

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

What is the Trouble with Transfer?

Permalink

<https://escholarship.org/uc/item/87m565jx>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN

1069-7977

Authors

Opfer, John E.
Thompson, Clarissa A.

Publication Date

2007

Peer reviewed

What is the Trouble with Transfer?

John E. Opfer (opfer.7@osu.edu)
245 Psychology Building, 1835 Neil Ave. Mall
Columbus, OH 43210

Clarissa A. Thompson (thompson.1345@osu.edu)
125 Psychology Building, 1835 Neil Ave. Mall
Columbus, OH 43210

Abstract

Spontaneous transfer of learning is widely reported to be difficult to elicit. We hypothesized that one reason this finding is so widespread is that pretests proactively interfere with performance on transfer. To test this hypothesis, we examined transfer of learning across two numerical tasks (number line estimation and categorization) in which similar representational changes have been observed. Children who were given feedback on the number line estimation task learned to use a linear representation of numerical quantity instead of a logarithmic one, but simply providing children with practice on the categorization pretest led them to continue using a logarithmic representation on the same task, which they otherwise abandoned with surprising frequency.

Keywords: transfer of learning, representational change, proactive interference, numeric cognition, and pretests

The Trouble with Transfer

Despite the importance of transferring knowledge—whether from domain to domain, from school to everyday life, or from everyday life to school—spontaneous transfer is notoriously difficult to elicit, with learners typically generalizing new approaches to a much narrower set of problems than is optimal (Barnett & Ceci, 1992; Gick & Holyoak, 1983; Singley & Anderson, 1989; Thorndike, 1922). This difficulty is not simply because initial learning is incomplete or unstable; even in microgenetic studies of children’s learning, which allow the stability of new knowledge to be established by obtaining a dense sampling of children’s thinking over time by means of trial-to-trial assessments, children very often under-extend novel solutions (see Siegler, 2006, for a review of 105 microgenetic studies examining the issue).

Numerous explanations have been proposed to identify why transfer is so difficult for learners to achieve. These explanations have included the degree of overlap in production rules between the base and transfer domain (Singley & Anderson, 1989), where attention is directed during learning (Anderson, Reder, & Simon, 1996), lack of structural similarity between the base domain and transfer domain (Gick & Holyoak, 1983), the stability and encoding of initial learning (Opfer & Siegler, 2004), and insufficient prior knowledge in the base domain (Brown, Kane, & Echols, 1986). Indeed, the many possible impediments to transfer have led some investigators to posit the somewhat pessimistic conclusion that learning is universally narrow and “situation specific” (Lave, 1988).

To add to this already daunting list, we propose a novel (yet somewhat more optimistic) explanation for why transfer is difficult to elicit in microgenetic studies in particular and possibly other studies more generally—proactive interference from previous practice. The basic premise of our account is that when children have practice on a task without any feedback (e.g., when they complete a pretest for an experimenter), children must use some representation to complete that task, and the more they use that representation, the greater the strength of the representation, and the more likely children will continue using it (Siegler & Shipley, 1995). Under circumstances in which the representations used on the task are already appropriate, pretests can facilitate transfer (“practice makes perfect”) (Roediger & Karpicke, 2006; Gick & Holyoak, 1983). However, when representations are inappropriate, practice on a pretest makes imperfect because practice merely strengthens the inappropriate representation and thereby blocks transfer of the more optimal representations learned during training (Gick & Holyoak, 1980). If true, this explanation is not a trivial one: it suggests that the trouble with transfer is at least partly an experimental artifact of studies that pretest the treatment and control groups.

To test our practice interference hypothesis, we used Solomon and Lessac’s (1968) four-group design to assess the independent and interactive effects of treatment (feedback) and pretesting on children’s judgments of numerical magnitude. The tasks we chose were ones where previous cross-sectional (Laski & Siegler, 2005; Siegler & Opfer, 2003) and microgenetic (Opfer & Siegler, in press) studies had shown that young children use an inappropriate (logarithmic) representation before using an appropriate (linear) one (e.g., younger children judge 150 to be closer to 1000 than to 1, whereas older children judge 150 to be closer to 1 than to 1000). In the first task, children were asked to make estimates of numerical quantity (e.g., where 150 would fall on a line flanked by 0 and 1000), and they were given feedback on their estimates so they would learn to use a linear representation (as in Opfer & Siegler, in press). In the second task, children were asked to categorize numerals by their magnitude (e.g., whether 150 is a small or big number in the context of the 0-1000 range); this is the task where we hoped children would transfer their learning. On the assumption that the two tasks tapped a common representation, we predicted that children’s learning of the linear representation on the number line estimation task would transfer to their category

judgments (at least when they were not given a pretest). Further, based on our proactive interference hypothesis, we predicted that pretesting categorization would also strengthen the logarithmic representation and thus block spontaneous transfer of learning.

Although there are many possible interactions between pretest and treatment, we illustrate two of the most interesting ones for our study in Figure 1: the case where pretesting the experimental group inhibits the effect of treatment, and the case where pretesting the experimental group enhances the effect of treatment. The first case directly corresponds to the predictions of the proactive interference hypothesis, whereas the second case corresponds to findings that would falsify the hypothesis.

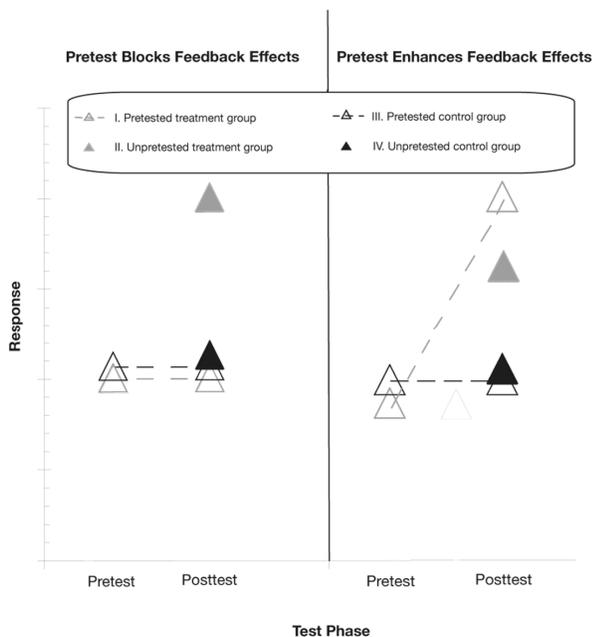


Figure 1: Pretesting Predictions

Method

Participants

Participants were 56 first and second graders ($M = 7.85$, $SD = .65$; 29 females, 27 males). The children attended schools in middle class suburbs near a large Midwestern city.

Tasks

Numberline Task All number line problems consisted of a 20 cm line with the left endpoint labeled “0,” the right endpoint labeled “1000,” and with the number to be estimated appearing 2 cm above the midpoint of the number line. Participants were asked to place the following numbers on a number line by making a hatch mark: 2, 5, 18, 27, 34, 42, 56, 78, 100, 111, 122, 133, 147, 150, 156, 162, 163, 172, 179, 187, 246, 306, 366, 426, 486, 546, 606, 666, 722, 725, 738,

754, 818, 878, and 938. These numbers maximized the discriminability of logarithmic and linear functions by oversampling the low end of the range, minimized the influence of specific knowledge (that 500 is halfway between 0 and 1000), and tested predictions about the range of numbers where estimates would most differ between pretest and posttest.

Categorization Task In the categorization task, children were asked to say how large numbers were when compared to 0 (really small) and 1000 (really big). To do this, children categorized numbers by using five aptly-labeled boxes as mnemonic devices: ‘really small’, ‘small’, ‘medium’, ‘big’, and ‘really big’ (adapted from Laski and Siegler, 2005). The 10 numbers children were asked about comprised a subset of the numbers used in the number line task (2, 5, 78, 100, 150, 246, 486, 606, 725, and 938) and were randomized for each participant.

Design and Procedure

Children were randomly assigned to one of four groups: a pretested treatment group (Group I: categorization pretest + feedback), an unpretested treatment group (Group II: no categorization pretest + feedback), (3) a pretested control group (Group III: categorization pretest + no feedback), or an unpretested control group (Group IV: no categorization pretest + no feedback).

Children in all groups completed the number line estimation task at pretest, over three training trial blocks, and at posttest. On pretest and posttest, children in all groups were presented the same 22 problems without feedback. For children in the two treatment groups, each training trial block included a feedback phase and a test phase. The feedback phase included either one item on which children received feedback (Trial Block 1) or three items on which they received feedback (Trial Blocks 2 and 3). The test phase included 10 items on which children did not receive feedback; this test phase occurred immediately after the feedback phase in each training trial block. Children in the control groups received the same number of estimation problems, but they never received feedback.

Feedback was administered to the two treatment groups (Group I and Group II) following the same procedure used in Opfer and Siegler (in press). On the feedback problems, children were told to make a hatch mark indicating where they believed the numerosity was supposed to go, and then the experimenter would show him/her how close the mark was to the actual location of the numerosity by making a second hatch mark on the number line. Children’s answers that deviated from the correct hatch mark by less than 10% were described to the children as being “really quite close”, whereas children’s answers that deviated by more than 10% were described as “a bit too high/too low”.

Results

We organized our results into two sections: (1) the process of change in numerical estimation (Siegler, 1996), and (2)

transfer of knowledge from numerical estimation to numerical categorization.

Process of Change in Numerical Estimation

Source of Change We first examined the source of change in estimation performance on the number line task. Specifically, we wanted to test whether the experiences that children received during the training phase of the experiment improved their estimation accuracy and influenced the degree to which their estimates came to follow a linear function. To find out, we compared pretest and posttest performance of the treatment groups to the control groups.

We examined whether or not there was a hypothesized logarithmic-to-linear shift. On the pretest, children's mean estimates for each number were in fact fit better by the logarithmic function than by the linear one regardless of experimental condition. The precision of the fit of the logarithmic function, and the degree of superiority of that function to the linear function, was similar across the treatment ($\log R^2 = .95$; $\text{lin } R^2 = .80$) and control groups ($\log R^2 = .93$, $\text{lin } R^2 = .82$). In contrast, the groups differed considerably in their posttest estimation patterns. Children in the control groups continued to generate estimates that fit the logarithmic function better than the linear one ($\log R^2 = .91$, $\text{lin } R^2 = .85$). In contrast, children in the treatment groups generated posttest estimates that fit the linear function substantially better than the logarithmic one ($\text{lin } R^2 = .96$, $\log R^2 = .69$).

Rate of Change To address the rate of change among the four groups, we analyzed the performance of all children who initially provided a logarithmic pattern of estimates and then compared the estimates of the two control groups to the two treatment groups on a trial-block to trial-block basis. We assigned a 1 to the trial blocks of each child that were best fit by the linear function and a 0 to the trial blocks that were best fit by the logarithmic function. The key prediction was that training group and trial block would interact, with the interaction due to children learning fastest in the feedback groups and slowest (if at all) in the no-feedback groups.

A 2 (training group: treatment, control) X 5 (trial block: pretest, 1, 2, 3, posttest) repeated-measures ANOVA indicated effects for training group, $F(1, 36) = 55.31$, $p < .001$, for trial block, $F(4, 144) = 10.71$, $p < .001$, and for the interaction between the two variables, $F(4, 144) = 5.27$, $p < .001$. The linear function more frequently fit the estimates of children in the treatment groups (56.25% of trial blocks) than the estimates of children in the control groups (11.82% of trial blocks, $p < .001$, $d = 2.34$). The effect of trial block was due to the linear function providing the better fit more often on trial blocks 1, 2, 3, and posttest (37%, 39%, 39%, and 37% respectively) than on the pretest (0%, p 's $< .001$).

The interaction between training group and trial block (Figure 2) reflected different rates of learning between the treatment and control groups. On pretest, there were no differences among groups in the percentage of children for whom the linear function provided the better fit (it was 0% by

definition). On trial block 1, the linear function fit more children's estimates in the treatment groups ($M = 69\%$, $SD = .23$) than in the control groups ($M = 14\%$, $SD = .12$) ($t[36] = 4.10$, $p < .001$, $d = 3.0$). On trial block 2, children in the treatment groups continued to generate linear patterns of estimates more frequently than children in the control groups ($M = 63\%$, $SD = .25$ versus $M = 23\%$, $SD = .18$, $t[36] = 2.63$, $p < .05$, $d = 1.84$). On trial block 3, children in the treatment groups also continued to generate linear patterns of estimates more frequently than children in the control groups ($M = 75\%$, $SD = .2$ versus $M = 14\%$, $SD = .12$, $t[36] = 4.74$, $p < .001$, $d = 3.7$). Finally, on posttest, children in the treatment groups generated linear patterns of estimates more frequently than children in the control groups ($M = 75\%$, $SD = .2$ versus $M = 9\%$, $SD = .09$, $t[36] = 5.48$, $p < .001$, $d = 4.3$). What this meant was that treatment effects manifested and persisted after feedback on a single estimate.

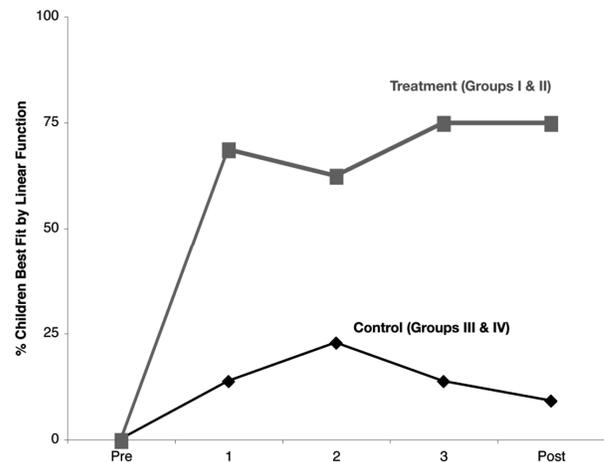


Figure 2: Rate of Change

Path of Change Children could have moved from a logarithmic to a linear representation via several paths. To examine which path(s) they actually took, we examined trial-block to trial-block changes in the fit of the linear function to individual children's estimates. In particular, we identified the first trial block on which the linear function provided the best fit to a given child's estimates, and we labeled it "trial block 0." The trial block immediately before each child's trial block 0 was that child's "trial block -1", the trial block before that was the child's "trial block -2", and so on.

These assessments of the trial block on which children's estimates first fit the linear function made possible a backward-trials analysis that allowed us to test alternative hypotheses about the path of change from a logarithmic to a linear representation. One hypothesis, suggested by incremental theories of representational change (Brainerd, 1983), was that the path of change entailed gradual, continuous improvements in the linearity of estimates. According to this hypothesis, the fit of the linear model would have gradually increased, from Trial Block -3 to Trial Block

+3. In this scenario, Trial Block 0 — the first trial block in which the linear model provided the better fit — would simply mark an arbitrary point along a continuum of gradual improvement, rather than the point at which children first chose a different representation.

A second hypothesis was that the path of change involved a discontinuous switch from a logarithmic to a linear representation, with no intermediate state. This would have entailed no change in the fit of the linear model from Trial Block -3 to -1, a large change from Trial Block -1 to Trial Block 0, and no further change after Trial Block 0. This second hypothesis clearly fit the data. As shown in Figure 3, from Trial Block -3 to -1, there was no change in the fit of the linear function ($F < 1$, ns). There also was no change from Trial Block 0 to Trial Block 3 in the fit of the linear function ($F < 1$, ns). However, from Trial Block -1 to Trial Block 0, there was a large increase in the fit of the linear function to individual children's estimates, from an average $R^2 = .46$ to an average $R^2 = .68$, $F(1, 80) = 12.20$, $p < .001$, $d = 2.67$. Thus, rather than Trial Block 0 reflecting an arbitrary point along a continuous path of improvement, it seemed to mark the point at which children switched from a logarithmic representation to a linear one.

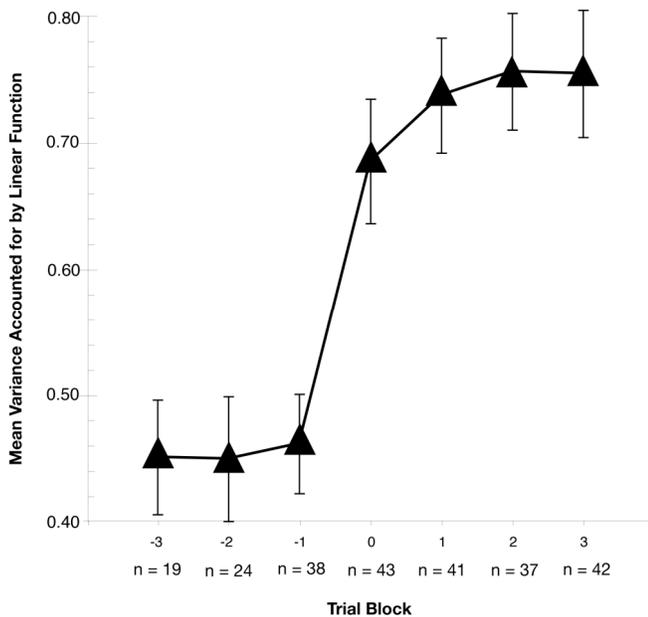


Figure 3: Path of Change

Transfer of Learning to Numerical Categorization

Finally, to test our proactive interference hypothesis, we examined the breadth of changes in children's estimates by examining whether children transferred learning on number line problems to their performance on the number categorization task and whether this transfer was impeded by previous practice.

To examine these issues, we first analyzed the relation between numerical value and categorization on children's pretest and posttest performance on the categorization task. To do so, category labels were converted to a numeric code (i.e., "really small" = 0, "small" = 1, "medium" = 2, "big" = 3, and "really big" = 4), and then we examined the fit of the linear and logarithmic regression functions to the mean judgments. As on the number line task, children's pretest magnitude judgments were again better fit by a logarithmic ($\log R^2 = .95$) than by a linear function ($\text{lin } R^2 = .69$). The fit of the logarithmic function did not result from aggregating over subjects: of all the children who received a pretest, 90% provided judgments that were better fit by the logarithmic than linear function, with the average fit of the logarithmic function ($\text{mean } \log R^2 = .78$, $SD = .03$) being significantly better than the average fit of the linear function ($\text{mean } \text{lin } R^2 = .59$, $SD = .05$), $t(30) = 5.71$, $p < .001$, $d = 4.61$. Furthermore, pretest performance on the two tasks was highly correlated: the more linear were the estimates on the number line task, the more linear were the judgments on the number categorization task ($r = .52$, $F[1, 30] = 10.92$, $p < .01$), and the more logarithmic were the estimates, the more logarithmic were the judgments on the number categorization task ($r = .72$, $F[1, 30] = 30.41$, $p < .001$). Thus, it appeared that a common, logarithmic representation of numeric value influenced children's categorization as well as estimation performance.

We next examined the effect of pretesting on transfer. As expected, category judgments on post-test varied substantially with the administration of treatment (feedback) and pretest (see Figure 4). At the group level, the linear function provided a better fit for the mean judgments of children in the unpretested treatment group ($\text{lin } R^2 = .84$) than in the pretested treatment ($\text{lin } R^2 = .72$), pretested control ($\text{lin } R^2 = .75$), and unpretested control ($\text{lin } R^2 = .68$) groups. The same pattern emerged when looking at the proportion of children who were best fit by each function, with 46% of children's judgments in the unpretested treatment group being best fit by the linear function versus 23% of children in the pretested treatment group, 21% of children in the unpretested control group, and 17% of children in the pretested control group.

Finally, to test for the predicted interaction between pretesting and treatment, we conducted a 2 (pretesting: yes, no) X 2 (treatment: yes, no) factorial ANOVA on the fit of the linear function for each child's judgments. Pretesting and treatment produced no main effects, but there was a substantial interaction between the two variables, $F(1, 56) = 5.32$, $p < .05$. The unpretested treatment group provided significantly more linear judgments ($\text{mean } R^2 = .76$, $SD = .02$) than did the pretested treatment group ($\text{mean } R^2 = .54$, $SD = .07$), $t(22) = 2.43$, $p < .05$, $d = 4.27$, and they also provided slightly more linear judgments than did the control groups (pretested, $\text{mean } R^2 = .64$, $SD = .03$; unpretested, $\text{mean } R^2 = .59$, $SD = .07$; p 's $< .07$), which did not differ from each other. To appreciate just how powerful transfer was when children received treatment but no pretest, it is useful to compare performance shown in Figure 4 (which depicts

linearity of number categorization) to the last trial blocks shown in Figure 3 (which depicts linearity of number line estimates): the variance accounted for by the linear function in the two tasks is nearly identical (estimation, $mean \text{ lin } R^2 = .75$; categorization, $mean \text{ lin } R^2 = .76$), which is consistent with an almost perfect transfer of a linear representation across the two contexts.

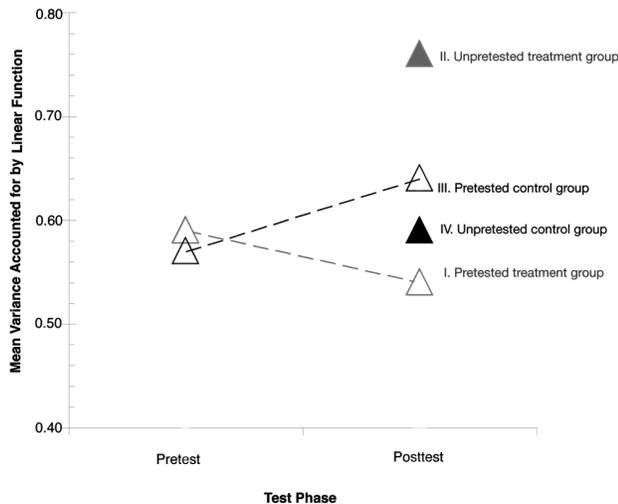


Figure 4: Breadth of Change

General Discussion

Spontaneous transfer of learning is notoriously difficult to elicit (Barnett & Ceci, 2002; Gick & Holyoak, 1983; Thorndike, 1922), even in microgenetic studies that allow one to statistically control for pre-transfer learning (Siegler, 2006). We hypothesized that one source of this difficulty comes from previous experience with the transfer task, such as that provided on a pretest.

To test our hypothesis about the difficulty with transfer, we used Solomon-Lessac's four-group design to examine how children acquired linear representations of numerical magnitude on an estimation task (see Process of Change in Numerical Estimation section), whether they transferred these representations to a new context in which they were asked to categorize numbers, and whether pretesting interfered with this transfer (see Transfer of Learning to Numerical Categorization section). Consistent with previous findings (Opfer & Siegler, in press), our study of microgenetic changes on the estimation task revealed that a single trial of feedback on the linear magnitude of 150 immediately elicited large and abrupt improvements in estimation, ones consistent with use of a linear representation of number. (Whether looking at this study alone or by comparing across studies, pretesting had no effect on the source, rate, and path of change.) Further, our study of transfer also revealed quite robust transfer of learning to the categorization context—at least when no pretest was administered. Indeed, the most striking evidence for transfer of representations across the two contexts came from the average fit of the regression functions

on the two tasks: the fits of the logarithmic and linear functions to children's category judgments were nearly identical to the fit of the same functions to children's estimates before (categorization, $\log R^2 = .78$; estimation, $\log R^2 = .72$) and after learning to use the linear representation (categorization, $\text{lin } R^2 = .76$; estimation, $\text{lin } R^2 = .75$). Thus, without the administration of pretest, we observed nearly perfect transfer; with the administration of pretest, we observed virtually no transfer at all.

Why Did the Pretest Interfere with Transfer?

One mechanism that might account for the pretest interfering with transfer is simple fatigue. Within this hypothesis, our giving pretested children so many problems to solve reduced their performance relative to the non-pretested children because the pretest caused children to become tired of the experiment. As a direct test of whether fatigue caused the pretested and unpretested groups to differ, we ran a follow-up study consisting of 20 children, who were given an equally long, but non-numeric pretest (animal naming) before estimation training. For these children, pre-testing did nothing to block transfer of estimation training: their numeric categorizations were more linear ($R^2 = .73$) than both children pretested for numeric categorization and given feedback ($R^2 = .54$) and children who received no estimation training ($R^2 = .47$), and their numeric categorizations were almost identical to those of the unpretested treatment group in Figure 4 ($R^2 = .76$). Thus, our follow up study indicated that *the numeric content of the pretest--and not its length--blocked transfer*.

In our view, pretests interfered with transfer for an altogether different reason: pretests strengthen default representations. In this view, practice on the categorization pretest made for imperfect performance on that task (and no other) because practice merely strengthened the logarithmic representation that children possessed, and thereby blocked transfer of the more optimal (but under-practiced) linear representation learned during training.

How broadly does prior experience inhibit the transfer of learning? One previous example is provided by Duncker's (1945) classic study of problem-solving. Duncker tested whether participants could apply novel functions to various familiar objects (e.g., a matchbox, tacks, and candles). For example, to solve the problem of placing three small candles on a door at eye-level, participants had to conceive of a novel function for the matchbox (i.e., to serve as a platform). Normally, about 86% of such problems were solved; however, if subjects were given a pretest to determine their knowledge of the original functions of the familiar objects, the rate of problem solving dropped to 58%. Duncker's explanation was that previous experience with the objects induced a "functional fixedness" that inhibited their novel solutions. It may be that prior experience more broadly induces a kind of 'conceptual fixedness' that might also prevent children from transferring their knowledge. For example, one reason that children may fail to transfer from school lessons to familiar real-world problems (yet succeed in transferring to novel problems like throwing darts at an

underwater dartboard; Hendrickson & Schroeder, 1941) is that their previous successes on real-world problems interferes with the application of novel school lessons, much like Duncker's pretest interfered with his subjects' ability to think of novel solutions. That is, previous experiences in school and real-world settings interfere with transfer because they lead children to think that symbolic operations are "for school" and not "for real-world problems", much like Duncker's subjects thought of matchboxes as "for holding matches" and not "for supporting candles".

Although our application of Duncker's findings to transfer is novel, it is also consistent with many previous findings on the formation of associations between strategies and problems (Siegler & Shipley, 1995), the formation of undesirable memory traces through practice (Roediger & Payne, 1982), with findings of set effects in analogy (Gick & Holyoak, 1980), and with the effects of progressive alignment on transfer (Markman & Gentner, 1997). To our knowledge, however, the implications of this research for the possibly harmful effects of repeated testing on transfer has not been examined previously, despite its importance in interpreting the narrow transfer of learning sometimes observed in microgenetic studies. Thus, in our view, the finding that pretesting can inhibit later transfer is not merely a methodological issue, but a much broader one concerning the general difficulty of transfer.

References

- Anderson, J. R., Reder, L. M., & Simon, H. A. (1996). Situated learning and education. *Educational Researcher*, 25, 5-11.
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128, 612-637.
- Brainerd, C. J. (1983). Young children's mental arithmetic errors: A working-memory analysis. *Child Development*, 54, 812-830.
- Brown, A. L., Kane, M. J., & Echols, C. H. (1986). Young children's mental models determine analogical transfer across problems with a common goal structure. *Cognitive Development*, 1, 103-121.
- Duncker, K. (1945). On problem-solving. *Psychological Monographs*, 58.
- Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, 12, 306-355.
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- Hendrickson, G., & Schroeder, W. H. (1941). Transfer of training in learning to hit a submerged target. *Journal of Educational Psychology*, 35, 205-213.
- Laski, E., & Siegler, R.S. (2005, October). *Big, Small, or Just Right? Children's Number Categories and Understanding of Numerical Magnitude*. Poster Session at the biannual meeting of Cognitive Development. San Diego, CA.
- Lave, J. (1988). *Cognition in practice: Mind, mathematics, and culture in everyday life*. Cambridge, UK: Cambridge University Press.
- Markman, A. B., & Gentner, D. (1997). The effects of alignability on memory. *Psychological Science*, 8, 363-367.
- Opfer, J. E., & Siegler, R. S. (in press). Representational change and children's numerical estimation. *Cognitive Psychology*.
- Opfer, J. E., & Siegler, R. S. (2004). Revisiting preschoolers' living things concept: A microgenetic analysis of conceptual change in basic biology. *Cognitive Psychology*, 49, 301-332.
- Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249-255.
- Roediger, H. L., III, & Payne, D. G. (1982). Hypermnnesia: The role of repeated testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 66-72.
- Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford.
- Siegler, R. S. (2006). Microgenetic analyses of learning. In W. Damon & R. M. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Volume 2: Cognition, perception, and language* (6th ed., pp. 464-510). Hoboken, NJ: Wiley.
- Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, 14, 237-243.
- Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling*. Hillsdale, NJ: Erlbaum.
- Singley, K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.
- Solomon, R. L., & Lessac, M. S. (1968). A control group design for experimental studies of developmental processes. *Psychological Bulletin*, 70, 145-150.
- Thorndike, E. L. (1922). The effect of changed data upon reasoning. *Journal of Experimental Psychology*, 5, 33-38.