

# Increased Practice with “Set” Problems Hinders Performance on the Water Jar Task

Noelle M. Crooks ([ncrooks@nd.edu](mailto:ncrooks@nd.edu))  
University of Notre Dame  
Department of Psychology, 118 Haggard Hall  
Notre Dame, IN 46556 USA

Nicole M. McNeil ([nmcneil@nd.edu](mailto:nmcneil@nd.edu))  
University of Notre Dame  
Department of Psychology, 118 Haggard Hall  
Notre Dame, IN 46556 USA

## Abstract

Luchins’s (1942) classic Einstellung (mental set) phenomenon has been demonstrated across a wide variety of samples and problem solving tasks. However, it is unclear how increased practice with the initial “set” affects subsequent performance. Although Luchins anecdotally reported no effect of increased practice with set problems, current theories would suggest otherwise. In this study, we varied the number of set problems (0, 1, 5, 10, 20, 40) in the water jar task and examined the effects on subsequent performance. Results suggest that problem-solving flexibility decreases linearly as the number of set problems increases. Contrary to the predictions of a dual-process theory, we found no evidence of a U-shaped association between flexibility and number of set problems, at least not in the range of 0-40 problems. These findings highlight the need for further investigation into the situations in which practice and automatization lead to change resistance versus reflection and conscious control.

**Keywords:** problem solving; practice; mental set; change resistance; cognitive control

Almost everyone is familiar with the popular phrase “practice makes perfect.” The principle behind the saying is intuitive—the more you practice something, the better you should become at it. Indeed, the assumption that practice makes perfect underlies many of our educational practices. Students in mathematics classes drill their arithmetic facts over and over, and those learning a musical instrument are encouraged to spend hours practicing each week.

However, practice may also come with undesirable consequences. Research has shown that practice with a single strategy can have a negative effect when people are presented with new problems that cannot be solved with the practiced strategy (e.g., Bilaic, McLeod, & Gobet, 2008; Chen, 1999; Diamond & Kirkham, 2005; Knoblich, Ohlsson, Haider, & Rhenius, 1999; Langer, 2000; Luchins & Luchins, 1950; Luchins, 1946; Luchins, 1942; McNeil & Alibali, 2005; Munakata, 2001; Wiley, 1998). This phenomenon, termed the “Einstellung” phenomenon by Luchins (1942), refers to people’s tendency to use previously learned and practiced strategies even after those strategies cease to be efficient or effective.

Many scientists over the past century have described psychological constructs that overlap with Einstellung. For example, constructs like “habit” (James, 1890), “direction” (Maier, 1931), “perseveration” (Kendig, 1937), “set” (Gibson, 1941; Wiley, 1998), “functional fixedness” (Duncker, 1945), “deep attractor state” (Thelen & Smith, 1994), “mindlessness” (Langer, 2000), “strong representation” (Munakata, 2001), “entrenchment” (Zevin & Seidenberg, 2002), and “attentional inertia” (Diamond & Kirkham, 2005) all reflect the overarching view that the cognitive system tends to resist change, even in the face of external pressure to change.

The classic example of change resistance is performance in Luchins’s water jar studies (Luchins, 1942, 1946; Luchins & Luchins, 1950, see also Bugelski & Huff, 1962; Cunningham, 1965; McGraw & McCullers, 1979; McKelvie, 1984). In these studies, participants were presented with three jars (A, B, and C), each of which held a certain amount of water. The goal was to determine how the jars could be used to obtain a designated amount of water. Both children and adults who practiced several “set” problems that required the same complicated strategy (e.g.,  $A - B - 2C$ ) often persisted in using that strategy on target problems that could be solved by a much simpler strategy. They did so even when the more complicated strategy did not lead to a correct solution. These participants were said to be operating according to an *Einstellung*, or mental set. They rigidly applied their knowledge of the practiced strategy, which apparently made them less open to generating alternative strategies. Participants who did not initially practice the complicated strategy used the simpler strategy on the target problems.

Despite the fact that Einstellung and other change resistant behaviors have been demonstrated across a wide variety of participant groups and on a wide variety of problem solving tasks, one key question has been largely ignored—how does the amount of practice with the initial strategy relate to the strength of a problem solver’s resistance to change? This question is the focus of the present study.

The absence of previous attempts to address this question may be due to the assumption that it has already been answered. Indeed, Luchins (1942) specifically stated that the number of water jar problems solved with the initial strategy (beyond 1) does not affect participants' tendency to apply the mental set. However, he never provided any systematic evidence to support this claim.

Despite Luchins's claim, many theories suggest that participants should become less flexible as the number of set problems increases (e.g., Langer, 2000; Diamond & Kirkham, 2005; Thelen & Smith, 1994; Munakata, 2001; Wiley, 1998; Zevin & Seidenberg, 2002). For example, according to Munakata's graded representations account, behavior is determined by the strength of the underlying representations, with stronger representations exerting more control over behavior. This account suggests that repeated practice with a single strategy strengthens the representations required for that strategy, making it more likely to be used again in the future. This prediction seems somewhat intuitive—as practice with a given strategy increases, the likelihood of using an alternative strategy decreases.

However, intuition is not always correct, and some dual-process theories predict a nonmonotonic relationship between the number of set problems and participants' subsequent flexibility (e.g., Marcovitch, Zelazo, & Schmuckler, 2002). For example, according to the hierarchical competing systems account, problem solving behavior is governed not only by a response-based system that operates according to low-level learning mechanisms, but also by a conscious representational system that has the potential to override the response-based system (Marcovitch et al. 2002; Marcovitch & Zelazo, 2006; cf. Crowley, Shrager, & Siegler, 1997). According to this account, a problem solver's flexibility in the water jar task should decrease gradually as the number of set problems increases, until it reaches an asymptote. Once the asymptote has been reached, increased practice with the set problems should lead to automatization, which in turn should increase the likelihood that the problem solver will be able to reflect on the task and override the practiced strategy when it is beneficial to do so. Based on this account, one would expect an initial decrease in flexibility as practice with the initial strategy increases, followed by an increase in flexibility as the conscious representational system begins to exert more influence over problem solving behavior (i.e., a U-shaped function).

Does problem solvers' flexibility decrease linearly as a function of the number of set problems, or is the association U-shaped? The answer has yet to be determined. Although many studies have shown that practice with an initial strategy hinders subsequent flexibility on water jar problems, no experimentally sound attempts have been made to manipulate the amount of practice participants engage in. Thus, in the current study, we varied the number of "set" problems and examined the effects on problem solving flexibility.

## Method

### Participants

Participants were 107 undergraduates from a mid-sized private university in the Midwest. Ten participants were excluded due to experimenter or participant error. An additional 16 participants were excluded because they had previously heard of Luchins's water jar problems. Thus, the final sample contained 81 undergraduates (42 male, 39 female; 3% African American or black, 7% Asian, 17% Hispanic or Latino, 73% white) ranging in age from 18-22 years ( $M = 19$  years, 0 months). Participants received one extra credit point toward an introductory psychology class for their participation.

### Task and Procedure

Participants were asked to solve Luchins's classic water jar problems (Luchins, 1942, 1946; Luchins & Luchins, 1950). In these problems, participants are presented with three jars (A, B, and C), each of which holds a certain amount of water. The goal is to determine how the jars can be used to measure out a fourth amount of water. Jars are not graduated, thus requiring participants to use a combination of addition and subtraction to solve each problem.

The task involves three different types of water jar problems: (a) *set problems*, which can be solved only using a complicated, multi-step strategy:  $B - A - 2C$  (i.e., the so-called "Einstellung method"), (b) *critical problems*, which can be solved either by the multi-step strategy, or by a much simpler, single-step strategy:  $A - C$  (i.e., the so-called "direct method"), and (c) *extinction problems*, which can be solved only using the single-step strategy (Luchins, 1942).

Participants in our experiment were seated at computers situated in individual cubicles. They were told: (a) that they would be solving a set of problems, (b) that they should record all answers on the given answer sheet, and (b) that they should solve the problems as quickly as possible, while still maintaining accuracy. After participants solved each problem, they pressed the spacebar to indicate that they had reached a solution, and then they pressed the spacebar again when they were ready to see the next problem.

As in previous experiments, participants completed a sample problem, followed by a specified number of "set" problems (see experimental conditions below), 2 critical problems, 1 extinction problem, and 2 additional critical problems. All critical and extinction problems were identical to those used by Luchins (1942, 1950) and set problems were drawn from Luchins (1942, 1950) and Bugelski, et al. (1962).

### Experimental Conditions

Participants were randomly assigned to one of six experimental conditions. The conditions differed in terms of the number of set problems that participants were required to solve: 0, 1, 5, 10, 20, or 40.

## Coding

We first examined performance on each phase of the task separately: set problems, pre-extinction critical problems, extinction problem, and post-extinction critical problems. Responses on the set problems were coded as correct or incorrect. Responses on the critical problems were coded as single-step correct, multi-step correct, or incorrect. The response on the extinction problem was coded as correct or incorrect. In addition to analyzing each phase of the task separately, we also created an overall task performance score to summarize how flexibly participants performed on the water jar task. Participants received 1 point for every critical problem solved correctly with the efficient, single-step strategy, and 1 point for solving the extinction problem correctly. Thus, the maximum task performance score (i.e., 5) reflects ideal performance on the water jar task. Response time, measured as the time from problem presentation until the participant pressed the space bar to indicate that they had reached a solution, was also recorded for each problem, although for the sake of conciseness, we focus mainly on the strategy data here.

## Results

### Set Problems

Performance on the set problems was near ceiling, and the percentage of set problems solved correctly was similar across conditions (1 set problem  $M = 100\%$ , 5 set problems  $M = 95\%$ , 10 set problems  $M = 99\%$ , 20 set problems  $M = 97\%$ , 40 set problems  $M = 98\%$ ). As expected, the time it took participants to solve their final set problem decreased as the number of set problems increased (1 set problem  $M = 32.35$  s, 5 set problems  $M = 41.08$  s, 10 set problems  $M = 25.99$  s, 20 set problems  $M = 21.54$  s, 40 set problems  $M = 18.22$  s). We performed the appropriate orthogonal polynomial contrasts with number of set problems as the independent variable and reaction time on the final set problem as the dependent variable. The coefficients used in this analysis accounted for the unequal spacing between conditions. The overall effect of condition was significant,  $F(4, 55) = 2.61, p = .04, \eta_p^2 = .16$ . Moreover, the linear trend was significant,  $F(1, 55) = 24.96, p = .01$ . None of the other polynomial contrasts approached significance. Results were comparable when we used average reaction time (instead of reaction time on the final set problem) as the dependent variable. Importantly, there were no significant differences in reaction times between conditions on the first set problem,  $F(4, 55) = 0.26, p = .90$ . Taken together, these results support the assumption that participants gained fluency as the number of set problems increased.

### Pre-extinction Critical Problems

Performance on the pre-extinction critical problems was near ceiling in terms of correctness. All participants solved at least one of the two problems correctly, and 79 of 81 participants solved both problems correctly. However, use of the efficient, single-step strategy was not widespread,

with only 24 of 81 participants (29%) using the single-step strategy on both problems (see Table 1 for use on both problems by condition). We used binomial logistic regression to predict the log of the odds of using the single-step strategy on both pre-extinction critical problems (see Agresti, 1996). Note that the conclusions are unchanged when we predict the log of the odds of using the single-step strategy on at least one problem (instead of both). The predictor variables included number of set problems (centered) and number of set problems (centered) taken to the second power. As predicted by a change-resistance account, the log of the odds of using the single-step strategy decreased as the number of set problems increased,  $\hat{\beta} = -0.15, z = -3.87, Wald(1, N = 81) = 14.99, p < .001$ . The model estimates that the odds of using the single-step strategy decrease by a factor of 1.16 for every additional set problem solved. The quadratic term was also significant,  $\hat{\beta} = 0.006, z = 3.00, Wald(1, N = 81) = 5.34, p = .02$ . However, this seems to reflect the fact that performance neared floor levels at 10 set problems and remained there; there was no evidence of a significant rebound in performance.

### Extinction problem

Performance on the extinction problem was good, with 67 of 81 participants (83%) solving it correctly. Table 1 displays performance by condition. We used binomial logistic regression to predict the log of the odds of solving the extinction problem correctly. The predictor variables included number of set problems (centered) and number of set problems (centered) taken to the second power. As predicted by a change-resistance account, the log of the odds of solving the extinction problem correctly decreased as the number of set problems increased,  $\hat{\beta} = -0.089, z = -1.98, Wald(1, N = 81) = 3.89, p = .048$ . The model estimates that the odds of solving the extinction problem correctly decrease by a factor of 1.10 for every additional set problem solved. The quadratic term was not significant,  $\hat{\beta} < 0.001, p = .98$ .

Table 1: Participants (%) in each condition who used a correct strategy on the extinction problem and the single-step strategy on both pre- and post-extinction problems.

| Condition | Pre | Extinction | Post |
|-----------|-----|------------|------|
| 0         | 76  | 95         | 95   |
| 1         | 22  | 89         | 89   |
| 5         | 27  | 100        | 73   |
| 10        | 0   | 100        | 60   |
| 20        | 12  | 75         | 38   |
| 40        | 7   | 43         | 29   |

## Post-extinction Critical Problems

Performance on the post-extinction critical problems was near ceiling in terms of correctness. All participants solved at least one of the two problems correctly, and 76 of the 81 participants solved both problems correctly. Use of the efficient, single-step strategy was higher than it was on the pre-extinction critical problems, but it was still far from universal, with 44 of the 81 participants (54%) using the single-step strategy on both problems (see Table 1 for use on both problems by condition). We used binomial logistic regression to predict the log of the odds of using the single-step strategy on both post-extinction critical problems. The predictor variables included number of set problems (centered) and number of set problems (centered) taken to the second power. Again, the linear term was significant,  $\hat{\beta} = -0.055$ ,  $z = -2.20$ ,  $Wald(1, N = 81) = 4.87$ ,  $p = .03$ , but the quadratic term was not,  $\hat{\beta} < 0.001$ ,  $p = .64$ .

## Overall Task Performance

We created an overall task performance score to summarize how flexibly participants performed on the water jar task. Averaging across conditions, the overall task performance score for participants was 2.80 (out of 5). Scores ranged from 0-5 ( $SD = 1.75$ ). As shown in Figure 1, participants' overall task performance decreased as the number of set problems increased. We performed a set of orthogonal polynomial contrasts with condition (number of set problems) as the independent variable and overall task performance score (out of 5) as the dependent variable. The coefficients used in this analysis accounted for the unequal spacing between conditions. There was a large significant effect of condition overall,  $F(5, 75) = 10.10$ ,  $p < .001$ ,  $\eta_p^2 = .40$ . Consistent with our previous analyses, the linear trend was significant,  $F(1, 75) = 35.72$ ,  $p < .001$ , and the quadratic trend was not,  $F(1, 75) = 2.20$ ,  $p = .14$ . None of the higher-order trends were significant. The linear component accounted for 71% of the variance that is due to the condition effect.

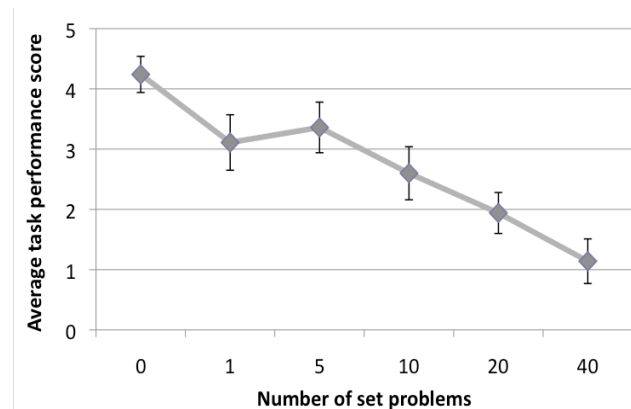


Figure 1: Overall performance on the water jar task as a function of the number of set problems solved. Error bars represent standard errors.

## Discussion

Given the seminal nature of Luchins's (1942) Einstellung phenomenon, it is surprising that the present study was the first to test the effect of number of set problems on performance on the water jar task. Luchins anecdotally reported no effect of the number of set problems (beyond 1) on participants' tendency to apply the mental set. However, our results suggest that Luchins was incorrect. We found a large effect of the number of set problems on problem-solving flexibility. As the number of set problems increased, participants' flexibility decreased. This result places important constraints on theoretical explanations of the mental set phenomena and contributes to our understanding of the role of practice in learning and performance.

Several prevailing theories (and intuition alike) would have predicted a negative linear association between number of set problems and problem-solving flexibility (e.g., Langer, 2000; Diamond & Kirkham, 2005; Thelen & Smith, 1994; Munakata, 2001; Wiley, 1998). For example, according to the graded representations account, problem-solving flexibility decreases with increasing practice because "latent" memory representations of the original strategy compete with "active" memory representations of the problem at hand (Munakata, 1998). Increased practice with set problems strengthens latent representations of the original strategy, and these representations come to exert a larger influence over behavior than the active representation of the given problem. This central tenet is shared by Wiley (1998) in her seminal study of expertise as a mental set. According to Wiley, experts weigh their prior domain knowledge more heavily than the information in a given problem when they construct their problem representation. Although both of these accounts (and others) would have predicted the results of the present study, the literature pointed to at least two other possible outcomes.

First, the number of set problems (beyond 1) could have had no discernable effect on participants' subsequent problem-solving flexibility. This would have been consistent with Luchins's (1942) conclusions. It also would have been consistent with conclusions of other researchers who have examined the effect of number of "set" trials on performance in other problem-solving tasks. For example, Wellman, Cross, and Bartsch (1986) conducted a meta-analysis of babies' performance on the A-not-B task and concluded that the number of A trials has no effect on babies' tendency to reach for location A after the object has been hidden in location B. Similarly, Zelazo, Frye, and Rapus (1996, Experiment 2) studied children's performance on the dimensional change card sort (DCCS) task and concluded that the number of pre-switch trials has no effect on children's tendency to keep sorting based on the pre-switch rules after the rules have been changed. It should be noted that although the A-not-B and DCCS tasks include a motor component that is absent in the water jar task, Marcovitch and Zelazo (2009) assert that it may be possible for representational habits (like those in the water jar task) to influence behavior in a similar way.

Second, the number of set problems could have had a nonmonotonic effect on subsequent problem-solving flexibility. This outcome would have been consistent with the hierarchical competing systems account (Marcovitch et al., 2002; Marcovitch & Zelazo, 2006). According to this account, the initial decrements in flexibility that accompany repeated practice with a strategy are eventually overcome as the strategy becomes automatized and the problem solver becomes increasingly likely to reflect on the task at hand. Marcovitch et al. (2002) provided support for this account in a study of babies' performance on the A-not-B task. Specifically, the authors found an inverted U-shaped association between the number of A trials and babies' tendency to reach for location A after the object has been hidden in location B. Although previous evidence supporting this account has been limited to babies and young children, Marcovitch and Zelazo (2009) state that the theory is intended as a foundation on which studies of executive function across the lifespan can be built.

Indeed, the hierarchical competing systems account shares the implicit assumption of several theories of skill acquisition. Several theories suggest that increased practice with a strategy eventually leads to fluency and automatization, which in turn reduces demands on working memory and supports higher-level cognitive processes (Anderson, 2002; Kotovsky, Hayes, & Simon, 1985; Haverty, 1999; Logan, 1988). The primary role of practice, according to these theories, is to "free resources" so they are available for other aspects of problems solving. In fact, numerous studies have shown that increased practice with a single strategy frees cognitive resources and makes them available for other problem-solving processes such as: (a) noticing novel problem features (Chase & Simon, 1973), (b) generating new solution strategies (Shrager & Siegler, 1998), (c) extracting meaning from unfamiliar problems (Haverty, 1999; Sweller, 1988), and (d) suppressing inappropriate solution strategies that have been retrieved many times in the past (Rosen & Engle, 1998).

Of course, in the present study the association between the number of set problems and flexibility was not U-shaped, at least not in the range of 0-40 set problems. However, it is entirely possible that flexibility would have started to increase if participants had been given even more set problems. We assumed 40 set problems would be sufficient based on the number of set problems informally investigated by Luchins (1942) and because the babies in Marcovitch et al.'s (2002) study only needed between six and 11 trials to overcome the initial decrements in flexibility that come with repeated practice. However, water jar problems involve a complex, multi-step strategy, and participants may need more than 40 problems to achieve the automaticity required for conscious reflection. Although reaction times on the set problems indicated that participants gained fluency as the number of set problems increased, it is still possible that it was not enough for participants to consciously reflect on the task.

Still another possibility is that participants did achieve the automaticity required for conscious reflection, but that the reflection did not occur because there was not an explicit push to do so (e.g., feedback or incentive). When students in one of Luchins's (1942) water jar experiments were given an explicit warning, "don't be blind," they were less likely to rely on the practiced strategy. It is possible that this type of explicit push to reflect on the task would be most beneficial for participants who have practiced many set problems (e.g., 40). Marcovitch and Zelazo (2009) do argue, however, that there is a baseline probability of reflection on each trial and that it increases with task experience.

Although the present results support the hypothesis that problem-solving flexibility decreases as practice with set problems increases, it is possible that the results could be explained by mental fatigue or boredom. Participants who solved 40 set problems spent more time in the experiment than did participants who solved 0 or 1 set problem; thus, they may have been more mentally fatigued. We argue that fatigue did not play a major role because: (a) the difference in total participation time between participants who solved 0 versus 40 set problems was never more than 30 minutes and (b) reaction times on the 40th set problem were significantly faster than reaction times on the 1st set problem. Still, time on task is a confound that could be addressed in a future study.

In terms of practical implications, the present findings may be of interest to educators who are weighing the pros and cons of repeated practice and drill on future learning and performance in domains such as mathematics. For example, previous research has shown that repetitive practice with arithmetic facts may improve some types of inductive reasoning (Haverty, 1999), but hinder understanding of some algebraic concepts (McNeil, 2007). A better understanding of the way in which practice affects later problem-solving flexibility might help us gain a better understanding of these seemingly contradictory findings.

More generally, the present results contribute to a growing body of work suggesting that the effects of practice may not be as straightforward as we like to think. Practice is typically something that we think of as being beneficial to learning and performance, so cases like this where it hinders flexibility can provide a unique window onto how the mind works. One important next step will be for researchers to develop a general framework for understanding the differences between the situations when increased practice increases versus hinders problem-solving flexibility.

## Acknowledgments

This paper is based, in part, on a senior honors thesis conducted by Crooks under the direction of McNeil. Thanks to April Dunwiddie, Jen Hames, Tom Merluzzi, and the members of the Cognition Learning and Development Lab.

## References

- Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, 26, 85-112.

- Bilalic, M., McLeod, P., & Gobet, F. (2008). Why good thoughts block better ones: The mechanism of the pernicious Einstellung (set) effect. *Cognition*, *108*, 652-661.
- Bugelski, B. R. & Huff, E. M. (1962). A note on increasing the efficiency of Luchins' mental set. *American Journal of Psychology*, *75*, 665-667.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.
- Chen, Z. (1999). Schema induction in children's analogical problem solving. *Journal of Educational Psychology*, *91*, 703-715.
- Crowley, K., Shrager, J., & Siegler, R. S. (1997). Strategy discovery as a competitive negotiation between metacognitive and associative mechanisms. *Developmental Review*, *17*, 462-489.
- Cunningham, J. D. (1965). Einstellung rigidity in children. *Journal of Experimental Child Psychology*, *2*, 237-247.
- Diamond, A., & Kirkham, N. (2005). Not quite as grown up as we like to think: Parallels between cognition in childhood and adulthood. *Psychological Science*, *16*, 291-297.
- Duncker, K. (1945). On problem-solving. *Psychological Monographs*, *68*(5, Whole No. 270).
- Gibson, J. J. (1941). A critical review of the concept of set in contemporary experimental psychology. *Psychological Bulletin*, *38*, 781-817.
- Haverty, L. A. (1999). *The importance of basic number knowledge to advanced mathematical problem solving*. Unpublished Doctoral Dissertation, Carnegie Mellon University, Pittsburgh, PA.
- James, W. (1890). *The Principles of Psychology, Chapter IV: Habit*. New York: Henry Holt.
- Kendig, I. (1937). Studies in perseveration. II. Determining factors in the development of compulsive activity. *Journal of Psychology*, *3*, 231-246.
- Knoblich, G., Ohlsson, S., Haider, H., & Rhenius, D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1534-1555.
- Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problem hard? Evidence from the tower of Hanoi. *Cognitive Psychology*, *17*, 248-294.
- Langer, E. J. (2000). Mindful learning. *Current Directions in Psychological Science*, *9*, 220-223.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.
- Luchins, A. S., & Luchins, E. H. (1950). New experimental attempts at preventing mechanization in problem solving. *The Journal of General Psychology*, *42*, 279-297.
- Luchins, A. S. (1946). Classroom experiments on mental set. *American Journal of Psychology*, *59*, 295-298.
- Luchins, A. S. (1942). Mechanization in problem solving: The effect of Einstellung. *Psychological Monographs*, *(6, Whole No. 248)*.
- Maier, N. R. F. (1931). Reasoning in humans I. On direction. *Journal of Comparative Psychology*, *12*, 115-143.
- Marcovitch, S., Zelazo, P. D., & Schmuckler, M. A. (2002). The effect of the number of A trials on performance on the A-not-B task. *Infancy*, *3*, 519-529.
- Marcovitch, S., & Zelazo, P. D. (2006). The influence of number of A trials on 2-year-olds' behavior in two A-not-B-type search tasks: A test of the hierarchical competing systems model. *Journal of Cognition and Development*, *7*, 477-501.
- Marcovitch, S., & Zelazo, P. D. (2009). A hierarchical competing systems model of the emergence and early development of executive function. *Developmental Science*, *12*, 1-25.
- McGraw, K. O., & McCullers, J. C. (1979). Evidence of a detrimental effect of extrinsic incentives on breaking a mental set. *Journal of Experimental Social Psychology*, *15*, 285-294.
- McKelvie, S. J. (1984). Relationship between set and functional fixedness: A replication. *Perceptual and Motor Skills*, *58*, 996-998.
- McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology*, *43*, 687-695.
- McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*, 883-899.
- Munakata, Y. (1998). Infant perseveration and implications for object permanence theories: A PDP model of the A-not-B task. *Developmental Science*, *1*, 161-211.
- Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences*, *5*, 309-315.
- Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. *Journal of Memory and Language*, *39*, 418-436.
- Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*, 405-410.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*, 257-285.
- Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Wellman, H. M., Cross, D., & Bartsch, K. (1986). Infant search and object permanence: A meta-analysis of the A-not-B error. *Monographs of the Society for Research in Child Development*, *51*.
- Wiley, J. (1998). Expertise as mental set: The effects of domain knowledge in creative problem solving. *Memory and Cognition*, *26*, 716-730.
- Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*, 37-63.
- Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in reading and other tasks. *Journal of Memory and Language*, *47*, 1-29.