Cognition 111 (2009) 397–403

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Discussion

When and how less is more: reply to Tharp and Pickering

Marci S. DeCaro [a], Krista D. Carlson [a], Robin D. Thomas [a], Sian L. Beilock [b,*]

[a] Miami University, Department of Psychology, Oxford, OH 45056, USA
[b] The University of Chicago, Department of Psychology, Chicago, IL 60637, USA

ARTICLE INFO

Article history:
Received 29 October 2008
Revised 23 February 2009
Accepted 1 March 2009

Keywords:
Category learning
Categorization
Working memory
Attention
Individual differences
Strategy selection
Hypothesis testing

ABSTRACT

In DeCaro et al. [DeCaro, M. S., Thomas, R. D., & Beilock, S. L. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition*, *107*, 284–294] we demonstrated that sometimes less working memory (WM) has its advantages. The lower individuals' WM, the faster they achieved success on an information-integration (II) category learning task adopted from Waldron and Ashby [Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review*, *8*, 168–176]. We attributed this success to the inability of lower WM individuals to employ explicit learning strategies heavily reliant on executive control. This in turn, we hypothesized, might push lower WM individuals to readily adopt procedural-based strategies thought to lead to success on the II task. Tharp and Pickering [Tharp, I. J., & Pickering, A. D. (2009). A note on DeCaro, Thomas, and Beilock (2008): Further data demonstrate complexities in the assessment of information-integration category learning. *Cognition*] recently questioned whether the II category learning task DeCaro et al. used really reflects procedural learning. In an effort to investigate Tharp and Pickering's assertions with respect to individual differences in WM, we replicate and extend our previous work, in part by modeling participants' response strategies during learning. We once again reveal that lower WM individuals demonstrate earlier II learning than their higher WM counterparts. However, we also show that low WM individuals' initial success is not because of procedural-based responding. Instead, individuals lower in WM capacity perseverate in using simple rule-based strategies that circumvent heavy demands on WM while producing above-chance accuracy.

1. Introduction

We (DeCaro, Thomas, & Beilock, 2008) recently demonstrated that the correlation between individual differences in working memory (WM) and performance is not always positive. Rather, we showed in a category learning paradigm that the relation between WM and the ability to acquire new categories depends on the nature of the category structure being tested. The higher individuals' WM, the faster they learned rule-based (RB) categories but the slower they learned information-integration (II) categories.

We predicted these findings by drawing on the COVIS model of categorization (Ashby & Maddox, 2005). This model states that II category learning will be slower to the extent that an individual persists in relying on explicit strategies to generate category membership in lieu of adopting the procedurally-driven strategies most optimal for learning II tasks (e.g., Zeithamova & Maddox, 2006). We reasoned that, because individuals higher in WM capacity are better able to explicitly attend to multiple stimulus features and test complex hypotheses (Dougherty & Hunter, 2003), they should exhaust explicit strategies before allowing the procedural system to dominate responding. On the other hand, individuals lower in WM, who have less capacity to test and switch among multiple rules, should instead be quicker to rely on the optimal

* Corresponding author. Tel.: +1 773 834 3713.
E-mail address: beilock@uchicago.edu (S.L. Beilock).

0010-0277/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2009.03.001

procedural system and thus faster to learn II category structures. As a result, less WM should support quicker learning on the II task – a counterintuitive prediction given the predominant view in the WM literature that greater capacity always translates into superior performance (Conway et al., 2005; Stanovich & West, 2000).

Recently, Tharp and Pickering (2009) called into question whether individuals actually use procedural learning when performing the II category task originally developed by Waldron and Ashby (2001) and used in DeCaro et al. (2008). Tharp and Pickering identified several simple explicit rules that can be used to classify the II stimuli correctly about 75% of the time. By modeling response strategies, Tharp and Pickering showed that individuals often utilize these simple rules and suggest that this rule use may lead to success on the II task. Tharp and Pickering support their claims by demonstrating that individuals able to reach 8 consecutive correct responses (CCR) – the learning criterion used by Waldron and Ashby (2001) and by DeCaro et al. (2008) – are often unable to maintain perfect accuracy across the following 8 trials (i.e., at a 16 CCR criterion). Thus, individuals could be attaining 8 CCR by using a simple rule that produces some success but falls short when higher levels of performance are required.

Tharp and Pickering's (2009) work leads to an interesting alternate explanation for our previous findings. Specifically, individuals lower in WM capacity, who do not have the capacity to support complex explicit hypothesis testing, may demonstrate quicker II category learning not because the procedural learning system dominates responding earlier but because these learners use simple explicit strategies. Such simple strategies are "good enough" to produce better-than-chance accuracy (e.g., Gigerenzer, 2007) but fall short when perfect or near-perfect performance is expected. Both simple rule-based and procedural strategies can be performed with relatively little demand on WM resources but differ qualitatively in their reliance on explicit versus implicit learning systems, respectively.

In the current work, we reassess the mechanism by which low WM individuals attain better II learning at 8 CCR in DeCaro et al. (2008). We do this by examining category learning at both 8 CCR and a more stringent 16 CCR criterion. If low WM individuals are using a simple strategy that circumvents complex hypothesis testing in WM but does not produce optimal category learning, then they should not excel on the II task in question at a stricter 16 CCR learning criterion. We then extend this performance data by directly modeling the response strategies individuals use to perform the II task.

2. Current experiment

2.1. Method

2.1.1. Participants

Undergraduates at Miami University scoring in the upper ($N = 16$) or lower ($N = 14$) quartiles on the average of two commonly used measures of WM capacity were selected to participate. These WM measures were administered in a previous session during the same semester as the current study. Although DeCaro et al. (2008) measured WM as a continuous variable, we chose to sample the extreme ends of the continuum so that we could explicitly focus on a contrast of the strategies used by those lowest and highest in WM.

WM composite scores were created for each individual by averaging absolute scores (see Conway et al., 2005) on the same two WM span tests used in DeCaro et al. (2008): the Automated Reading Span (ARspan) and Automated Operation Span (AOspan; Unsworth, Heitz, Schrock, & Engle, 2005). Cutoff scores for high and low WM groups (i.e., upper and lower quartile scores) were determined from a prior sample of individuals at the same university (high WM score range: 58–75; low WM score range: 0–33). Participants received either course credit or payment for participation.

2.1.2. Procedure

Individuals completed four different category learning sets, order counterbalanced across participants. These four sets consisted of two RB and two II category structures. The specific tasks and procedure were borrowed directly from Waldron and Ashby (2001) and were the same as those used in DeCaro et al. (2008), with one primary exception: we created 13 learning blocks, and within each block all 16 stimuli were randomly sampled without replacement (for a total of 208 trials maximum – see Supplementary material for more detail). We chose the sampling without replacement method in an attempt to maximally differentiate the response patterns that different decision strategies predict for each block over the course of learning.

3. Results

3.1. Category learning

We first examined rule-based and information-integration category learning as a function of WM group. We obtained four learning scores for each individual: the number of trials taken to reach 8 CCR and 16 CCR averaged across the two RB categories and the number of trials taken to reach 8 and 16 CCR averaged across the two II categories.

If an individual did not attain the learning criterion for a particular set, he or she received a score of 208 for that set (the maximum number of trials performed). This was done so that we could perform an assessment of response strategies (see below). An individual's data was removed from the dataset if the average score for a particular category structure was higher than 2 standard deviations from the mean of their WM group. The data from three individuals were excluded for the 8 CCR analyses (1 high WM, 2 low WM) and the data from three individuals were excluded for the 16 CCR analyses (1 high WM, 2 low WM).

3.1.1. 8 CCR

Scores for the 8 CCR criterion were analyzed using a 2 (WM group: low, high) $\times$ 2 (category structure: RB, II) mixed ANOVA. A main effect of category structure, $F(1,22) = 65.77$, $p < .001$, was qualified by a significant WM $\times$ category structure interaction, $F(1,22) = 4.11$, $p = .05$. As depicted in Fig. 1, for RB categories, high WM

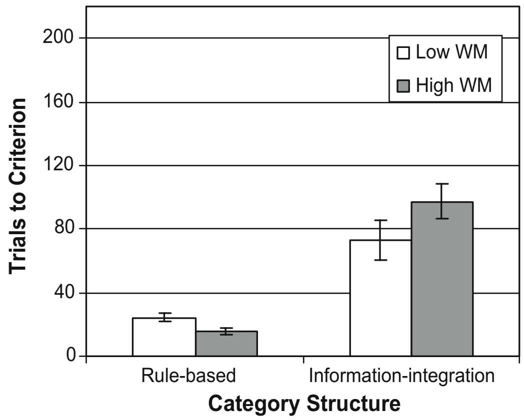

Fig. 1. Mean trials to reach a criterion of 8 consecutive correct responses (CCR) as a function of category structure and individual differences in working memory. Error bars represent standard errors.

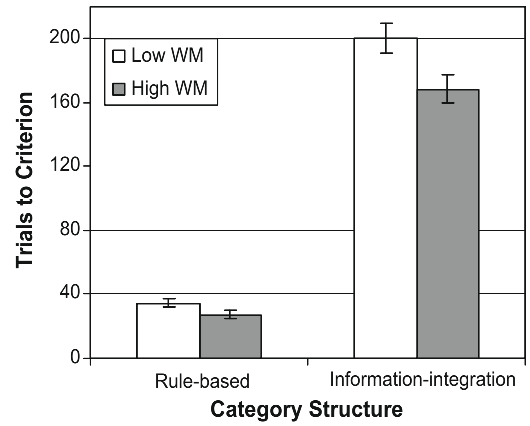

Fig. 2. Mean trials to reach a criterion of 16 consecutive correct responses (CCR) as a function of category structure and individual differences in working memory. Error bars represent standard errors.

individuals (high WMs) took fewer trials to reach the learning criterion (M = 15.97, SD = 5.26; CI: 11.87–20.06) than low WM individuals (low WMs; M = 24.42, SD = 9.97), $d = -1.10$. In contrast, high WMs required more trials to learn II categories to the 8 CCR criterion (M = 97.33, SD = 52.89; CI: 74.47–120.20) than low WMs (M = 73.25, SD = 26.36), $d = .56$.

With a criterion of 8 CCR, we replicate DeCaro et al.'s (2008) findings. More WM capacity leads to faster learning of RB categories thought to rely heavily on explicit hypothesis testing (Ashby & Maddox, 2005). In contrast, less WM leads to faster learning of II categories. By demonstrating this "less-is-more" effect for II learning holds for the current sample, we can now examine the potential constraints on this finding. Specifically, will low WMs fall short when the learning criterion is made more stringent (e.g., 16 CCR)? And what might a closer examination of response strategies tell us about the mechanism(s) underlying category learning as a function of individual differences in WM?

3.1.2. 16 TTC

We next examined the impact of imposing the more stringent 16 CCR learning criterion on performance outcomes. The number of trials taken to reach 16 CCR was analyzed in a 2 (WM group: low, high) $\times$ 2 (category structure: RB, II) mixed ANOVA, revealing only main effects of category structure, $F(1,25) = 565.30$, $p < .001$, and WM group, $F(1,25) = 8.16$, $p < .01$. As seen in Fig. 2, for the RB task, high WMs (M = 27.40, SD = 7.77; CI: 22.08–32.72) outperformed low WMs (M = 34.54, SD = 12.27), $d = -.71$. This is consistent with our findings at 8 CCR. However, in contrast to the 8 CCR findings, at 16 CCR high WMs (M = 168.43, SD = 42.97; CI: 150.90–185.96) also learned the II categories faster than low WMs (M = 200.58, SD = 10.93), $d = -.97$.

This reversal in II learning at 8 versus 16 CCR as a function of WM group aligns with Tharp and Pickering's suggestion that learning to 8 CCR on Waldron and Ashby's (2001) task may not always reflect reliance on a stable proceduralized learning system. Otherwise, low WMs' learning advantage at 8 CCR would remain at 16 CCR. But, it did not. How, then, are lower (and higher) WM individuals learning the II category task?

From an inspection of Figs. 1 and 2, one can see that low WMs' performance changes more than high WMs' performance as a function of the specific learning criterion we used to analyze the data. If the extremity of the difference in II performance across 8 CCR and 16 CCR depends on WM, then this might reflect a fundamental difference in the learning strategies adopted by these groups. Although one would still need to determine what these strategies are, showing that changing how a performance criterion is used to indicate learning impacts low WMs more than high WMs would suggest that the former may be adopting a learning strategy that is effective in terms of quickly reaching 8 CCR, but fails to produce stable learning needed to reach a 16 CCR criteria. In contrast, high WMs may be more likely to rely on a learning strategy that leads to success when a stricter 16 CCR learning criteria is imposed.

A 2 (WM group: low, high) $\times$ 2 (criterion: 8 CCR, 16 CCR) ANOVA on II learning revealed a significant WM $\times$ criterion interaction, $F(1,24) = 16.22$, $p < .001$. This confirms that category learning performance was more dependent on the learning criterion we used to analyze the data for low than high WM individuals. To explore how low and high WMs differ in terms of the strategies they use to perform the II task, we next modeled individuals' response strategies over the course of learning.

3.2. Response strategies for the information-integration task

When setting out to model response strategies on the II task, we noticed that explicit rules could potentially be instantiated to categorize these stimuli at least 75% of the time (see also Tharp & Pickering, 2009). Specifically, the II tasks could be categorized with rules involving one, two or three-dimensions, at 75–87.5% accuracy (see Supplementary material for details). Thus, it is conceivable that one could master, to a fairly high level of accuracy, an II category structure when actually utilizing a simple explicit rule – particularly when a less stringent learning criterion is imposed (e.g., 8 CCR). Using the optimal strategy (i.e., categorizing the stimuli exactly the way they were set up by the experimenter) of course would lead to 100% accuracy.

For each 16-trial block we assessed the strategy most likely used by an individual (for response modeling details, see Supplementary material). We found that strategy sampling differed over time as a function of WM group. This was confirmed by a 2 (WM group: low, high) × 4 (strategy type: optimal, one-dimension, two-dimension, three-dimension) × 13 (block) interaction, $F(36,900) = 1.46$, $p = .04$. As one can see in Fig. 3, across all learning blocks, low WMs favored one-dimension rules above all other strategies. High WMs also employed one-dimension rules, yet to a lesser degree than low WM individuals – switching instead among several different strategies.

Of note is the pattern of optimal strategy selection over time. As shown in the leftmost column of Fig. 3, high WMs incrementally increased their use of the optimal strategy across learning blocks, whereas low WMs did not. This observation was confirmed by the fact that optimal strategy use over time was significantly fit by a linear function for high WMs, $F(1,14) = 13.03$, $p < .01$, but not low WMs, $F(1,11) = 3.32$, $p > .05$. This finding is consistent with the notion that high WMs gradually accrue stimulus-response associations (indicative of a procedural-based learning system) that guide responding more than lower WMs. However, as Tharp and Pickering (2009) suggest and as we discuss below, these data cannot sufficiently rule out the possibility that, instead of a procedural classification strategy, high WMs may be using extremely complex explicit rules to learn the II category structures we employed (but see Ashby, Alfonso-Reese, Turken, & Waldron, 1998).

These strategy trends can also be seen in Table 1, which displays proportions of strategy use collapsed across all 13 blocks for low and high WMs. One should note that, because individuals almost always begin II tasks with explicit rule use and gradually implement the optimal strategy over time (Ashby et al., 1998), explicit rules will constitute a large proportion of learning trials across all individuals when collapsing across all 13 blocks. Nonetheless, submitting these data to a 2 (WM group) × 4 (strategy type) ANOVA revealed a significant main effect of strategy type, $F(3,75) = 28.04$, $p < .001$, qualified by a WM × strategy type interaction, $F(3,75) = 4.62$, $p < .01$. Using 95% confidence intervals (CIs), low WMs were more likely to use one-dimension rules than high WMs (CI: 0.30–0.43), $d = -1.24$. In contrast, high WMs were overall more likely to use three-dimension rules (CI: 0.20–0.26), $d = .72$, and

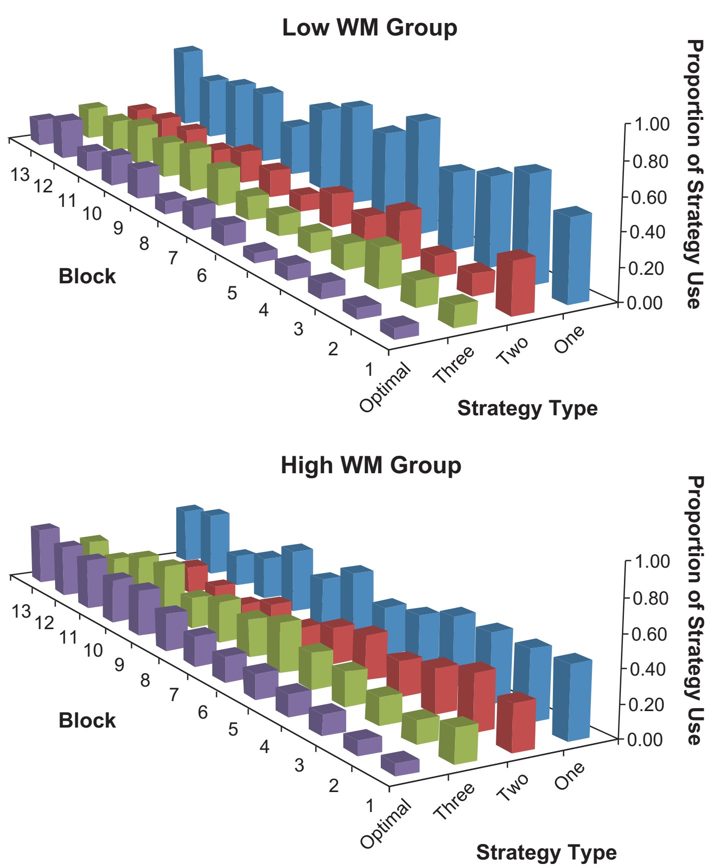

Fig. 3. Proportion of strategy use (optimal, three-dimension rule, two-dimension rule, and one-dimension rule) by strategy type and block for low and high working memory groups.

Table 1
Overall proportions of strategy use for low and high WM groups. Standard errors are in parentheses.

	Strategy type			
	Optimal	One-dimension	Two-dimension	Three-dimension
Low WM	.13 (.04)	.51 (.04)	.18 (.02)	.19 (.02)
High WM	.22 (.04)	.36 (.03)	.19 (.02)	.23 (.01)

the optimal strategy (CI: 0.15–0.29), $d = .66$, than low WMs. The groups did not differ in their use of two-dimension rules.

3.3. Response strategies and category learning

We next asked how these response strategies relate to the II learning scores reported above. Because we were specifically interested in exploring how strategy use leads low WMs to perform better than high WMs at 8 CCR but worse at 16 CCR, we correlated the number of trials to reach each criterion with the proportion of strategy use across the first 5 blocks (i.e., the first 80 trials). This is the point at which individuals generally reached 8 CCR (see Fig. 1), and thus modeling this data provides insight into how low WMs may achieve early success on the II task. As shown in Table 2, the more low WMs used the one-dimension rule, the fewer trials they took to reach 8 CCR. In addition, the more low WMs tried a three-dimension rule, the worse they performed in terms of reaching the 8 CCR criterion. High WMs did not show these correlations. Instead, performance tended to improve as high WMs used the optimal strategy.

Two key conclusions can be drawn from our data: (1) Low WMs depend heavily on simple one-dimension rules throughout learning, relying little on either multi-dimensional rules or the optimal II learning strategy. By perseverating on these simple rules despite receiving error feedback, low WMs are quicker to attain 8 CCR than high WMs. (2) High WMs also initially employ a variety of explicit rules, but this type of rule use does not relate to early learning because high WMs are less likely to stick with one-dimension rules that lead to early success. Unlike low WMs, high WMs incrementally increase their use of the optimal classification strategy across time. Compared to using a one-dimension "shortcut", such complex strategy selection slows accurate performance early on, but it generally proves successful overall.

Table 2
Correlations between the number of trials taken to reach 8 and 16 consecutive correct responses (CCR) and proportions of strategy use in blocks 1–5.

	Strategy type			
	Optimal	Three-dimension	Two-dimension	One-dimension
Low WM group				
8 CCR	−.27	.71**	.53[a]	−.65*
16 CCR	−.17	−.09	.08	.11
High WM group				
8 CCR	−.49[b]	−.18	.33	.22
16 CCR	−.75**	−.23	.61*	.23

Note: Positive correlations indicate worse performance (i.e., greater use of a given strategy correlated with more trials taken to reach criterion).
[a] $p < .08$.
[b] $p < .07$.
* $p < .05$.
** $p < .01$.

4. Discussion

4.1. Re-conceptualizing II task strategies for low WMs

We demonstrate here, as in our previous work (DeCaro et al., 2008), that high WMs take longer than low WMs to attain 8 CCR on Waldron and Ashby's (2001) II task. Using the COVIS model of category learning as a guide (Ashby & Maddox, 2005), we had previously reasoned that, because high WMs are better able than low WMs to hypothesis test and keep multiple pieces of information active in WM, high WM individuals would be more likely to test complex rules for a longer period of time. We hypothesized that this strategy selection would result in slower learning of II category structures based on a proceduralized learning system whose efficacy is hindered by the use of WM-dependent strategies (Maddox, Love, Glass, & Filoteo, 2008). In other words, we reasoned that less WM might be better than more during II category learning.

In some ways, we did find that less could be more in the current work. High WMs – at the outset of learning – tested both simple and complex explicit categorization rules that did not allow for the quick attainment (8 CCR) of II learning. Low WMs perseverated on simple one-dimension rules that actually helped them obtain early success. This type of perseveration is consistent with the response perseveration often seen with older adults (e.g., Raz, Gunning-Dixon, Head, Dupuis, & Acker, 1998) and patients with prefrontal lobe damage (Shallice & Burgess, 1991; Shimamura, 2000) typically presenting with lower WM. Yet, such simple strategy perseveration was not beneficial for meeting the 16 CCR criterions. Low WMs seem to have learned a "good enough" strategy for II category structures that worked only when a less stringent learning criterion was imposed. This is in contrast to higher WMs who eventually adopted an optimal classification response that aided performance using a 16 CCR criterion. Thus, in other ways, less is not more.

Compared to other II category learning tasks, the Waldron and Ashby (2001) II task we used has the *potentially* unique characteristic that simple explicit rules can be adopted for above-chance accuracy. From our findings it appears that if a task can be sufficiently performed using a simple explicit strategy, some people (i.e., low WMs) will likely opt for this strategy above all else. This is consistent with individual difference research in other domains, such as insight problem solving (Beilock & DeCaro, 2007; Ricks, Turley-Ames, & Wiley, 2007), correlation detection (Gaissmaier, Schooler, & Rieskamp, 2006), and probability matching (Wolford, Newman, Miller, & Wig, 2004). Here, individuals lower in executive function abilities have been shown to capitalize on simpler ways to solve problems than their higher ability counterparts (see DeCaro & Beilock, in press, for a review). For example, in Beilock and DeCaro (2007), we found that lower WMs were more likely than higher WMs to abandon a complex multi-step math problem solving strategy in favor of a simple one-step strategy that also led to accurate performance. The tendency for low WMs in the current work to rely on a one-dimension rule that leads to above-chance accuracy and early success is consistent with this other research.

Of course, it is an open question as to what the strategy use of lower WMs will look like when a simple rule does not lead to early success. We speculate that it may still be the case that low WMs will be quicker to rely on procedural-based responses than high WMs if the only other option is complex explicit hypothesis testing heavily dependent on WM. That is, if a simple explicit rule does not lead to above-chance performance and early success, low WMs may instead rely on a different strategy that skirts the burden on WM resources: procedural-based responding. Although this is certainly a question for future work, findings derived from studies using continuous-dimension stimuli (e.g., Gabor patches) align with this idea. For instance, Markman, Maddox, and Worthy (2006) demonstrated that a distracting performance environment improved II category learning. The distracting environment Markman et al. employed has been shown, across a number of different studies, to reduce the WM available for the task at hand (see Beilock, 2008 for a review). Maddox et al. (2008) offer confirmatory evidence for this finding, showing that II learning is impaired when the attention devoted to performance is increased.

Thus, when one cannot turn to a simple one-dimension rule for above-chance performance, low WMs (who come to the table in what might be described as a reduced capacity state to begin with) may rely on a procedural-based system earlier than their higher WM counterparts. Future work using alternate II category tasks will help to address this possibility. However, other II category structures (such as the continuous-dimension stimuli mentioned above) may also afford above-chance accuracy for those who perseverate on simple one-dimension explicit rules even in the face of negative feedback. In other words, uni-dimensional rule selection may be a factor in all II learning tasks, not just with the Waldron and Ashby task (cf. Zeithamova & Maddox, 2006). Without employing individual difference measures in category learning, it may be hard to ascertain the extent of this rule use in any II task.

4.2. II Task strategies of high WMs

Although we show that low WM individuals have a tendency to rely on explicit one-dimension rules throughout II category learning, we also found that those higher in WM incrementally increase their use of optimal response strategies over time. We believe that these optimal strategies reflect reliance on a procedural-based learning system because of the gradual adoption of these strategies shown in the current work and because of the limited dual-task interference found in other work (Waldron & Ashby, 2001; Zeithamova & Maddox, 2006; see also Ashby et al., 1998). If so, then studies using the category learning task employed in the current work may still be assessing II learning for some individuals – especially when high-achieving college students with above-average WM capacities are used as participants.

However, one limitation of this particular II task is that response modeling cannot truly differentiate a procedural strategy from a complex multi-dimensional rule that yields an identical response pattern. Thus, our data cannot speak directly to whether the optimal strategy high WMs ultimately adopt to classify the II categories is reflective of a complex multi-dimensional rule or a procedural strategy. Other converging operations, such as dual-task interventions with individual difference measures, may provide additional insight into the actual strategy used.

5. Conclusion

We show that lower WM individuals rely on simple one-dimension explicit rules to perform the II task developed by Waldron and Ashby (2001) and employed by DeCaro et al. (2008). And given that a simple rule can lead to early success on the II task in question (i.e., above-chance accuracy and quicker achievement of an 8 CCR criterion), we show when less is better than more. Importantly, our interpretation of how less is more has changed, at least with respect to the II category learning task used in DeCaro et al. (2008). Rather than low WMs' early success being due to a reliance on procedural-based responding, it appears to be due to a heavier reliance on simple one-dimension rules. Although such rules produce above-chance accuracy and success when an 8 CCR criterion is imposed, such shortcut strategies fall short in the current category learning paradigm when higher performance levels are required. The same category learning tasks can be approached very differently as a function of individual differences in WM, with those lower in WM capacity opting for strategies that make the least demands on executive control. These findings reinforce the need to consider individual differences in general cognitive capacities both in terms of the strategies individuals adopt for performance and the implications these strategies carry for cognitive and neural mechanisms driving learning.

Acknowledgements

This research was supported by NSF CAREER Grant No. DRL-0746970 to Sian L. Beilock and NSF Grant No. BCS-0544688 to Robin D. Thomas.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cognition.2009.03.001.

References

Ashby, E. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review, 105*, 442–481.

Ashby, F. G., & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology, 56*, 149–178.

Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science, 17*, 339–343.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 983–998.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786.

DeCaro, M. S., & Beilock, S. L. (in press). The benefits and perils of attentional control. In M. Csikszentmihalyi, B. Bruya (Eds.), *Effortless attention: A new perspective in the cognitive science of attention and action*. MIT Press.

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

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

Gaissmaier, W., Schooler, L. J., & Rieskamp, J. (2006). Simple predictions fueled by capacity limitations: When are they successful? *Journal of Experimental Psychology: Learning, Memory & Cognition, 32*, 966–982.

Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. London: Penguin Books.

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

Markman, A. B., Maddox, W. T., & Worthy, D. A. (2006). Choking and excelling under pressure. *Psychological Science, 17*, 944–948.

Raz, N., Gunning-Dixon, F. M., Head, D., Dupuis, J. H., & Acker, J. D. (1998). Neuroanatomical correlates of cognitive aging: Evidence from structural magnetic resonance imaging. *Neuropsychology, 12*(1), 95–114.

Ricks, T. R., Turley-Ames, K. J., & Wiley, J. (2007). Effects of working memory capacity on mental set due to domain knowledge. *Memory & Cognition, 35*, 1456–1462.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*, 727–741.

Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology, 28*, 207–218.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*, 645–726.

Tharp, I. J., & Pickering, A. D. (2009). A note on DeCaro, Thomas, and Beilock (2008): Further data demonstrate complexities in the assessment of information-integration category learning. *Cognition, 111*, 411–415.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37*, 498–505.

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. *Psychonomic Bulletin & Review, 8*, 168–176.

Wolford, G., Newman, S. E., Miller, M. B., & Wig, G. S. (2004). Searching for patterns in random sequences. *Canadian Journal of Psychology, 584*, 221–228.

Zeithamova, D., & Maddox, W. T. (2006). Dual-task interference in perceptual category learning. *Memory & Cognition, 34*, 387–398.