Abstract Title Page
Not included in page count.

Title: The Self-Explanation Effect When Learning Mathematics: A Meta-Analysis

Author(s): Kelley Durkin

Abstract Body

Background / Context:

Most people would agree that helping students learn new information is important; however, there are numerous methods for achieving this goal, and not all of them have been proven to be effective. One method with intuitive appeal and some empirical backing is prompting students to self-explain. A self-explanation can be defined as generating explanations to oneself in an attempt to make sense of new information (Chi, 2000). The self-explanation effect has been studied since the 1980s, and has been examined in many learning domains, from chemistry to mathematics to argumentation (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989).

Self-explanations are thought to be useful because they foster the integration of new knowledge with existing knowledge, and they support the learner in updating his or her current mental model (Chi, de Leeuw, Chiu, & LaVancher, 1994). Having students self-explain might also help them monitor their levels of understanding more accurately (Chi et al., 1989). In addition, self-explanation may encourage students to discover underlying principles in a domain by increasing their ability to generalize information (Lombrozo, 2006). One learning domain in which the self-explanation effect has been studied frequently is mathematics (e.g., Aleven & Koedinger, 2002; Rittle-Johnson, 2006; Siegler, 2002). Self-explanations can help students improve their conceptual and procedural knowledge of mathematics by integrating knowledge of the problem solving process and knowledge of the underlying principles in the learner's mental model. For the purposes of this meta-analysis, procedural knowledge was defined as "the ability to execute action sequences to solve problems" (Rittle-Johnson, Star, & Durkin, 2009, p. 837). Conceptual knowledge was defined as an understanding of underlying principles in a domain (Rittle-Johnson et al., 2009). While there are some findings that support the previously mentioned hypotheses of the potential benefits of self-explanation in mathematics, the actual empirical findings are ambiguous. There are studies that reveal a positive effect of self-explanation in mathematics (e.g., Aleven & Koedinger, 2002; Mitrovic, 2005); however, there are also studies that reveal no effect (e.g., Große & Renkl, 2007; Matthews & Rittle-Johnson, 2009) or a negative effect for self-explanations (e.g., Berthold & Renkl, 2009; Hilbert, Renkl, Kessler, & Reiss, 2008). Consequently, there is a need to determine if self-explanation in mathematics has actually been shown to be useful overall in past research.

Purpose / Objective / Research Question / Focus of Study:

The purpose of the current meta-analysis reported in this paper is to determine if there is a significant, positive self-explanation effect in mathematics domains or not. In addition, this meta-analysis will examine the possibility that study setting might relate to the effectiveness of self-explanation in mathematics.

Setting:

Due to the nature of this paper as a meta-analysis, the included studies come from a variety of research settings and countries, as listed in Table 1 (please insert Table 1 here). These studies were often conducted by the researchers or with researchers closely monitoring a teacher, and six studies occurred in a lab setting, while twelve studies occurred in classrooms.

Population / Participants / Subjects:

The coded study characteristics are summarized in Table 1. The mean number of participants in each study was around 55.28. Also, the mean age of participants was 15.62 years old, with ten studies involving pre-college students and eight studies including college students. The included studies contained a mean of 39.55% male students. A majority of these studies involved interventions teaching probability or geometry, but there were quite a few different math domains studied. The majority of studies were conducted with participants in the United States of America or in Germany, but there were also studies conducted in Australia, Japan and New Zealand.

Intervention / Program / Practice:

A detailed list of eligibility criteria was developed to determine the eligibility of studies for inclusion in this meta-analysis, including requirements for the intervention used in each study. These criteria included:

1) The study included at least 2 participants.
2) The study assessed an intervention that was in a mathematics domain (economics was not considered mathematics for the purposes of this study).
3) The study assessed an intervention that included instructional time during which participants engaged in self-explanation.
4) The study included a control group that did not engage in any type of self-explanation.
5) The study used an experimental or quasi-experimental design.
6) The study used a quantitative measure of conceptual and/or procedural knowledge of mathematics. Studies that implemented only self-report measures were not eligible.
7) The study was reported in English.

Research Design:

The current paper is a meta-analysis that used data from studies that met the above eligibility criteria and included studies that used an experimental or quasi-experimental design. Standardized mean difference effect sizes were calculated for each study and used to calculate an overall mean effect size.

Data Collection and Analysis:

Search and Retrieval of Studies

An attempt was made to locate all studies that fit the above criteria, including both published and unpublished work. Searches were completed in March 2010 using the electronic databases of PsycINFO, ERIC, ProQuest, the ISI Web of Knowledge, the National Technical Information Service (NTIS), the IES What Works Clearinghouse, and Google Scholar. In addition to these resources, the Vanderbilt library catalog was searched, along with available conference archives from the Cognitive Development Society and the Society for Research in Child Development. Typically the search terms self-explanation and mathematics were used in these searches. The NTIS, What Works Clearinghouse, Google Scholar, and ProQuest Theses and Dissertation databases were searched along with conference proceedings in an effort to identify all possibly eligible grey literature.

After completing these searches, there were 199 references found after removing duplicate references. The abstract from each of these references was scanned to determine whether it met the eligibility criteria, and after doing so, 47 potentially relevant references remained eligible. These references were read in full. Eleven of these references were excluded after being read in their entirety; most often, these references were excluded because they did not contain a control group or the intervention was not in a mathematics domain. Four references were excluded because they could not be found or because they could only be obtained by purchasing them. One reference could not be retrieved in time for inclusion in this meta-analysis abstract. Nine references were excluded because they contained data that was already reported in another reference, and three references were excluded because they were not reported in English. Finally, three references were excluded because they were missing data that was necessary to calculate an effect size. This left eighteen separate studies (from sixteen references) that met the eligibility criteria and were coded for this meta-analysis.

Coding of Studies

Studies were coded to determine a mean difference between conditions (participants who self-explained and those who did not) on procedural knowledge at posttest. The procedural knowledge measure was coded because all studies contained a procedural knowledge measure. Effect sizes were calculated using data from posttest measures because that was the assessment time closest to the intervention. In addition to coding a standardized mean difference effect size for each study, a variety of other variables were coded. These included knowledge type assessed (procedural, conceptual, or both), number of assessment items, whether it was an experimenter-created or standardized measure, total number of participants, total number of participants in each condition, grade of participants, average age of participants, participants' gender (percent male), study setting (lab or classroom), country of study, publication type (journal article, book, etc.), publication year, specific mathematics domain (algebra, geometry, etc.), and additional condition activities (not including self-explanation) (see Table 1).

Analysis Strategies

A standardized mean difference effect size was calculated for each study; the standardized mean difference is the difference between the treatment and control group means on their procedural knowledge scores divided by their pooled standard deviation (Lipsey & Wilson, 2001). Hedge's g correction was used on each of these effect sizes to correct for small sample sizes. This involved multiplying each effect size by $1 - (3/4n - 9)$, where n is the total number of participants in the study (Lipsey & Wilson, 2001). Each study was then weighted using its inverse variance. The mean standardized mean difference effect size was calculated using a random effects model. A random effects model was chosen because it was desirable for these results to be generalizable, and it did not seem likely that the calculated value would be the one true population effect size. When multiple effect sizes could be calculated in a given study, the effect size related to studying correct examples was used. If all possible effect sizes in the study used correct examples, then an effect size was chosen at random for inclusion in the meta-analysis. In addition, a heterogeneity analysis, a meta-regression moderator analysis, and a publication bias analysis were run. The results from these analyses are reported below.

Findings / Results:

The overall standardized mean effect size was 0.373 with a 95% confidence interval ranging from 0.118 to 0.629 (please insert Figure 1 here). This effect size of 0.373 was statistically significant ($p = 0.004$). Consequently, there does appear to be a small, positive effect for students who self-explain when learning mathematics.

After examining Figure 1, it seemed as if there was some heterogeneity in these effect sizes. After running an analysis for heterogeneity, $Q(17) = 63.62$, $p < 0.001$, we should reject the null hypothesis that there was not heterogeneity in these effect sizes. To find out how much of this variance was due to true variability, an I^2 value was calculated. $I^2 = 73.3\%$ which indicated that 73.3% of this heterogeneity is true heterogeneity. In addition, $\tau^2 = 0.2118$, which indicated that there was some variability between studies. When taking all of these heterogeneity analyses into account, this was a sign that there were most likely moderators influencing these effect sizes.

Moderator Analysis

Due to the small number of studies included in this meta-analysis, only one moderator analysis could be run. Frequently in many areas of research, effect sizes of interventions are found to be affected by whether they are conducted in a lab setting or in a classroom setting. A moderator analysis was run to examine whether performing the intervention in a lab setting was a moderator for the standardized mean difference effect size. After running a meta-regression model including study setting, it appeared that being performed in a lab slightly increased a study's effect size, but this difference was not significant, $\beta = 0.235$, $p = 0.377$, $\tau^2 = 0.194$. Consequently, it seems that study setting was not a significant moderator contributing to the heterogeneity in effect sizes. One possible reason this was not a significant moderator is that even the classroom-based studies were either run by researchers in classrooms or with the researchers very closely guiding the teachers. Thus the differences between the lab studies and classroom studies in this case may not have been very large.

Publication Bias

Although there are not many studies included in this meta-analysis, it is important to test for any possible publication bias. An Egger's regression test for small-study effects indicated that there was possibly publication bias. The bias coefficient was 3.563 with $p = 0.03$. Thus we rejected the null hypothesis and concluded that there was evidence of small-study effects in these studies. As an additional measure of publication bias, a trim and fill analysis was run. As a result of this analysis, 4 "filled" effect sizes were found. By including these filled effect sizes into the random effects model, the overall standardized mean difference effect size changed from 0.373 (reported previously) to 0.220. This effect size of 0.220 was no longer statistically significant ($p = 0.095$). This finding led to the conclusion that there was publication bias, with 4 studies with higher positive effect sizes increasing the mean effect size for all of the studies. Once 4 studies were filled in to balance out the higher positive effect sizes, the overall mean effect size was no longer significantly different from 0.

Summary of Results

The above analyses led to several findings about the self-explanation effect in mathematics. First, there was a small, positive mean effect for self-explanation in math domains. There was also significant heterogeneity in the effect sizes from these studies with a large portion of the heterogeneity being true heterogeneity; however, the one moderator tested here, study setting, was not significant. In addition, there were multiple tests that indicated there may be publication bias in this meta-analysis, and these results must be examined with caution.

Conclusions:

While an attempt was made to include every possible study that could contribute to the findings of this meta-analysis, there were some limitations. First, there was publication bias, which means that the findings from this meta-analysis must be interpreted with caution. Also, almost all of the studies were in journal articles, and journal articles are more likely to find some significant effects than other unpublished works. There were also only a small number of studies in this analysis, which must be taken into consideration. In addition, this study only examined the effect size of self-explanation in mathematics domains. It is possible that these same findings would not be found in other domains, such as sciences or the humanities. As a result, the findings from this meta-analysis can only be applied to self-explanation in connection to mathematics and not to the broader concept of the self-explanation effect. Another limitation of this study was that no effect sizes based off of conceptual knowledge measures were used, due to the smaller number of conceptual knowledge data available. It may be possible that the effect of self-explanation would be different for procedural and conceptual knowledge. In fact, past work has illustrated differential self-explanation effects for these two knowledge types (Berthold & Renkl, 2009). One must also consider that most all of the studies were implemented by researchers, and it is unclear if these same effects would be found in a less controlled setting. Finally, several references that could have been included based off of the eligibility criteria could not be found, and they could affect the overall mean effect size.

This meta-analysis illustrates that there is some evidence for the benefits of self-explanation in math; however, the evidence is not as strong as some researchers might think. It seems that self-explanation is not harmful to students, but whether or not it is worth the time it takes to go through them is unclear. Future studies may want to include a control group that does not self-explain so that a direct comparison can be made to test for the self-explanation effect. Many past studies did not include a control group and could not be included in this meta-analysis. Also, future research should continue to examine this self-explanation effect because it seems likely that there are moderators that have yet to be accounted for. Additional research could help identify and test such moderators. Further research is also needed to make conclusions about the effect of self-explanations on conceptual knowledge and in other learning domains. Due to the large number of self-explanation studies in science domains, it would be beneficial for a future meta-analysis to be done examining the effectiveness of self-explanation in the sciences. Self-explanation shows potential for being a useful instructional strategy, and future research will examine the intricacies of self-explanation more closely to see if self-explanation should be incorporated into more classrooms in the future.

Appendices

Not included in page count.

Appendix A. References

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26*(2), 147-179.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology, 95*(4), 774-783.

Berthold, K., Eysink, T. H. S., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science, 37*(4), 345-363.

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology, 101*(1), 70-87.

Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 161-238).

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439-477.

Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction, 16*(2), 104-121.

Große, C. S., & Renkl, A. (2006). Effects of multiple solution methods in mathematics learning. *Learning and Instruction, 16*(2), 122-138.

Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction, 17*(6), 612-634.

Hilbert, T. S., Renkl, A., Kessler, S., & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction, 18*(1), 54-65.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical Meta-Analysis*. Thousand Oaks, CA: SAGE Publications.

Lombrozo, T. (2006). The structure and function of explanations. *Trends in cognitive sciences, 10*(10), 464-470.

Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology, 104*, 1-21.

Mitrovic, A. (2005). Scaffolding answer explanation in a data normalization tutor. *Facta universitatis-series: Electronics and Energetics, 18*, 151–163.

Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction, 16*(2), 173-199.

Nathan, M. J., Mertz, K., & Ryan, R. (1994, April 8). *Learning through Self-Explanation of Mathematics Examples: Effects of Cognitive Load.* Paper presented at the Annual Meeting of The American Educational Research Association, New Orleans.

Rittle-Johnson, B. (2006). Promoting Transfer: Effects of Self-Explanation and Direct Instruction. *Child Development, 77*(1), 1-15.

Rittle-Johnson, B., Saylor, M., & Swygert, K. E. (2008). Learning from explaining: Does it matter if mom is listening? *Journal of Experimental Child Psychology, 100*(3), 215-224.

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: Influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology, 101*(4), 836-852.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Garnott & J. Parziale (Eds.), *Microdevelopment: A process-oriented perspective for studying development and learning* (pp. 31-58). Cambridge, MA: Cambridge University Press.

Tajika, H., Nakatsu, N., Nozaki, H., Neumann, E., & Maruno, S. (2007). Effects of self-explanation as a metacognitive strategy for solving mathematical word problems. *Japanese Psychological Research, 49*(3), 222-233.

Wong, R. M. F., Lawson, M. J., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction, 12*(2), 233-262.

Appendix B. Tables and Figures

Table 1

Characteristics of the Studies in the Meta-Analysis

Variable	# of Studies	Variable	# of Studies
Assessment Characteristics			
Knowledge type		*Gender (percent male)*	
Procedural	8	0 to 20	2
Conceptual	0	20 to 40	3
Procedural and Conceptual	10	40 to 50	7
Number of Assessment Items		> 50	2
1 to 10	6	Unknown	4
11 to 20	8	**Other Characteristics**	
> 20	3	*Setting*	
Unknown	1	Lab	6
Measure Creation		Classroom	12
Experimenter-Created Measure	18	*Country*	
Standardized Measure	0	USA	9
Participant Characteristics		Germany	5
Number of Participants		Australia	2
0 to 40	5	New Zealand	1
41 to 80	9	Japan	1
> 80	4	*Mathematics Domain*	
Grade Level		Geometry	4
preK to 2nd	1	Probability	6
3rd to 5th	3	Other (algebra, math equivalence, etc.)	8
6th to 8th	1	*Publication Type*	
9th to 12th	5	Journal	17
College or higher	8	Conference Paper	1
Age		*Publication Year*	
5 to 10 years	3	1994 to 2000	2
10 to 15 years	2	2001 to 2005	6
15 to 20 years	4	2006 to 2010	10
> 20 years	3		
Unknown	6		

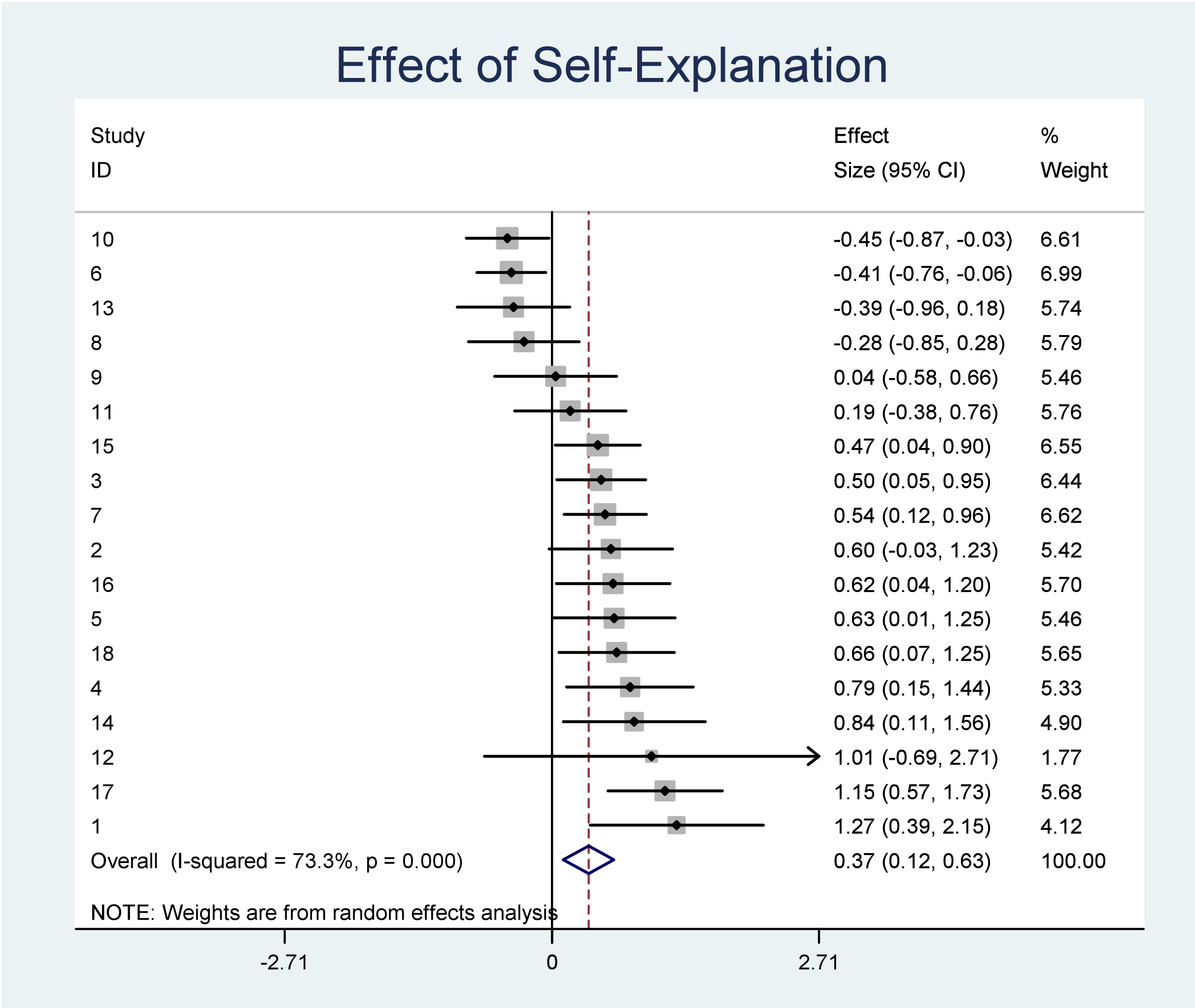

Figure 1. Mean effect sizes and 95% confidence intervals for the effects of self-explanation