety index ($\alpha = .87$) and a self-reported handicap index ($\alpha = .69$). These indexes assessed how participants considered the SPMs as a whole, regardless of problem demand (a within-subjects factor).

RESULTS

SPMs

Participants were assigned to low and high working memory groups (LWMs, $M = .48$, $SE = .01$; HWMs, $M = .73$, $SE = .01$), using a median split of RSPAN scores (ranging from .33 to .95) as in Beilock and Carr (2005). The performance data (SPM accuracy and inspection time) were analyzed in a 2 (WMC: LWM vs. HWM) \times 2 (problem demands: low vs. high) \times 2 (pressure: low vs. high) mixed ANOVA, with problem demands as a within-subjects factor.

Accuracy (% correct). The main effect of problem demands was clearly significant [$F(1,63) = 243.65$, $p < .0001$, $\eta_p^2 = .81$]. Accuracy was better for low-demand ($M = 94\%$, $SE = .01$) than for high-demand ($M = 58\%$, $SE = .03$) problems. As was expected, this effect was qualified by a significant three-way interaction [$F(1,63) = 4.21$, $p < .05$, $\eta_p^2 = .07$]. Planned comparisons derived from Beilock and Carr's (2005) findings showed that on low-demand problems HWMs did not differ from LWMs in either the low-pressure ($M = 95\%$ vs. 92% , respectively) or high-pressure ($M = 94\%$ vs. 93% , respectively) condition. Also, as was expected, on high-demand problems (see Figure 1, left panel), HWMs outperformed LWMs in the low-pressure condition [$t(63) = 3.72$, $p < .001$] (unless otherwise mentioned, all t tests in the Results section are two-tailed). This difference was not ob-

served when pressure was applied [$t(63) = 0.69$, n.s.]. The LWMs' performance did not suffer from pressure [$t(63) = -1.32$, $p = .19$], and even tended to benefit from it, whereas the HWMs' performance did suffer [$t(63) = 1.67$, $p = .05$, one-tailed]. Moreover, the probability of replicating the three-way interaction was high ($p_{rep} = .89$; see Killeen, 2005).

To clarify the findings, performance accuracy on high-demand problems was regressed over continuous RSPAN scores, pressure, and their cross product. In line with the cognitive literature on working memory (e.g., Unsworth & Engle, 2005), RSPAN scores predicted accuracy ($\beta = .38$, $p < .001$). However, as could be expected from the results above, the RSPAN \times pressure interaction was clearly significant ($\beta = -.31$, $p < .01$). Although the RSPAN scores predicted accuracy under low pressure ($\beta = .60$, $p < .0001$), this was not the case when pressure was applied ($\beta = .09$, $p = .64$). This interaction remained consistently significant when the participants' comparative evaluations—including reasoning and mathematical ability¹—were controlled. Gender had no effect either.

Inspection time (IT) on correct exercises. The main effect of problem demands was significant [$F(1,63) = 286.69$, $p < .0001$, $\eta_p^2 = .82$]. ITs were longer for high-demand ($M = 24.20$ sec, $SE = 0.93$) than for low-demand ($M = 10.13$ sec, $SE = 0.31$) problems. No other effects were found.

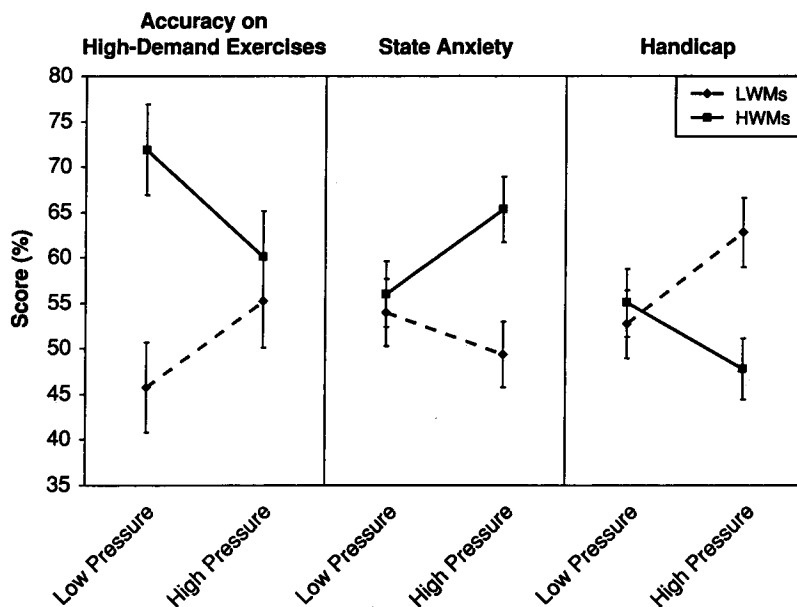


Figure 1. Performance accuracy on high-demand problems, state anxiety, and self-reported handicap for low (LWM) versus high (HWM) working memory capacity individuals. To allow the anxiety and handicap indexes to be portrayed on the same measurement scale as performance on the standard progressive matrices (SPMs), these indexes were transformed to percentages (by first summing the values of their constituent items and then dividing by the maximum sum obtainable). By doing so, we hope to visually render how performance on the SPMs was linked to the scores on the anxiety and handicap scales. Error bars represent standard errors.

Performance Expectation Ratings, Anxiety, and Handicaps

No effects were found on the expectation ratings ($M = 4.69$, $SE = 0.13$, overall).

On state anxiety, the WMC \times pressure interaction was marginally significant [$F(1,63) = 3.63$, $p = .06$, $\eta_p^2 = .06$]. HWMs and LWMs did not differ from each other under low pressure [$t(63) = 0.40$, n.s.], whereas the former reported significantly more anxiety than the latter when pressure was applied [$t(63) = 3.07$, $p < .01$] (see Figure 1, middle panel). Anxiety increased in HWMs under pressure [$t(63) = 1.81$, $p < .04$, one-tailed], but not in LWMs [$t(63) = 0.89$, n.s.].

The WMC \times pressure interaction was significant on self-reported handicaps [$F(1,63) = 5.29$, $p < .03$, $\eta_p^2 = .08$]. HWMs and LWMs were indistinguishable under low pressure [$t(63) = 0.44$, n.s.], whereas (contrary to our prediction) the latter reported more handicaps than the former when the pressure was great [$t(63) = -2.79$, $p < .01$] (see Figure 1, right panel). The self-handicapping tendency tended to increase in LWMs under pressure [$t(63) = -1.87$, $p = .06$], but not in HWMs [$t(63) = 1.38$, $p = .17$], quite the contrary.

The anxiety and handicap indexes were not correlated overall or within each condition taken separately.

Mediation

Anxiety mediated (see, e.g., Baron & Kenny, 1986) choking among HWMs. Additional tests (MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002) showed that the direct path coefficient from pressure to performance ($\beta = -.33$, $p < .05$) decreased significantly ($p < .05$) when performance was regressed on both pressure and anxiety ($\beta = -.22$, n.s.).²

DISCUSSION

The present findings merit special attention. They offer new evidence that, on tasks with high WM demands, the individuals who are most likely to fail under performance pressure are those with high WMC. Although potential confounding factors were addressed (WMC was measured before pressure was induced; pressure was manipulated between subjects via task characterization only, and comparative evaluations were controlled), our performance effects were clearly consistent with those reported by Beilock and Carr (2005). Our findings also show that choking under pressure is due to increased state anxiety, which is in line with Beilock and Carr's (2005) explanation that WMC is jeopardized by pressure via increased anxiety. In accordance with this explanation, however, pressure should increase anxiety in LWMs as well, which was clearly not the case here. Does this mean that we simply failed to induce pressure in our LWMs? Probably not, given that the LWMs reported feeling more handicapped under pressure than did the HWMs. This rules out the possibility that they were totally unaffected by pressure. Taken together, our results suggest that not all individu-

als experience high-performance-pressure situations in the same way: Reactions to pressure appear to depend on WMC. It seems, therefore, that choking occurs in HWMs alone, because of their anxious perception of high-stakes situations at the onset.

Alternatively, our findings could be taken as evidence that those with high spans tend to use a much more controlled processing approach (to the focal task) than do those with low spans, who primarily make use of automatic processes (e.g., automatic spreading activation; see Kane & Engle, 2000; Rosen & Engle, 1997). According to this account, the additional load (in this case, probably worries or intrusive thoughts) interfered with controlled processing and not with automatic processing. But for this alternative account to be valid, the load that the low spans have to deal with must equal the one that the high spans must deal with, which does not seem true here (at least when the anxiety findings are considered). Once more, it can be reasonably assumed that pressure induced a higher load in the HWMs than in the LWMs. Future research is needed on this crucial point, however.

One may also wonder whether the way pressure was manipulated in this experiment was isomorphic with the method used in previous work (Beilock & Carr, 2005). After all, varying the diagnosticity of the task comes close to the pressure manipulation that is typically used in the stereotype-threat paradigm (see Steele, Spencer, & Aronson, 2002, for a review). In this paradigm, however, the performance deficit that is frequently observed on difficult tests—including those measuring WMC (see Schmader & Johns, 2003)—is not due to the pressure that may be associated with taking a difficult intellectual test, per se. This deficit is due to the extra pressure that is associated with the possibility that one's performance will confirm—to others, to oneself, or to both—a negative stereotype about one's group. Although there is some reason here to believe that the HWMs experienced more pressure than their LWM counterparts, there is no reason to think that the former felt threatened by the possibility of confirming a negative stereotype. The question arising here is, exactly which stereotype could threaten the HWMs in the diagnostic condition? A stereotype-threat approach in the present study would lead toward expecting a suboptimal performance in women under this condition, because of the well-known negative stereotype about their gender's mathematical and scientific abilities. No gender effects were found in the present study, however. Furthermore, the assumption that our pressure manipulation came close to that used in the stereotype-threat literature would be a somewhat weak criticism. There is indeed no guarantee that pressure and stereotype-threat manipulations do not produce their effects through a common mediating variable (e.g., working memory). Most people in the field would accept the term *performance pressure* as describing any manipulation (including stereotype threat) that motivates people to try harder to perform well.

Finally, earlier in this article, we assumed that low-demand problems of the SPMs were more visuospatially

oriented whereas the high-demand ones were more verbally oriented. To the extent that the diagnostic or high-pressure condition was associated with increased worries about the situation and that the related consequences were verbal in nature (see, e.g., Beilock et al., 2004; Cadinu, Maass, Rosabianca, & Kiesner, 2005), it could be that the impact on the high-demand problems was not due to the degree of working memory reliance per se, but to the type of reliance. Although Kane et al. (2004) offered evidence that WMC captures domain-general information processing capabilities (see, however, Baddeley, 1996; Shah & Miyake, 1996), this possibility might also be explored in future research.

Above all, our findings show that pressure-induced performance decrements in HWMs are not limited to tasks involving acquired skills and knowledge; rather, they cover fluid reasoning or intelligence (*Gf*). As such, they may have important implications. *Gf* tests are widely used to distinguish high achievers from low achievers in academic, clinical, occupational, and research settings (see, e.g., Kane et al., 2004; Raven et al., 1998; Schmidt & Hunter, 1998). Because pressure is generally high in these settings, individuals with different capacity-related potentials may ironically display approximately the same level of fluid ability—as revealed by our results. Combined with the results of other studies (Beilock & Carr, 2005; Croizet et al., 2004; Huguet, Galvaing, Monteil, & Dumas, 1999; Schmader & Johns, 2003; Steele, 1997), the present findings offer new evidence that the test situation per se matters: It is an integral part of cognitive functioning, not just a surrounding context.

REFERENCES

- ASHCRAFT, M. H., & KIRK, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*, 224-237.
- BADDELEY, A. D. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology*, *49A*, 5-28.
- BARON, R. M., & KENNY, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology*, *51*, 1173-1182.
- BARRETT, L. F., TUGADE, M. M., & ENGLE, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin*, *130*, 553-573.
- BAUMEISTER, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality & Social Psychology*, *46*, 610-620.
- 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.
- 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.
- 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.
- CADINU, M., MAASS, A., ROSABIANCA, A., & KIESNER, J. (2005). Why do women underperform under stereotype threat? *Psychological Science*, *16*, 572-578.
- CATTELL, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin*, *40*, 153-193.
- CROIZET, J.-C., DESPRÉS, G., GAUZINS, M.-E., HUGUET, P., LEYENS, J.-P., & MÉOT, A. (2004). Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality & Social Psychology Bulletin*, *30*, 721-731.
- DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, *19*, 450-466.
- DESMETTE, D., HUPET, M., SCHELSTRAETE, M.-A., & VAN DER LINDEN, M. (1995). Adaptation en langue française du "Reading Span Test" de Daneman et Carpenter (1980) [A French version of M. Daneman and P. A. Carpenter's (1980) Reading Span Test]. *L'Année Psychologique*, *95*, 459-482.
- ENGLE, R. W., & KANE, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 44, pp. 145-199). New York: Academic Press.
- EYSENCK, M. W., & CALVO, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, *6*, 409-434.
- HUGUET, P., GALVAING, M. P., MONTEIL, J. M., & DUMAS, F. (1999). Social presence effects in the Stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality & Social Psychology*, *77*, 1011-1025.
- KANE, M. J., & ENGLE, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *26*, 336-358.
- KANE, M. J., HAMBRICK, D. Z., TUHOLSKI, S. W., WILHELM, O., PAYNE, T. W., & ENGLE, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*, 189-217.
- KILLEEN, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, *16*, 345-353.
- LYNN, R., ALLIK, J., & IRWING, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, *32*, 411-424.
- MACKINNON, D. P., LOCKWOOD, C. M., HOFFMAN, J. M., WEST, S. G., & SHEETS, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, *7*, 83-104.
- RAVEN, J. C., RAVEN, J. E., & COURT, J. H. (1998). *Progressive matrices*. Oxford: Oxford Psychologists Press.
- RHODEWALT, F. (1990). Self-handicappers: Individual differences in the preference for anticipatory self-protective acts. In R. L. Higgins, C. R. Snyder, & S. Berglas (Eds.), *Self-handicapping: The paradox that isn't* (pp. 69-106). New York: Plenum.
- ROSEN, V. M., & ENGLE, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*, 211-227.
- SCHMADER, T., & JOHNS, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality & Social Psychology*, *85*, 440-452.
- SCHMIDT, F. L., & HUNTER, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262-274.
- SHAH, P., & MIYAKE, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, *125*, 4-27.
- SPIELBERGER, C. D., GORSUCH, R. L., & LUSHENE, R. E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- STEELE, C. M. (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 threat and social identity threat. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 34, pp. 379-440). San Diego: Academic Press.
- UNSWORTH, N., & ENGLE, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between operation span and Raven. *Intelligence*, *33*, 67-81.

VAN DER VEN, A. H. G. S., & ELLIS, J. L. (2000). A Rasch analysis of Raven's standard progressive matrices. *Personality & Individual Differences*, 29, 45-64.

[$t(65) = -1.58$, n.s.], reasoning ability [$t(65) = -0.93$, n.s.], or general intelligence [$t(65) = -0.25$, n.s.]. This suggests that differences in a comparative evaluation do not explain why HWMs reported more anxiety than LWMs (see the Results section on anxiety).

NOTES

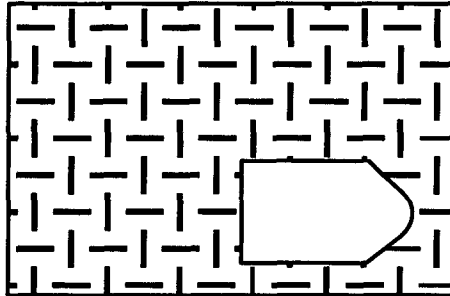
1. Additional analysis revealed that the LWM and HWM groups did not differ on perceived relative standing regarding their math ability

2. Among LWMs, the anxiety index correlated negatively, although not significantly, with accuracy on the high-demand exercises in both high-pressure ($r = -.38$) and low-pressure ($r = -.42$) conditions.

APPENDIX A

Example of a Low-Demand Exercise Based on Gestalt Continuation

Item 01

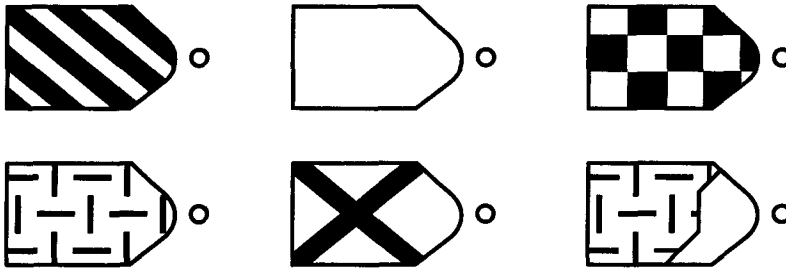


Choose the missing piece among the proposed ones.

To select your answer, click on the small round button next to it.

You can correct your answer by simply selecting another one.

When sure of your answer, validate it by clicking on the "Validate" button below.

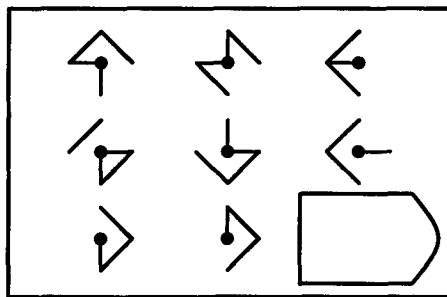


Validate

APPENDIX B

Example of a High-Demand Exercise Based on Verbal-Analytic Reasoning

Item 02

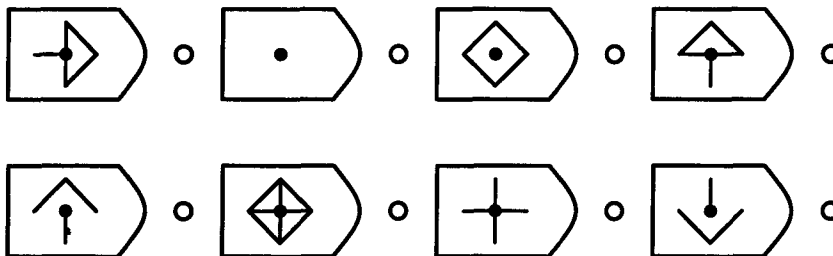


Choose the missing piece among the proposed ones.

To select your answer, click on the small round button next to it.

You can correct your answer by simply selecting another one.

When sure of your answer, validate it by clicking on the "Validate" button below.



Validate