Research Report

Choking and Excelling Under Pressure

Arthur B. Markman, W. Todd Maddox, and Darrell A. Worthy

University of Texas at Austin

ABSTRACT—*Decrements in performance on cognitive tasks resulting from pressure to perform (i.e., choking) are thought to be caused by interference with the ability to use explicit strategies (the distraction theory). This view suggests that pressure should improve performance on tasks for which explicit strategies hamper performance. This hypothesis was tested by giving subjects one of two nearly identical learning tasks, a task that required learning a rule or one that required using a holistic information-integration strategy. Explicit rule use would hurt performance in the latter task. As predicted by the distraction theory, pressure decreased performance on the rule-based task but enhanced performance on the information-integration task.*

Psychological research provides evidence for the anecdotal phenomenon that pressure causes decrements in performance on cognitive and motor tasks (e.g., Beilock & Carr, 2005; Beilock, Kulp, Holt, & Carr, 2004; Masters, 1992). This phenomenon is often called *choking under pressure*, because people become unable to perform a task that they would otherwise perform well.

A prominent explanation for choking under pressure in cognitive tasks is the *distraction hypothesis* (Beilock & Carr, 2005; Beilock et al., 2004). In this view, pressure leads to a decrease in available working memory resources, which in turn has a negative influence on cognitive performance. In a study supporting this view, for example, Beilock and Carr (2005) found that individuals with high working memory capacity were more strongly affected by a pressure manipulation than were those with low working memory capacity.

An interesting implication of this distraction hypothesis is that pressure might enhance performance on tasks for which performance is harmed by interference from processes that rely intensively on working memory. For example, Figure 1 shows the stimulus structure for two simple categorization tasks. The stimuli for each task can be described by two dimensions (e.g., the width and orientation of sine-wave gradients). Figure 1a depicts a simple rule-based task in which the participant must focus on one of the two dimensions and determine the location on that dimension that separates the two categories. This task is typically thought to involve explicit hypothesis testing, and so it should be harmed by any procedure that decreases working memory capacity (Ashby, Alfonso-Reese, Turken, & Waldron, 1998; Maddox & Ashby, 2004; Maddox, Filoteo, Hejl, & Ing, 2004). In research consistent with this proposal Zeithamova and Maddox (2006; see also Maddox, Ashby, Ing, & Pickering, 2004; Waldron & Ashby, 2001) showed that learning of this particular category structure was impaired in a dual-task setting relative to a control condition in which a second task did not have to be performed.

In contrast, the stimulus structure in Figure 1b rotates the category boundary 45° in stimulus space. Thus, the rule that separates the categories cannot be stated easily. This stimulus configuration is called an *information-integration* structure and is thought to be best learned by a procedural or similarity-based process that is limited in its demands on working memory (Maddox & Ashby, 2004; Maddox, Ashby, & Bohil, 2003). Explicit hypothesis-testing strategies can be used to solve such a task, but they lead to suboptimal performance. Even so, previous research suggests that early in learning, people start by using hypothesis-testing strategies even in information-integration tasks (Ashby et al., 1998). Thus, if people's working memory capacity were disrupted by pressure, they might abandon hypothesis-testing strategies more quickly than usual and thereby improve their ability to acquire the information-integration structure. That is, people would excel under pressure in this condition.

We tested this hypothesis in the present study. The manipulation of pressure was similar to that used by Beilock and Carr (2005). In their study, subjects in the low-pressure condition were told that they were going to perform a category-learning

Address correspondence to Arthur B. Markman, Department of Psychology, University of Texas, 1 University Station, A8000, Austin, TX 78712, e-mail: markman@psy.utexas.edu.

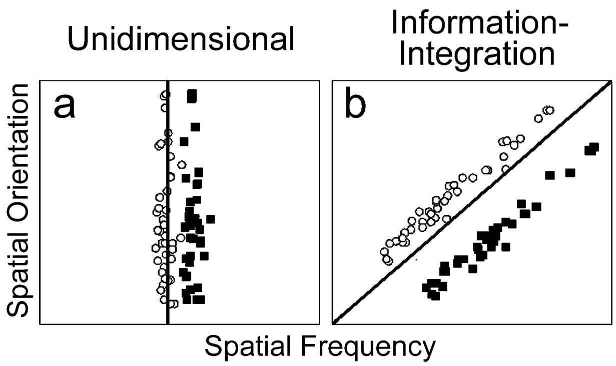

Fig. 1. The two category structures used in the experiment: (a) a unidimensional rule-based structure and (b) a two-dimensional information-integration structure.

task and that they should do their best. Subjects in the high-pressure condition were told that they and a partner would each receive a performance bonus of \$6 in the study if they both exceeded a performance criterion, but that neither would receive the bonus if one of them failed to exceed the criterion. Furthermore, they were told that their partner had already exceeded the criterion, so the bonus was contingent on their own performance. Beilock and Carr found that the high-pressure group performed reliably worse than the low-pressure group on a mathematics task. The distraction view predicts that, relative to subjects in the low-pressure condition, subjects in the high-pressure condition should have a more difficult time learning the rule-based categories shown in Figure 1a, but an easier time learning the information-integration categories shown in Figure 1b.

METHOD

Subjects were 80 members of the University of Texas community who received course credit for their time. They were randomly assigned to one of four between-subjects conditions defined by the factorial combination of two levels of pressure (low vs. high) and two category types (rule-based categories vs. information-integration categories). Pressure level was manipulated through the task instructions.

The low-pressure group was asked to do their best. The high-pressure group was told that they would receive a monetary bonus if both they and a (fictional) partner exceeded a performance criterion (80% accuracy over the last 80 trials of the study), and that neither would receive a bonus if one failed to reach the criterion. Each high-pressure participant was then informed that his or her partner had already exceeded the criterion, so the bonus for the participant and the partner rested on the participant's performance.

The two category types were the unidimensional rule-based and information-integration structures shown in Figure 1. Stimuli were Gabor patches varying in frequency (spacing of bars) and orientation of the grating relative to the *x*-axis of the computer screen. The category structures were developed by finding stimulus sets that equated performance accuracy in a low-pressure condition run on other subjects.

Subjects performed eight blocks of 80 learning trials. On each trial, they saw an item presented on a computer and pressed a key to indicate the category to which they thought it belonged. They received corrective feedback. Two points were added to a point meter on the screen when subjects were correct, and zero points were added when subjects were incorrect. The counter was reset to zero at the start of each block. A line at 128 points (80% accuracy) indicated the level required for the performance bonus.

RESULTS

Figure 2 graphs the mean accuracy for each block in each condition. The accuracy data were subjected to a 2 (pressure level) × 2 (category type) × 8 (block) analysis of variance. As expected, there was a significant Pressure Level × Category Type interaction, $F(1, 76) = 7.08$, $p_{\text{rep}} = .96$, $\eta^2 = .09$. There was also a significant main effect of block, $F(7, 532) = 14.09$, $p_{\text{rep}} > .99$, $\eta^2 = .16$, as well as a significant Block × Pressure Level interaction, $F(7, 532) = 2.17$, $p_{\text{rep}} = .90$, $\eta^2 = .03$.

To illuminate these interactions, we contrasted the performance of subjects in the low- and high-pressure conditions in each block, separately for the two category structures. As expected, subjects learning rule-based categories had higher accuracy on average in the low-pressure condition than in the high-pressure

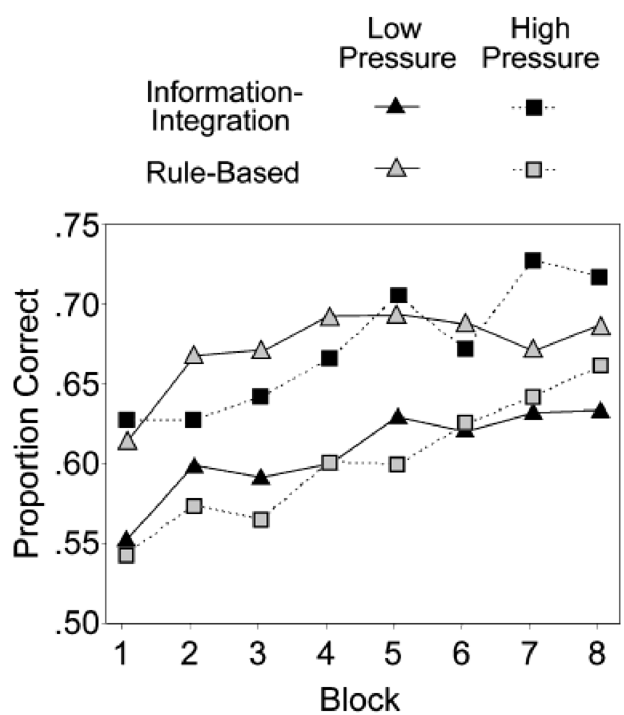

Fig. 2. Mean accuracy for the two category structures in the low- and high-pressure conditions, as a function of block.

condition in each block of the experiment (p_{rep} = .98 by sign test). Planned contrasts revealed that the difference in accuracy was statistically significant in Blocks 3 and 4. Also as predicted, subjects learning information-integration categories showed the opposite pattern: On average, accuracy in the high-pressure condition was higher than accuracy in the low-pressure condition in each block of the experiment (p_{rep} = .98 by sign test). Planned contrasts revealed that the difference was statistically significant in Blocks 1, 5, and 7.

One advantage of using simple stimuli was that we could fit models to individual subjects' data to determine the strategy they used to perform the task. We fit decision-bound models to the data from individual subjects on a block-by-block basis (Maddox, 1999; Maddox & Ashby, 1993). Four models were fit. One was a two-parameter rule-based model that assumed a unidimensional decision bound along the spatial-frequency dimension (another variant assumed a decision bound along the spatial-orientation dimension). The location of the decision bound was a free parameter, and a noise parameter represented variability in trial-by-trial memory for, and application of, the decision bound. The second model was a three-parameter rule-based model that assumed a conjunctive strategy. The model assumed one decision bound along the spatial-frequency dimension and a second along the spatial-orientation dimension, and the decision rule was as follows: Respond "A" if the spatial frequency is low and the orientation is steep; otherwise, respond "B." A second variant assumed a different rule: Respond "B" if the spatial frequency is high and the orientation is low; otherwise, respond "A." The conjunctive model was applied only to the data from the information-integration condition. The third model was a three-parameter information-integration model that assumed a linear decision bound. The slope and the intercept of this boundary were free parameters, along with the noise parameter outlined earlier. The fourth model was a one-parameter guessing model that assumed that the probability of responding "A" (a free parameter in the model) was not affected by the location of the stimulus in space.

Akaike's (1974) information criterion (AIC) was used to determine the model that provided the best account of the data. AIC penalizes a model for each free parameter, and thus a model with fewer parameters can provide a better account of the data than a model with more free parameters. Table 1 shows the proportion of subjects whose data were best fit by the rule-based models and by the information-integration model, separately for each block of the study.

Previous research suggests that early in learning, people start by using a rule-based strategy (even in information-integration tasks; Ashby et al., 1998). Thus, the poor performance of subjects in the high-pressure, rule-based condition should have been caused by difficulty applying the rule, rather than by a shift in strategy toward information integration. Figure 3a shows the proportion of subjects learning rule-based categories whose data were best fit by a rule-based model in each block of trials. These data suggest that there was a tendency for the rule-based model to fit the data from the low-pressure condition better than the data from the high-pressure condition. The model fits also indicate relatively little change in strategy use over the course of the study. Across the first six blocks, the proportion of subjects learning rule-based categories whose data were fit by a rule-based model was higher in the low-pressure condition than in the high-pressure condition; the proportion of subjects fit by a rule-based model was about equal for the two conditions in the last two blocks (p_{rep} = .77 by sign test).

TABLE 1

Proportion of Subjects in Each Condition Whose Data Were Best Fit by Rule-Based Versus Information-Integration Models

Model	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Mean
Rule-based categories: low pressure									
Rule-based	.45	.50	.55	.55	.70	.60	.50	.50	.54
Information-integration	.30	.20	.20	.40	.10	.15	.20	.25	.23
Rule-based categories: high pressure									
Rule-based	.45	.40	.35	.45	.40	.35	.55	.60	.44
Information-integration	.20	.45	.15	.15	.25	.35	.20	.20	.24
Information-integration categories: low pressure									
Rule-based	.45	.45	.55	.45	.50	.40	.20	.35	.42
Information-integration	.15	.25	.25	.40	.30	.25	.50	.40	.31
Information-integration categories: high pressure									
Rule-based	.50	.45	.65	.50	.35	.30	.30	.40	.43
Information-integration	.35	.40	.35	.45	.60	.60	.65	.60	.50

Note. Subjects not fit by either a rule-based or an information-integration model were best fit by a model assuming that they guessed randomly on each trial.

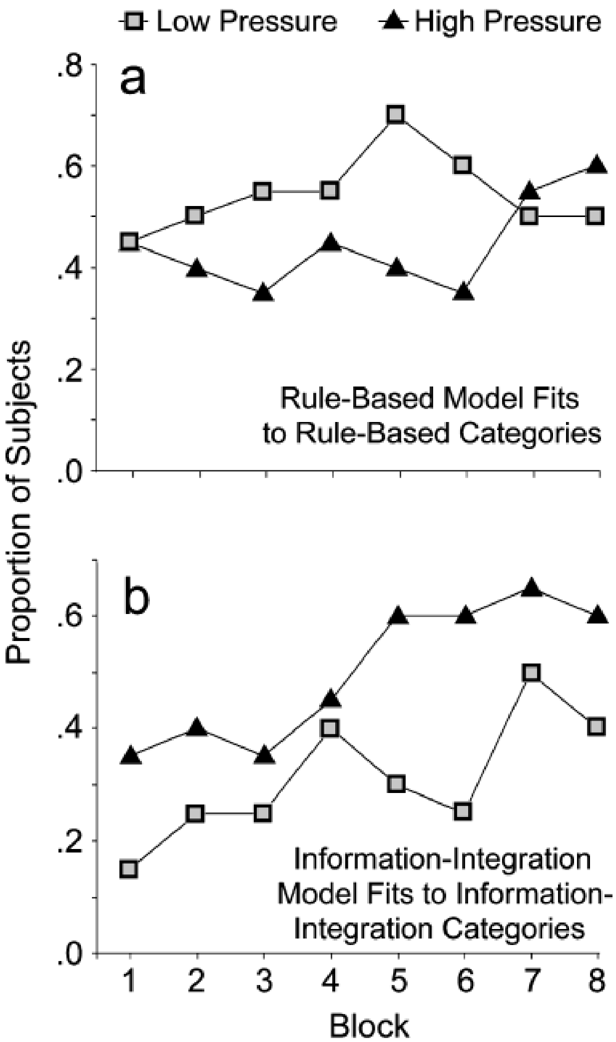

Fig. 3. (a) Proportion of subjects learning rule-based categories whose data were best fit by a rule-based model and (b) proportion of subjects learning information-integration categories whose data were best fit by an information-integration model.

Most subjects learning rule-based categories who were not fit by the rule-based model were fit best by the guessing model, which suggests that they were struggling to learn the categories. Results from a regression analysis were consistent with this interpretation: In regression models predicting the proportion of subjects best fit by a rule-based model as a function of block, the linear trend was not significant for either the low-pressure or the high-pressure condition (both p_{rep}s $< .81$).

This pattern contrasts with the model-based results for the information-integration categories. As shown in Figure 3b, data from subjects learning information-integration categories were better fit with an information-integration model when subjects were in the high-pressure condition than when they were in the low-pressure condition ($p_{\mathrm{rep}} = .98$ by sign test). There was a tendency in both conditions for subjects to increase their use of an information-integration strategy over the course of the study. Consistent with this interpretation, the linear trend was significant for both the low-pressure condition, $F(1, 7) = 7.76$, $p_{\mathrm{rep}} = .90$, and the high-pressure condition, $F(1, 7) = 27.98$, $p_{\mathrm{rep}} = .97$. Finally, Figure 3b indicates that the data from subjects in the high-pressure condition were best fit by an information-integration model (rather than a rule-based or guessing model) earlier in the study than were the data from subjects in the low-pressure condition.

DISCUSSION

These data are consistent with previous findings that pressure causes decrements in cognitive performance (i.e., choking) because it reduces working memory available for processing (Beilock & Carr, 2005; Beilock et al., 2004). Perhaps counterintuitively, the hypothesized decrement in working memory capacity actually improves performance on an information-integration classification task. We suggest that this improvement occurs because people with reduced working memory capacity are less likely to use and maintain a suboptimal rule-based hypothesis-testing strategy for learning the category distinction. Instead, they allow their performance to be driven by an implicit similarity-based process.

Another prominent explanation for choking phenomena is the *explicit-monitoring* view, which suggests that pressure may influence performance of highly learned skills by causing people to engage explicit processes that interfere with carrying out the learned procedure (e.g., Gray, 2004; Masters, 1992). This view seems less relevant to perceptual classification tasks, which focus on acquisition of representations rather than use of an expert skill. However, future research should explore the possibility that when information-integration categories are overlearned, there may be a decrement in classification accuracy because of explicit monitoring.

The present findings are consistent with previous research indicating that there are (at least) two learning systems responsible for perceptual classification. The explicit-rule-based system is affected by manipulations that influence working memory (Maddox, Ashby, et al., 2004; Waldron & Ashby, 2001; Zeithamova & Maddox, 2006). In contrast, the implicit-similarity-based system connects stimuli to particular motor responses, and so it is influenced by manipulations of the motor actions used to respond in the task, as well as by delays between response and feedback that prevent the implicit system from learning the contingency between stimulus and response (Ashby, Ell, & Waldron, 2003; Maddox et al., 2003; Maddox & Ing, 2005).

This work also demonstrates the value of examining motivational variables using cognitive tasks for which models can be fit to the performance of individual subjects (Maddox, Baldwin, & Markman, in press; Markman, Baldwin, & Maddox, 2005). Research on motivation often focuses on tasks for which only gross measures of performance, such as accuracy, are available.

Although accuracy is an important and useful measure, the use of perceptual classification tasks permits one to identify more subtle changes in people's behavior across blocks, and consequently one can address finer-grained aspects of the influence of incentives in performance.

Finally, these findings suggest an important avenue for future research. The performance incentive used in the experiment reported here is similar to incentives used in tasks of regulatory focus (Crowe & Higgins, 1997; Higgins, 1997; Markman et al., 2005). The high-pressure condition involved a potential gain in the form of a monetary bonus. Such a gain may induce a *promotion focus*, which is a general sensitivity to gains and nongains in the environment. The manipulation also induced the potential for a negative social outcome if the participant let down his or her partner. This aspect of the situation may have induced a *prevention focus*, which is a general sensitivity to losses and nonlosses. Because the task had elements of both a promotion focus and a prevention focus, there was a possible regulatory mismatch. We suggest that regulatory mismatches generally lead to inflexible cognitive processing, which may have contributed to the poor performance observed in the rule-based task in the present study (Maddox, Markman, & Baldwin, 2007). Future research should address this possibility.

Acknowledgments—This research was supported by National Institutes of Health (NIH) Grant 1 R21 DA015211-01A1 and a fellowship at the IC^2 institute to A.B.M. and by NIH Grant R01 MH59196 to W.T.M. The authors thank Scott Lauritzen for his help running subjects on this project.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723.

Ashby, F.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., Ell, S.W., & Waldron, E.M. (2003). Procedural learning in perceptual categorization. *Memory & Cognition*, *31*, 1114–1125.

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.

Crowe, E., & Higgins, E.T. (1997). Regulatory focus and strategic inclinations: Promotion and prevention in decision-making. *Organizational Behavior and Human Decision Processes*, *69*, 117–132.

Gray, R. (2004). Attending to the execution of a complex sensorimotor skill: Expertise differences, choking, and slumps. *Journal of Experimental Psychology: Applied*, *10*, 42–54.

Higgins, E.T. (1997). Beyond pleasure and pain. *American Psychologist*, *52*, 1280–1300.

Maddox, W.T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception & Psychophysics*, *61*, 354–374.

Maddox, W.T., & Ashby, F.G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*, 49–70.

Maddox, W.T., & Ashby, F.G. (2004). Dissociating explicit and procedure-learning based systems of perceptual category learning. *Behavioral Processes*, *66*, 309–332.

Maddox, W.T., Ashby, F.G., & Bohil, C.J. (2003). Delayed feedback effects on rule-based and information-integration category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 650–662.

Maddox, W.T., Ashby, F.G., Ing, A.D., & Pickering, A. (2004). Disrupting feedback processing interferes with rule-based but not information-integration category learning. *Memory & Cognition*, *32*, 582–591.

Maddox, W.T., Baldwin, G.C., & Markman, A.B. (in press). A test of the regulatory fit hypothesis in perceptual classification learning. *Memory & Cognition*.

Maddox, W.T., Filoteo, J.V., Hejl, K.D., & Ing, A.D. (2004). Category number impacts rule-based but not information-integration category learning: Further evidence for dissociable category-learning systems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 227–245.

Maddox, W.T., & Ing, A.D. (2005). Delayed feedback disrupts information-integration but not rule-based category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 100–107.

Maddox, W.T., Markman, A.B., & Baldwin, G.C. (2007). Using classification to understand the motivation-learning interface. *Psychology of Learning and Motivation*, *47*, 213–250.

Markman, A.B., Baldwin, G.C., & Maddox, W.T. (2005). The interaction of payoff structure and regulatory focus in classification. *Psychological Science*, *16*, 852–855.

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

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

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

(RECEIVED 11/7/05; REVISION ACCEPTED 2/17/06;
FINAL MATERIALS RECEIVED 2/24/06)