Title: The Effects of Feedback During Exploratory Mathematics Practice: Evidence for an Expertise Reversal Effect

Author(s): Emily R. Fyfe, Vanderbilt University; Bethany Rittle-Johnson, Vanderbilt University; Marci S. DeCaro, Vanderbilt University
Abstract Body

Limit 4 pages single spaced.

Background / Context:

Description of prior research and its intellectual context.

The notion that discovery learning is more beneficial than direct instruction has a long history in psychology and in education dating back to Bruner (1961), Piaget (1970), and Kuhn (1989). Indeed, educators often associate discovery learning with flexible and robust conceptual knowledge as opposed to static knowledge and rote memorization (e.g., Bredekamp & Copple, 1997; Fuson et al., 1997). Despite the proposed advantages of discovery learning, a number of critiques raise doubts as to its effectiveness. For example, cognitive load theorists note that discovery learning requires random search and simultaneous problem solving, which often overloads working memory capacity (Sweller, 1988). Additionally, learners left to “discover” principles on their own might never come into contact with the relevant information (Kirschner, Sweller, & Clark, 2006; Mayer, 2004). Because of these drawbacks, an emerging consensus suggests that people often learn best though some combination of exploration and instruction, as opposed to one or the other (e.g., Alfieri et al., 2010; Mayer, 2004; Schwartz & Bransford, 1998). In fact, recent evidence indicates that “guided discovery methods,” which include some form of hints or coaching during problem solving enable deeper learning than discovery learning alone (Alfieri et al., 2010). However, few studies have specifically investigated how best to combine instructional and exploratory elements to achieve optimal effects.

We suggest feedback as one instructional element that may be particularly effective in combination with exploration. Feedback is any information about performance or understanding that the learner can use to confirm, reject, or modify prior knowledge. This information may reduce the disadvantages of discovery learning by guiding the learner’s search for information without providing direct instruction. In fact, past research indicates that feedback’s primary function is to identify inaccurate information and encourage the adoption of correct alternatives (e.g., Anderson, Kulhavy, & Andre, 1971, 1972; Bardwell, 1981; Kulhavy, 1977). These studies demonstrate that participants who receive feedback have a higher probability of correcting their errors than participants who do not. Thus, we hypothesize that feedback during exploratory math practice will be more beneficial than no feedback. However, the effects of feedback are not always straightforward. For example, in one meta-analysis, over one third of the feedback effects were negative, indicating that feedback in those instances actually decreased performance (Kluger & DeNisi, 1996). Additionally, feedback may only benefit a certain subset of learners.

In addition to comparing feedback and no feedback, we examined if the type of feedback influences children’s active exploration. We focused on the two most relevant for mathematical problem solving: outcome feedback, which provides a judgment about the accuracy of the learner’s response, and strategy feedback, which provides a judgment about how the learner obtained that response. Outcome feedback that provides the correct answer is consistently related to strong, positive effects, particularly when compared to no-feedback controls (e.g., Kluger & DeNisi, 1996). Strategy feedback, on the other hand, has not been studied extensively. However, at least two studies have found that strategy feedback can be more beneficial than outcome feedback, particularly in terms of strategy selection (Ahmad, 1988; Luwel, et al., 2011).

Finally, we explored the robustness of guided exploration by examining learners with low and high knowledge of the target domain. Past work has found evidence for an expertise reversal effect, a phenomenon in which an instructional method that is effective for novices loses its benefits for learners with higher knowledge (Kalyuga et al., 2003). Generally, this is due to high-
knowledge learners experiencing extraneous cognitive load when additional, redundant guidance is provided. It may be that guided exploration is necessary for novices who have limited prior knowledge to rely on in their search for relevant information. However, high-knowledge learners may not need the extra guidance (e.g., feedback) and may perform better without it.

**Purpose / Objective / Research Question / Focus of Study:**

*Description of the focus of the research.*

The goal of this study was to examine the effects of a specific instructional element (i.e., feedback) in combination with exploratory practice for children with varying levels of prior knowledge. In particular, we examined the role of feedback during exploration on children’s learning of math equivalence. Math equivalence is the idea that two sides of an equation are equal or interchangeable and it is a fundamental concept in both arithmetic and algebra. In addition to comparing feedback to no feedback, we also examined two specific types of feedback (outcome feedback and strategy feedback) and their effects both during the exploration and on posttest assessments. We hypothesized that children who received feedback would exhibit better procedural knowledge of math equivalence than children who did not. Additionally, we expected strategy feedback to lead to the highest performance. Finally, we expected feedback to have a larger impact for children with low prior knowledge.

**Setting:**

*Description of the research location.*

We worked with children in 7 second- and third-grade classrooms in a public school in Middle Tennessee.

**Population / Participants / Subjects:**

*Description of the participants in the study: who, how many, key features, or characteristics.*

Participants were 115 second- and third-grade children. Twenty-four were excluded from participation: 22 for scoring above 80% on a pretest designed to assess children’s prior knowledge of math equivalence and 2 for not completing all activities. The final sample consisted of 91 children (*M age =* 8 yrs, 7 mo; 53 girls; 45% Caucasian; 40% African American, 7% Hispanic, and 8% Other). Just under half (47%) received free or reduced-price lunch.

**Intervention / Program / Practice:**

*Description of the intervention, program, or practice, including details of administration and duration. For Track 2, this may include the development and validation of a measurement instrument.*

The tutoring intervention began with exploratory practice. Children were asked to solve 12 math equivalence problems (e.g., \(5 + 3 + 9 = 5 + \square\)), presented one at a time on a computer screen. After each problem, children reported how they solved the problem and then received different kinds of feedback based on their condition. In the strategy feedback condition, children received feedback on how they solved each problem (e.g., “Good try, but that is not a correct way to solve the problem.”). In the outcome feedback condition, children received feedback on their answer to the problem (e.g., “Good try, but you did not get the right answer. \(X\) is the correct answer.”). Finally, in the no feedback condition, children did not receive any feedback and were simply told to try the next problem. After the exploratory practice, all children rated their subjective cognitive load, completed a short midtest, and then received brief conceptual instruction on the relational function of the equal sign. Providing a conceptual lesson after exploratory practice has been shown to improve learning (DeCaro & Rittle-Johnson, 2011).
**Research Design:**
*Description of the research design.*

We used a pretest – intervention – posttest design followed by a two-week retention test. Children were randomly assigned to one of three conditions: (a) strategy feedback, (b) outcome feedback, or (c) no feedback.

**Data Collection and Analysis:**
*Description of the methods for collecting and analyzing data.*

*For Track 2, this may include the use of existing datasets.*

Children completed a written pretest in their classrooms in one 30-minutes session. Within 1 week, they completed a one-on-one tutoring intervention and posttest in a single session lasting approximately 45 minutes. Two weeks later they completed the retention test in small-group sessions. A math equivalence assessment adapted from Rittle-Johnson, Matthews, Taylor, & McEldoon (2011) was administered at pretest, posttest, and retention test. The assessment included conceptual (8 items) and procedural (8 items) knowledge subscales. Items and scoring criteria are presented in Table 1. Procedural knowledge scores were based on children’s use of a correct strategy to solve each equation (see Table 2). A brief version of the assessment was given as a midtest (5 items—3 conceptual and 2 procedural knowledge items).

**Findings / Results:**
*Description of the main findings with specific details.*

To efficiently test our specific predictions of interest, we used contrast codes to represent our categorical “condition” variable (West, Aiken, & Krull, 1996). Because our condition variable has three groups (strategy feedback, outcome feedback, and no feedback), we created two coded variables. Based on our predictions, we selected the set of helmert contrast codes. The first coded variable (feedback) compared no feedback to the two feedback conditions combined. The second coded variable (feedback type) compared outcome feedback to strategy feedback.

**Procedural Knowledge.** To examine children’s procedural knowledge, we conducted a repeated-measures ANCOVA with feedback and feedback type as between-subject variables and time (midtest, posttest, retention test) as the within-subject variable. We included three covariates (procedural and conceptual knowledge pretest scores and children’s age) as well as two interaction terms (feedback x prior knowledge [i.e., procedural knowledge at pretest] and feedback type x prior knowledge). There were no main effects of feedback or feedback type, Fs < 1. However, there was a significant feedback by prior knowledge interaction, $F(1, 83) = 7.05$, $p = .01$, $\eta^2_p = .08$. To interpret this interaction, we categorized children as having high prior knowledge (scored above the median on procedural knowledge pretest measure) or low prior knowledge and looked at the simple main effects of feedback for each group. As shown in Figure 1, for the low knowledge group, children who received feedback had higher procedural knowledge scores ($M = 35\%, SE = 5\%$) than children who did not receive feedback ($M = 20\%, SE = 6\%$), $F(1, 83) = 3.84$, $p = .054$, $\eta^2_p = .05$. For the high knowledge group, the pattern was reversed. Children who received feedback had lower procedural knowledge scores ($M = 47\%, SE = 5\%$) than children who did not receive feedback ($M = 67\%, SE = 8\%$), $F(1, 83) = 4.61$, $p = .04$, $\eta^2_p = .05$. The results provide evidence for an expertise reversal effect (Kalyuga et al., 2003). Feedback during exploratory practice is more beneficial than no feedback, but only for children with low prior knowledge. For children with high prior knowledge, the reverse is true.
Subjective Cognitive Load. To better understand why this pattern of results occurred, we explored children’s cognitive load during the intervention. For the low-knowledge group, children who received feedback experienced levels of cognitive load ($M = 3.2$, $SE = 0.16$) similar to those of children who did not receive feedback ($M = 3.0$, $SE = 0.21$), $F(1, 83) = 1.15$, $p = .29$, $\eta^2_p = .01$. However, for the high-knowledge group, children who received feedback experienced higher levels of cognitive load ($M = 3.3$, $SE = 0.19$) than children who did not ($M = 2.5$, $SE = 0.27$), $F(1, 83) = 6.05$, $p = .02$, $\eta^2_p = .07$. These results suggest that feedback during exploration is less advantageous for children with higher prior knowledge because it increases cognitive load relative to no feedback—an effect not found for low knowledge children.

Strategy Variability. We also explored how feedback during problem solving impacted children’s performance. Overall, children who received feedback used a greater number of different strategies ($M = 3.2$, $SE = 0.2$) than children who did not ($M = 2.2$, $SE = 0.2$), $p < .001$. There were also differences in perseveration—using the same incorrect strategy on all problems. More children perseverated in the no feedback condition (23%) than in the feedback conditions (3%; $\chi^2 (2, N = 91) = 8.50$, $p = .004$). When we considered prior knowledge, this effect was only significant for children with low prior knowledge, $\chi^2 (2, N = 54) = 6.53$, $p = .01$. Thus, low-knowledge children used more diverse strategies if they received feedback, but tended to use only one incorrect strategy if they did not. This may explain some children’s performance on the posttest. Strategy generation can be an effective mechanism of cognitive change for low-knowledge children, and feedback was an effective motivator of strategy generation.

Conclusions:
Description of conclusions, recommendations, and limitations based on findings.

Recent evidence suggests that guided discovery approaches facilitate deeper learning than pure discovery alone (Alfieri, et al., 2010). We sought to clarify this claim by examining a specific instructional element (i.e., feedback) in combination with exploratory math practice for children with various levels of prior knowledge. As predicted, feedback led to higher procedural knowledge of math equivalence than no feedback. However, this was only true for children with low prior knowledge. For children with high prior knowledge, feedback actually led to lower procedural knowledge than no feedback. These results are consistent with prior work on expertise reversal effects, in which an instructional technique that is effective for novice learners loses its benefits for more experienced learners (Kalyuga et al., 2003). Additionally, these results clarify recent research on discovery learning. Guided exploration with feedback is beneficial, but only for some children. For children who enter the situation with some prior knowledge, unguided exploration appears more beneficial. Future studies should include measures of prior knowledge to examine the generalizability of these results across age groups and across domains.

We expected strategy feedback to be more beneficial than outcome feedback based on findings in prior work (Luwel et al., 2011). However, our results indicated no differences between the two. Perhaps strategy feedback is more beneficial when the learner is choosing between known, correct strategies than when generating strategies for the first time. Additionally, our strategy feedback was rather vague and may have been misinterpreted. Future studies should include more specific strategy feedback.

In addition to its theoretical contributions, this study also has practical implications. Our results can inform current math tutoring and teaching techniques. Specifically, they suggest that providing guidance (e.g., feedback) during exploratory practice should vary with children’s prior knowledge; it may only be beneficial for children with relatively low knowledge.
Appendices
Not included in page count.

Appendix A. References
References are to be in APA version 6 format.


Appendix B. Tables and Figures
Not included in page count.

Table 1

_Procedural and Conceptual Knowledge Items on the Mathematical Equivalence Assessment_

<table>
<thead>
<tr>
<th>Item Type</th>
<th>Task</th>
<th>Scoring Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Procedural</strong></td>
<td></td>
<td>(α = .83; based on 8 scored items)</td>
</tr>
<tr>
<td>Open-ended problems</td>
<td>Solve 1 problem in nonstandard format (8 = 6 + □)</td>
<td>Use correct strategy (or response must be within 1 of correct answer)</td>
</tr>
<tr>
<td></td>
<td>Solve 6 problems with operations on both sides (e.g., 3 + 4 = □ + 5)</td>
<td>Same as above</td>
</tr>
<tr>
<td></td>
<td>Solve 1 equation with an unknown variable (y + 4 = 8 + 2)</td>
<td>Same as above</td>
</tr>
<tr>
<td><strong>Conceptual</strong></td>
<td></td>
<td>(α = .64; based on 8 scored items)</td>
</tr>
<tr>
<td>Meaning of equal sign</td>
<td>Define equal sign</td>
<td>1 point for relational definition (e.g., the same amount)</td>
</tr>
<tr>
<td></td>
<td>Define equal sign in money context (1 dollar = 100 pennies)</td>
<td>1 point for relational definition</td>
</tr>
<tr>
<td></td>
<td>Rate 2 definitions of equal sign as good, not good, or don’t know</td>
<td>1 point for rating “the same as” as a good definition</td>
</tr>
<tr>
<td></td>
<td>Identify the best definition of equal sign from prior item</td>
<td>1 point for selecting “the same as” as the best definition</td>
</tr>
<tr>
<td></td>
<td>Select the pair of numbers equivalent to 6 + 2</td>
<td>1 point for selecting equivalent number pair</td>
</tr>
<tr>
<td>Structure of equations</td>
<td>Reproduce 3 equivalence problems from memory</td>
<td>1 point for correctly reconstructing each of the 3 problems</td>
</tr>
<tr>
<td></td>
<td>Indicate whether 5 equations are true or false</td>
<td>1 point for correctly recognizing 4 or more equations as true or false</td>
</tr>
<tr>
<td></td>
<td>Explain why 1 valid equation is in fact true</td>
<td>1 point for indicating the two sides of the equation are the same amount</td>
</tr>
</tbody>
</table>

*Note.* Cronbach alphas are for posttest. Alphas were somewhat lower at pretest, largely due to floor effects on some items.
Table 2

*Strategies Used to Solve Math Equivalence Problems*

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Sample explanation (4 + 5 + 8 = ___ + 8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Strategies</td>
<td></td>
</tr>
<tr>
<td>Equalize</td>
<td>I added 4, 5, and 8 and got 17, and 9 plus 8 is also 17.</td>
</tr>
<tr>
<td>Add-Subtract</td>
<td>I added 4, 5, and 8 and got 17, and 17 minus 8 is 9.</td>
</tr>
<tr>
<td>Grouping</td>
<td>I took out the 8’s and I added 4 plus 5.</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>8 divided by 8 is 0 and 4 plus 5 is 9.</td>
</tr>
<tr>
<td>Incorrect Strategies</td>
<td></td>
</tr>
<tr>
<td>Add All</td>
<td>I added the 4, 5, 8 and 8.</td>
</tr>
<tr>
<td>Add-to-Equal</td>
<td>4 plus 5 is 9, and 9 plus 8 is 17.</td>
</tr>
<tr>
<td>Add-Two</td>
<td>I added the 5 and the 8.</td>
</tr>
<tr>
<td>Carry</td>
<td>I saw a 4 here, so I wrote a 4 in the blank.</td>
</tr>
<tr>
<td>Ambiguous</td>
<td>I used 8 plus 8 and then 5.</td>
</tr>
</tbody>
</table>
Figure 1

*Children’s Procedural Knowledge Scores by Condition and Prior Knowledge*

![Bar graph showing procedural knowledge scores by condition and prior knowledge.

Note. Scores are estimated marginal means based on midtest, posttest, and retention test scores. Differences are between the no-feedback condition and the two feedback conditions combined: * p = .04, τ = .054. Error bars represent standard errors.