Support for struggling students in algebra: Contributions of incorrect worked examples

Christina Barbieri
University of Delaware

Julie L. Booth
Temple University

Abstract

Middle school algebra students (N = 125) randomly assigned within classroom to a Problem-solving control group, a Correct worked examples control group, or an Incorrect worked examples group, completed an experimental classroom study to assess the differential effects of incorrect examples versus the two control groups on students' algebra learning, competence expectancy, and sense of belonging to math class. The study also explored whether prior knowledge impacted the effectiveness of the intervention. A greater sense of belonging and competence expectancy predicted greater learning overall. Students' sense of belonging to math and competence expectancies were high at the start of the study and did not increase as a result of the intervention. A significant interaction between prior knowledge and incorrect worked examples on post-test scores revealed that students with low prior knowledge who struggle with learning math benefit most from reflecting on highlighted errors within an incorrect worked examples intervention. The unique contributions of these findings as well as educational implications are discussed.

Keywords: learning from errors; incorrect examples; worked examples; algebraic problem-solving; sense of belonging; competence expectancy

Corresponding author:
Christina Barbieri
barbieri@udel.edu
University of Delaware
School of Education
113 Willard Hall Education Building
Newark, DE 19716, USA
(302) 831-8695

Full Citation: Barbieri, C. & Booth, J.L. (2016). Support for struggling students in algebra: Contributions of incorrect worked examples. *Learning and Individual Differences, 48,* 36 – 44.

Funding: Research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through Grant R305A100150 to Strategic Education Research Partnership Institute. Continued work was supported by Grant R305B130012 from the Institute of Education Sciences, U.S. Department of Education to the University of Delaware. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

1. Introduction

Algebra I, commonly taken in the ninth or eighth grade, is often considered a gatekeeper course for higher level math and science courses (Matthews & Farmer, 2008), which influences college admissions (Schneider, Swanson, & Riegle-Crumb, 1998) and majoring in a Science, Technology, Engineering, or Math (STEM) field (Chen, 2009). Unfortunately, there is a decline in mathematics achievement and achievement motivation for mathematics during adolescence, which both impact students' likelihood of success in mathematics (Wang & Pomerantz, 2009). According to the 2012 Programme for International Student Assessment (PISA), adolescents in the United States ranked below average in mathematics performance compared to other nations, and 50% of US students reported low levels of interest in mathematics. (Kelly, Xie, Nord, Jenkins, Chan, & Kastberg, 2013). Thus, innovative work is needed to improve student learning in mathematics at this critical time.

Researchers have traditionally set out to improve learning in one of two ways: *cognitive* interventions, which aim to design instruction that is more suitable for students' cognitive capabilities (Sweller, 2012), or *motivational* interventions, which aim to increase student engagement or alter beliefs to increase the effectiveness of traditional instruction. However, focusing an intervention solely on cognition or motivation may not lead to increases in the other outcome. Thus, in domains where both achievement and motivation are critical to success—like Algebra—a combined approach may be beneficial. In the present study, we use an intervention that alters mathematics instruction to be both more *cognitively* and *motivationally* relevant: error reflection.

1.2 Learning from Error Reflection

Errors are a daily occurrence in schoolwork, but teachers differ greatly in how they handle students' errors (Schleppenbach, Flevares, Sims, & Perry, 2007). In many U.S. classrooms, consideration of errors is discouraged because of the belief that this may reinforce incorrect procedures or faulty knowledge (Santagata, 2004; Stigler & Perry, 1988). In contrast, errors are frequently discussed in Japanese classrooms and thought to be integral for learning (Stigler & Hiebert, 1999).

Having students explain incorrect examples (i.e., examples of fictitious students' errors on problems) is beneficial to learning (Booth, Lange, Koedinger, & Newton, 2013; Durkin & Rittle-Johnson, 2012). Potential cognitive and motivational benefits are discussed next.

1.2.1 Potential Cognitive Benefits of Incorrect Examples

Compared to problem-solving alone, studying *correct* worked examples during practice has been found to boost problem-solving skills (e.g., Sweller & Cooper, 1985; Zhu & Simon, 1987). Additional benefits of reflecting on *incorrect* examples include: (1) allowing students to accept that the strategies they may use are wrong (Siegler, 2002), and (2) drawing student attention to components that make the solution incorrect (Booth et al., 2013). Practice sessions involving both correct and incorrect examples yielded increases in both knowledge about problem features and procedural skills (Booth, Cooper, Donovan, Huyghe, Koedinger, & Pare-Blagoev, 2015). However, prior knowledge may impact the cognitive benefit of incorrect examples; some studies find benefits only for those who have already developed a strong foundational knowledge of the to-be-learned material (Heemsoth & Heinze, 2014) while others find equal benefits for low and high prior knowledge students (Adams, McLaren, Durkin, Mayer, Rittle-Johnson, Isotani, & van Velsen, 2014) or greater benefit for students with low prior knowledge (Booth et al., 2015). Thus, further study is needed to compare benefits of incorrect worked examples for students with varying levels of prior knowledge. This is a key purpose of the present study.

1.2.2 Potential Motivational Benefits of Incorrect Examples

In the present study, we focus on two motivational constructs—student beliefs about their own competence (Cho, Weinstein, & Wicker, 2011) and their sense of belonging to a math community (Good, Rattan, & Dweck, 2012)—both important for math learning and likely to decline during adolescence (Anderman, 2003; Wigfield, Eccles, MacIver, Reuman, & Midgley, 1991). These outcomes were chosen because a) the decline in competence beliefs occurs around the time that students are transitioning into algebra (Wang & Pomerantz, 2009), and b) the idea that one either is or is not a "math person" is prevalent in US culture (Rattan, Good & Dweck, 2012). As both are likely to lead to decreased math performance and fewer students pursuing STEM degrees, the current study focuses on how reflection on errors may relate to students' competence beliefs and feelings of belonging in math class.

Competence beliefs are an important component in many theories of achievement motivation (Bong & Skaalvik, 2004; Ferla, Valcke, & Cai, 2009; Pintrich & Schunk, 1996). We utilize Elliot and Church's (1997) terminology--competence expectancy—to refer to students' confidence in their ability to accomplish a task[1]. Competence beliefs are influential on learning in many content domains (see Eccles & Roeser, 2009, for a review)--including mathematics (Cho et al., 2011). Interventions which increase competence expectancy may lead to improvements in achievement, as students who feel competent engage in adaptive learning

---

[1] This construct is similar to Eccles and Wigfield's construct of *expectancies for success* (Eccles et al., 1993; Eccles & Wigfield, 1995; Wigfield & Eccles, 2002). Both constructs address students' expectations for future success, however, a major difference is that Eccles' measure asks students to make comparisons of their abilities to others and Elliot's measure does not. Early work in development of competence belief scales assumed that theoretical differences between competences beliefs and expectancies should result in distinct scales (Eccles et al., 1983), later work demonstrated that many of these measures load on the same factor (Eccles & Wigfield, 1995; Eccles et al., 1993).

behaviors, while students with low or threatened competence beliefs engage in maladaptive behaviors (Dweck, 1986). Wigfield and Eccles (2002) explain that students develop a more accurate view of their own competence as they become more aware of their own successes and failures during adolescence; thus, if feedback about successes and failures is relevant to students' developing sense of their competence, and algebra is a domain in which many students struggle, altering the way students view successes and failures may also influence their expectancies for success. However, students react differently to failure (Tulis & Ainley, 2011). Those who view mistakes positively and see feedback as a learning opportunity experience more positive emotions after failure. Therefore, students may vary greatly in whether they benefit from studying their own errors. Thus, we propose that confronting errors by studying incorrect examples (i.e., errors of a fictitious student) may increase focus on errors as learning tools rather than reflections of ability, potentially normalizing errors as part of learning and increasing students' expectations for their own competence.

Sense of belonging to mathematics is the feeling that one is a member of the math community (Good, et al., 2012), comprising feelings of acceptance, positive affect, trust, and the desire to actively participate. Sense of belonging to the *school* community has been established as influencing students' *school* experience (for a review, see Osterman, 2000); however, the influence of sense of belonging to *mathematics* on algebra learning has not been explored (but see Good et al., 2012). Failure experiences may lower feelings of belonging in that domain, and as algebra is a troubling topic for many students, failure experiences may be especially common. An error reflection intervention may foster the idea that errors are made by many students and thus are a normal part of learning math. If making errors is not viewed as an indicator of one not being a ‘math person’, all students can feel like they are a valuable part of their math class community.

2. The Present Study

The present study tests the effect of reflecting on errors in the form of incorrect examples on both cognitive and motivational factors for middle-school students learning Algebra. To isolate the effects of error reflection during problem-solving practice we compare the intervention to two different conditions: one practice-problem condition, in which no worked examples are included, and one condition in which only correct examples are included. The control condition (i.e., practice only) allows us to determine how inserting the study of incorrect examples into typical classroom practice may influence cognitive and motivational growth. The correct worked examples condition allows us to ensure that any observed benefits are not due simply to the inclusion of worked examples but instead to the experience of specifically reflecting upon and explaining *incorrect* solutions. Comparing incorrect worked examples to both correct worked examples and traditional problem-solving all within regular instruction allows us to consider the influence of the problems and examples themselves, controlling for classroom influences. The current study utilized within-class assignment, reducing potential teacher or classroom effects.

2.1 Research Questions

The present study has four research questions: (1) Does having students reflect on and explain incorrect worked examples during problem-solving practice benefit algebra learning overall? (2) Does studying and explaining incorrect worked examples increase students’ competence expectancy and sense of belonging to math? (3) Does prior knowledge moderate the effectiveness of the incorrect worked examples on post-test scores? (4) Finally, if there are boosts in students' competence expectancy or sense of belonging to math, do these increases partially explain any potential moderation of the effect of incorrect example on post-test scores by students' prior knowledge?

2.1.1 Hypotheses

Hypothesis 1: Studying and explaining incorrect worked examples will improve student learning overall.

Hypothesis 2: Working with incorrect worked examples will increase students' competence expectancy and sense of belonging to math.

Hypothesis 3: Incorrect worked examples will be especially beneficial for students with low prior knowledge in refining algebraic problem-solving as measured by post-test scores.

Hypothesis 4: If the hypothesized boosts in competence expectancy or sense of belonging to math are found, as well as an interaction between prior knowledge and incorrect worked examples on learning, increases in these motivation constructs may partially explain the moderating role of prior knowledge.

As previously explained, studying incorrect worked examples may normalize errors as part of the learning process, hence increasing students' competence expectancies and sense of belonging to math. This should be particularly beneficial to students with low prior knowledge who may be more likely to experience lowered feelings of belonging and competence expectancy in algebra. Testing this mediated moderation involves testing several paths (see Muller, Judd, & Yzerbyt, 2005) including testing whether prior knowledge moderates the effect of incorrect examples on post-test scores, testing whether prior knowledge moderates the effect of incorrect examples on changes in motivation, and testing whether changes in motivation mediate the moderating effect of prior knowledge on post-test scores by incorrect examples. This proposed mediated moderation is displayed in Figure 1 with the key paths in bold.

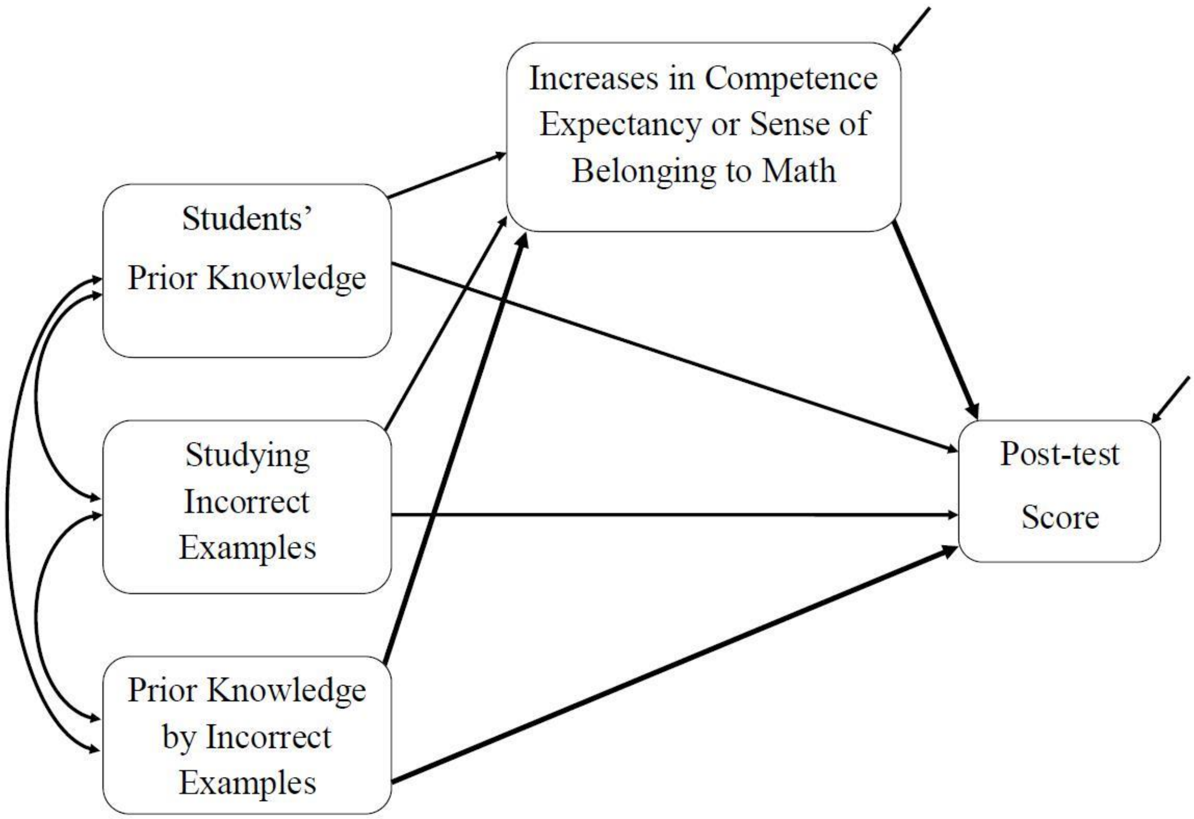


Figure 1. Proposed mediated moderation

Preconditions that must be met to test this hypothesis are that (1) competence expectancy and sense of belonging to math must be related to learning and (2) there must be change in motivation from Time 1 to Time 2 to address whether this change mediates the assumed moderating effect of prior knowledge on the influence of incorrect examples on learning.

3. Methods

3.1 Participants

Participants were middle school students (N = 140; 57 males; 97% 8th grade; 10.9% low income) from five Algebra I classrooms. Students were classified according to membership in racial and ethnic populations that are underrepresented in the domain of mathematics: Underrepresented minority (URM) students (African-American, Hispanic, and bi-racial; 23%) or non-URM (White and Asian; 77%) students. Students who did not complete all measures (N=15) were excluded from analysis.

3.2 Procedure

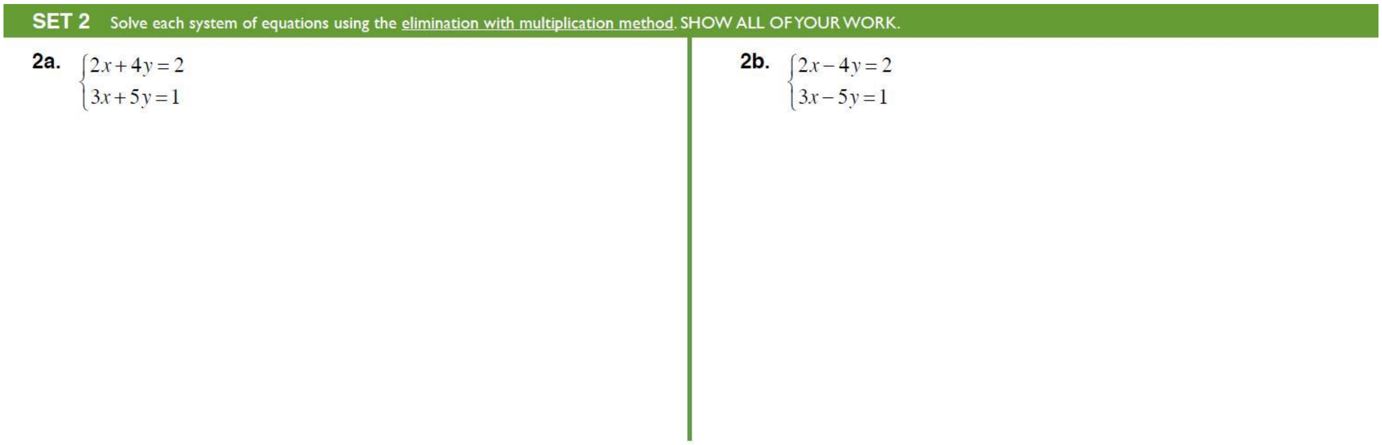
Students were randomly assigned within classroom to receive worksheets with incorrect worked examples (n = 40), worksheets with correct worked examples (n = 44), or control worksheets (n = 41) during their unit on solving systems of equations. At the start of the study (Time 1), students completed a motivation survey, followed by a pre-test during the next school day. Students completed four worksheets individually over a period of five to seven weeks[2]. Teachers allowed students approximately 20 minutes to complete each worksheet and did not provide incentive for completion. The order of worksheets administered varied by classroom to coincide with each teacher's lesson plans. The motivation survey was re-administered after the completion of the four worksheets (Time 2). A post-test was administered at the completion of the study. As within-class assignment was used, the only factor that varied systematically between conditions was the type of worksheets received.

3.3 Materials

Four topics were covered by the worksheets utilized for this study, each requiring students to find the values of a set of unknowns that satisfy a pair of equations utilizing a different method (graphing, substitution, elimination with addition and subtraction, and elimination with multiplication). Each worksheet contained eight problems; differences between conditions were only in what type of activity the worksheet required students to do with those problems. In the control condition, students were instructed to solve each of the four problem

[2] One accelerated Algebra 1 course completed the systems of equations unit in only two weeks.

pairs for a total of eight problems per worksheet. A sample control problem pair is displayed in Figure 2 below.



SET 2 Solve each system of equations using the elimination with multiplication method. SHOW ALL OF YOUR WORK.

2a. $\begin{cases} 2x+4y=2 \\ 3x+5y=1 \end{cases}$

2b. $\begin{cases} 2x-4y=2 \\ 3x-5y=1 \end{cases}$

Figure 2. Sample Problem Pair from control condition worksheets

In the two example conditions, students were given four example-problem pairs that included examples of fictitious students' correct or incorrect work. Students were asked to study the examples and then respond to self-explanation prompts designed to direct their attention to important concepts that students often struggle with in algebra (e.g., understanding of equal sign as indication of balance of equations), before solving the "Your Turn" problem paired with the example. Students studied four examples and solved four "Your Turn" problems per worksheet. Self-explanation prompts in the Correct and Incorrect conditions were similar in that they targeted the same concepts, regardless of the correctness of the example. In the correct worked examples condition, the correct procedure was carried out and the correct version of the key concept was highlighted within the prompt (Figure 3).

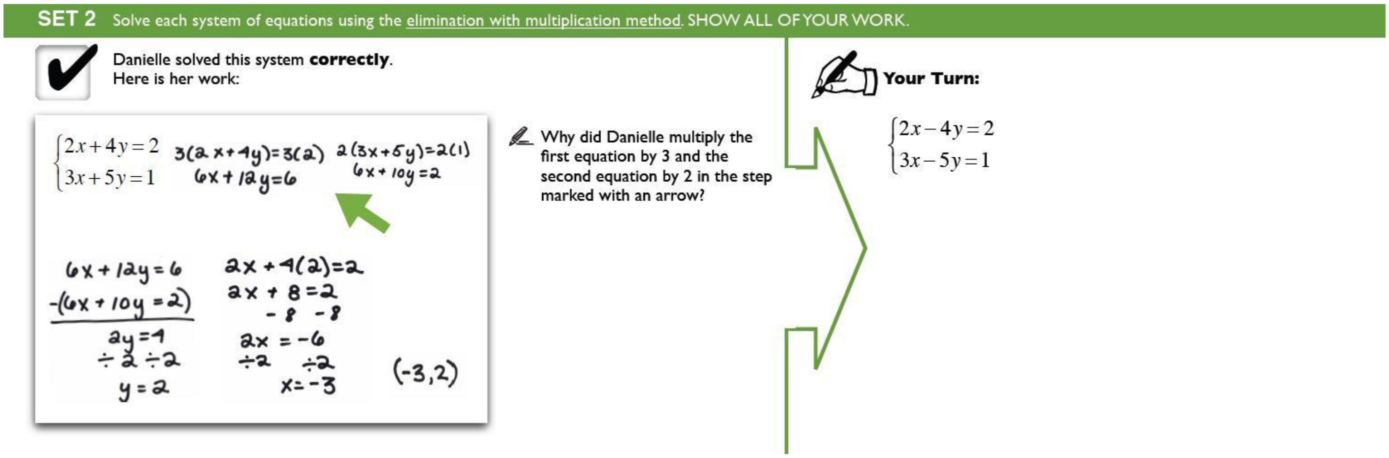


Figure 3. Sample Example-Problem Pair from correct example condition worksheets

In the incorrect worked examples condition, the incorrect procedure was displayed and highlighted as one would expect a student that held the corresponding misconception to solve the problem (Figure 4).

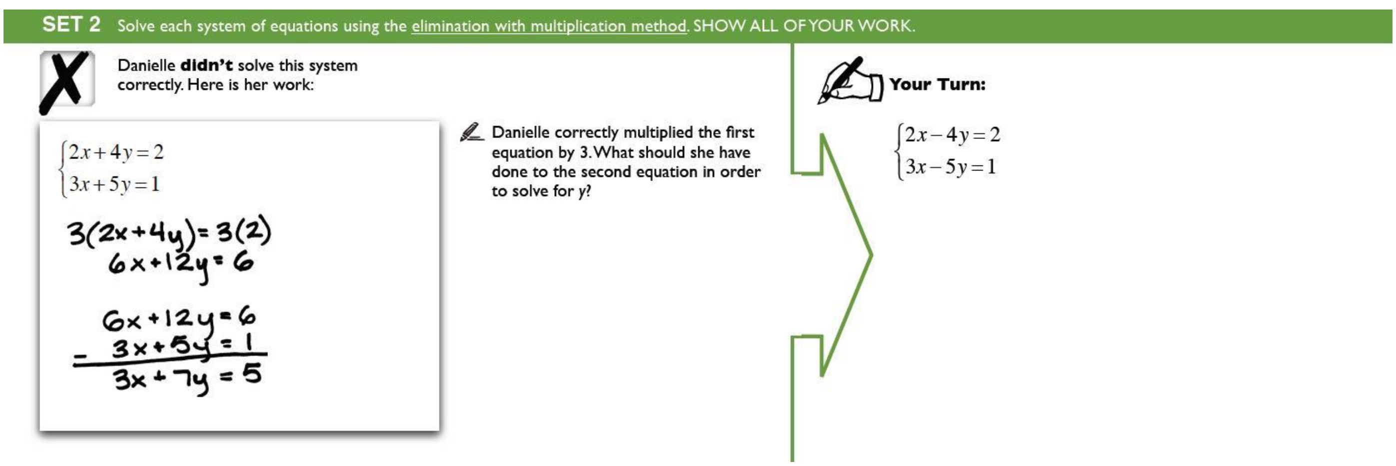


Figure 4. Sample Example-Problem Pair from incorrect example condition worksheets

Thus, the control group was asked to solve a total of 32 problems whereas the example groups were asked to explain 16 examples and solve 16 “Your Turn” problems. Participants attempted M = 13.74 of the 16 “Your Turn” problems overall; attempts did not differ by condition [$F(2,122) = 1.670$, $MSE = 9.879$, $p = .195$, $\omega^2 = .01$]. Participating teachers reported good alignment of the worksheet contents with their course material.

The motivation survey administered prior to the pre-test and post-test measured students' competence expectancy in algebra class and their sense of belonging to their math class community.

3.3.1 Competence expectancy

Competence expectancy was measured using a previously validated scale (Elliot & Church, 1997). Using a Likert scale, students indicated level of agreement on two statements: "I believe I will get an excellent grade in this class" and "I expect to do well in this class." The scale ranged from 1 (*No, Not at all*) to 7 (*Yes, definitely*). Students were instructed to respond to the statements using their thoughts about algebra at the time of the study. Scores were computed by averaging the items. Higher means represent higher competence expectancy. Internal consistency was sufficient (α = .922).

3.3.2 Sense of Belonging

Sense of belonging was measured using five statements selected from the original sense of belonging scale[3] (Good et al., 2012). Using a Likert scale, students indicated level of agreement about their sense of belonging in their math class. Items included: "I feel like I am part of the math community", "I feel accepted", "I feel comfortable", "I trust my teachers to help me learn", and "I enjoy participating". The scale ranged from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). Scores were computed by averaging the items. Higher means represent a stronger sense of belonging to math. Internal consistency was sufficient (α = .856).

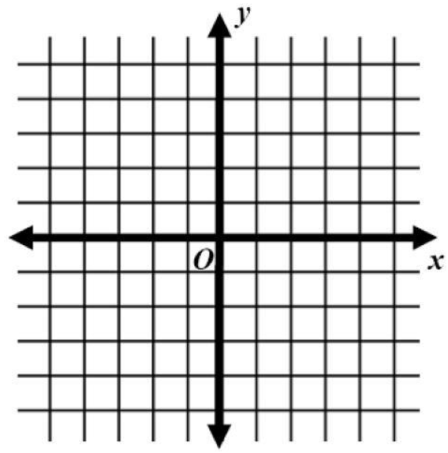
---

[3] Items were reworded from the original scale to meet the reading level of middle school students, as the original scale was used with undergraduates.

1. Solve the system by graphing:

   y=3x-7

   y=x+1



2. Look at each system of equations. <u>Without solving the systems</u>, circle the choice that describes the solution(s) each system will have.

y=x+2 y=2x+2	None	1 (at origin)	1 (not at origin)	Infinite
y=3x+2 y=3x-4	None	1 (at origin)	1 (not at origin)	Infinite
y=x y=x+1	None	1 (at origin)	1 (not at origin)	Infinite
y=3x y=-1/3x	None	1 (at origin)	1 (not at origin)	Infinite

Figure 5. Sample Assessment Items

3.3.3 Content knowledge test

Pre- and post-tests included a total of 25 items that assessed students' knowledge about solving systems of equations. These items targeted students' understanding of features in a problem (e.g., differences between variables and constants) and ability to solve problems (e.g., solving a system of equations using the substitution method). To avoid ceiling effects at posttest and ensure that we could distinguish between different levels of content knowledge, items were generally designed to be harder than those students would see on teacher-developed exams. Sample questions are displayed in Figure 5. Scores were computed as the percent of items answered correctly; internal consistency was acceptable (α = .763).

3.4 A priori and posthoc power analyses

Using G*Power (Faul, Erfelder, Lang & Buchner, 2007), a priori and post hoc power analyses were conducted to determine the minimum sample size to detect a significant effect on learning and motivation using α = .05. While a prior analyses suggested a sample size of N = 107 would provide power of .80 to detect a small-medium sized effect on learning, post hoc analyses revealed that the actual sample of N = 125 was slightly underpowered, providing power of only .70. Observed power to detect a small effect for the motivation variables was high at .95.

4. Results

4.1 Plan of analysis

Dummy coded variables were created for condition (Correct: 1 = Correct Examples Condition; Incorrect: 1 = Incorrect Examples Condition), URM status (1 = URM), and SES (1 = low-SES). Intra-class correlations revealed that post-test scores were related to cluster (ICC=.169). The current sample size did not allow for multilevel modeling; however, comparison of study variables by classroom revealed no significant differences. Implications are considered in the discussion.

Table 1.
Student cognitive, motivation, and demographic factors by condition

	Total Sample (N = 125)		Problem-Solving Group (n = 41)		Correct Worked Examples (n = 44)		Incorrect Worked Examples (n = 40)	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Pre-test Score	26.56%	14.59	29.27%	16.98	27.91%	12.40	22.30%	13.52
Post-test Score	48.83%	19.08	50.34%	20.53	47.91%	19.40	48.30%	17.53
Competence Expectancy (T1)	5.70	1.34	5.77	1.37	5.80	1.33	5.51	1.32
Competence Expectancy (T2)	5.49	1.43	5.48	1.51	5.77	1.26	5.19	1.53
Belonging (T1)	4.73	1.01	5.03	0.87	4.67	0.90	4.50	1.19
Belonging (T2)	4.66	1.00	4.72	1.10	4.74	0.84	4.53	1.07
Low SES	9.6%		7.3%		6.8%		15.0%	
URM	22.4%		22.0%		20.5%		25.0%	

To determine whether incorrect worked examples benefitted algebra learning (RQ1) we conducted a 3(condition) x 2(time) repeated measures analysis of variance (ANOVA) on post-test scores; we conducted parallel 3 x 2 ANOVAs on competence expectancy and sense of belonging to examine benefits for motivation (RQ2). To answer RQ3, linear regressions predicting post-test scores were conducted, with all Time 1 and Time 2 scores grand-mean centered to avoid issues of multicollinearity (see Table 1 for sample means prior to centering). All scores were normally distributed.

4.2 Testing main effects of condition

As shown in Table 2, participants demonstrated improvement from pre- to post-test[4] but this did not vary by condition providing no supporting evidence for Hypothesis 1.

[4] Improvement occurred across all conditions, but because content knowledge items were designed to be more difficult than typical classroom tests, posttest scores are lower than would be expected on teacher-developed exams.

Table 2.
3 x 2 RMANOVA on pre-test and post-test scores by condition

Source	*df*	*F*	*p*	*MSE*	ω_p^2
Condition	2,122	1.040	.357	398.078	.001
Time	1,122	175.129	<.001	178.102	.590
Condition x Time	2,122	1.182	.310	178.102	.003

As shown in Table 3, there was a slight but significant decline in competence expectancy from T1 (*M* = 5.70, *SD* = 1.34) to T2(*M* = 5.49, *SD* = 1.43) which did not vary by condition. However, participants seemed to maintain a high level of competence expectancy throughout the study.

Table 3.
3 x 2 RMANOVA on Time 1 and Time 2 competence expectancy by condition

Source	*df*	*F*	*p*	*MSE*	ω_p^2
Condition	2,122	1.162	.316	3.444	.010
Time	1,122	7.027	.009	.405	.054
Condition x Time	2,122	1.453	.238	.405	.007

As shown in Table 4, there was a significant interaction between condition and time on sense of belonging; follow-up paired samples t-test with Bonferroni correction revealed that the Problem-solving Control condition reported a reduced sense of belonging from Time 1 (*M* = 5.03, *SD* = .87) to Time 2 (*M* = 4.72, *SD* = 1.10; *t*(40) = 2.397, *p* = .021, *d* = .31), while students in the Incorrect condition ($M_{Time\ 1}$ = 4.50, *SD* = 1.19; $M_{Time\ 2}$ = 4.53, *SD* = 1.07) and the Correct condition ($M_{Time\ 1}$ = 4.67, *SD* = .90; $M_{Time\ 2}$ = 4.74, *SD* = 8.4) did not [*t*(39) = -.322, p = .749, *d* = .02; *t*(43) = -.807, p = .424, *d* = .01 respectively]. As such, Hypothesis 2 which proposed particular motivational benefits of incorrect examples was not supported. It is important to note that observed power to detect a main effect of condition on learning was lower than expected power but power to detect an effect on motivation was sufficient. Results are interpreted with consideration to power in the discussion section.

Table 4.
3 x 2 RMANOVA on Time 1 and Time 2 sense of belonging to math by condition

Source	*df*	*F*	*p*	*MSE*	ω_p^2
Condition	2,122	1.529	.221	1.751	.008
Time	1,122	1.271	.262	.248	.010
Condition x Time	2,122	3.802	.025	.248	.043

4.3 Exploring the role of prior knowledge

Table 5 displays the results of two multiple regressions predicting post-test scores. The first model was conducted to establish the relationship between post-test scores and the motivation variables of competence expectancy and sense of belonging at Time 1, controlling for pre-test scores, demographic variables, and condition. Post-test scores were regressed onto pre-test scores, student's URM status and SES, Time 1 competence expectancy (CE1), Time 1 sense of belonging to math (BL1), and dummy-coded condition variables for Correct and Incorrect. The model reveals no main effect of condition on learning, but CE1 and BL1 predicted post-test scores.

The second model tested whether the effect of condition varied by prior knowledge. Two interaction terms were created between centered pre-test scores and the Correct and Incorrect dummy variables. This moderation model revealed a significant Prior Knowledge by Incorrect condition interaction ($\beta = -.182$, $p = .045$).

Table 5.
Prediction models of post-test score

	Post-test Score			
	Main Effects Model		Moderation Model	
	β	*B* (*SE*)	β	*B* (*SE*)
Pre-test Overall	.26**	.34 (.10)	.34**	.44 (.14)
Low SES	-.06	-3.95 (4.84)	-.06	-4.03 (4.78)
URM	-.03	-1.32 (3.26)	-.05	-2.20 (3.24)
T1 Competence Expectancy (CE1)	.40**	5.72 (1.26)	.40**	5.66 (1.25)
T1 Belongingness (BL1)	.19*	3.63 (1.67)	.21*	4.03 (1.66)
Correct	-.02	-.84 (3.21)	-.02	-.619 (3.20)
Incorrect	.10	4.06 (3.35)	.08	3.14 (3.34)
Pre-Overall x Correct			.02	.04 (.22)
Pre-Overall x Incorrect			-.18*	-.44 (.22)
R^2	.454		.478	
ΔR^2 from main effects model			.023	
F	13.92**		11.68**	

* $p < .05$ level. ** $p < .01$ level.

To test for simple slopes, the referent group was changed to the correct examples condition and a parallel model was run. This included a dummy code and interaction term for the Problem-Solving control group by Pre-test scores which replaced the corresponding variables for the Correct example control group in Model 2. The results were consistent with Model 2 and the interaction term between Incorrect example and pre-test remained a significant predictor of post-test scores ($\beta = -.199$, $p = .05$). To explicate the significant interaction, follow-up split regressions were conducted based on condition using the same variables included in the main effect model (Model 1). Post-test scores were regressed onto pre-test scores controlling for demographic and T1 motivation variables. Results demonstrated that while prior knowledge positively predicted post-test scores in both the control group ($\beta = .342$, $p = .003$) and the correct examples group ($\beta = .301$, $p = .023$), it did not predict post-test scores for the incorrect examples group ($\beta = .039$, $p = .793$).

These results suggest that students with low prior knowledge benefitted more from the incorrect worked examples condition than the correct and problem-solving control groups. Therefore, Hypothesis 3 was supported. Students with high prior knowledge benefitted equally from all three conditions. This moderation is displayed in Figure 6 which displays the hypothetical post-test score of a student who was one standard deviation above, below, and at the mean on the pre-test.

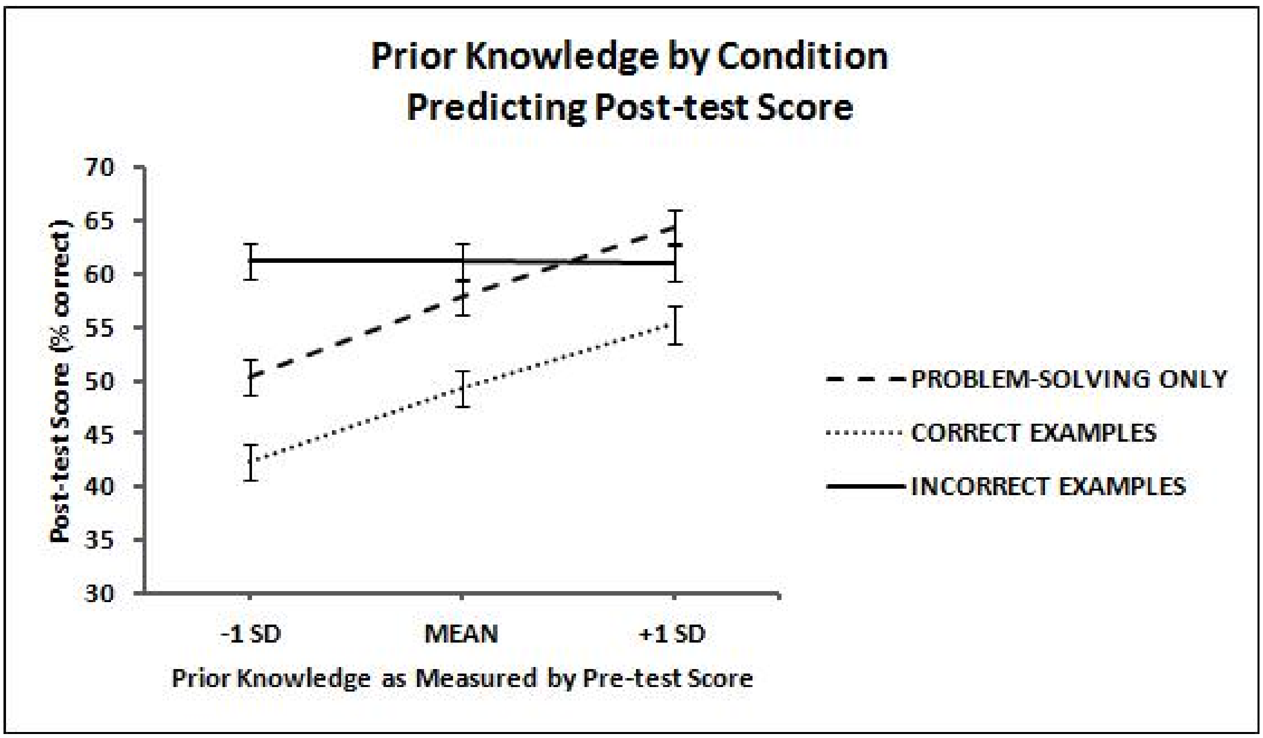


Figure 6. *Prior knowledge by condition predicting post-test score*

4.4 Proposed mediated moderation

The final research question was posed to assess whether increases in students' competence expectancy or sense of belonging to math partially explain the moderation of the effect of incorrect examples by prior knowledge. Hypothesis 4 stated that incorrect examples may be particularly beneficial for students with low prior knowledge due to fostering the idea that errors are a normal part of the learning process. This was expected to result in increases in competence expectancy and sense of belonging to math. However, because no increases in motivation were found, the proposed mediated moderation model[5]cannot explain the moderation of the effect of incorrect examples by prior knowledge on post-test scores.

5. Discussion

Our findings indicate that error reflection was equally beneficial for students across levels of prior knowledge, but more effective than correct worked examples or problem-solving alone for those who began the study with little knowledge on systems of equations. This is particularly important, as researchers and practitioners adamantly seek methods to improve learning for the many students who struggle with Algebra I. Still, the underlying mechanism for this treatment by aptitude interaction is unclear. Future research may explore differential benefits of incorrect examples based upon different types of prior knowledge. The assessment utilized in the current study measured a composite of conceptual and procedural skills but did not provide a more refined measure of prior knowledge. Different types of prior knowledge such as conceptual and procedural knowledge, encoding, and algebraic misconceptions should be examined. Booth and Davenport (2013) emphasize the impact that prior knowledge has on students' ability to encode relevant features of algebraic equations for successful problem-solving. Thus, one possible

[5] Although it was statistically unlikely to find the proposed mediated moderation due to the lack of increase in the motivation variables, structural equation models, as suggested in Figure 1, were conducted and no suggestion of mediated moderation was found.

explanation is that highlighting common errors in incorrect examples provides the support students need to fully benefit from worked examples on procedural knowledge in particular. This is consistent with Groβe & Renkl's (2007) finding that highlighting errors within incorrect worked examples enabled students with low prior knowledge to benefit from an otherwise difficult task. Another possibility is that highlighting errors that represent an underlying algebraic misconception may reduce the prevalence of that misconception and allow students to more fully benefit from instruction.

While competence beliefs have commonly been tied to achievement (Elliot, 2005; Elliot & Dweck, 1988), they have seldom been assessed in the context of a cognitively-based intervention on learning. Further, research on domain-specific sense of belonging is sparse. The findings from the current study support that sense of belonging to *mathematics* is indeed a contributor to algebra learning for middle school students, along with competence expectancy. However, studying incorrect examples over the course of one Algebra I unit did not result in significant changes in competence expectancy or sense of belonging to math, as initially hypothesized. One possible explanation is that the incorrect examples did not provide a salient message that errors are useful for learning. Beliefs and attitudes towards math, developed over an extended schooling experience, may require a more conspicuous intervention aimed at changing the classroom climate to one that is open and welcoming of mistakes as learning tools. Altering the intervention to target perspectives of mathematics that influence the development of sense of belonging and competence expectancy may provide a robust effect.

In contrast to hypothesis 2, slight declines were found in competence expectancy across all conditions. One potential explanation is that students may have initially had overly high expectations of competence (García, Rodríguez, González-Castro, González-Pienda, & Torrance,

2015) and became more accurate in their expectancies as the school year progressed (e.g., Wigfield & Eccles, 2002). As calibrated ability beliefs have direct effects on middle school mathematics performance (Chen, 2003), slight declines in competence expectancy may not be cause for concern. However, as competence expectancy was only measured at two time points, this speculation must be confirmed with further research.

Unfortunately, as increases in these motivation constructs were not found, we were unable to test the hypothesized mediation of the moderating effect of prior knowledge on the effect of incorrect examples on learning. Further, while the current study proposed that a normalization of errors may alter ability beliefs and sense of belonging to math, attitudes towards errors were not measured. Measuring students' perspectives of errors as a learning tool while working with incorrect examples may provide a clearer interpretation of these results. Further, other achievement motivation constructs may relate more closely to learning from errors. For example, it is possible that task specific self-efficacy or views of the malleability of intelligence may relate to their perceptions of how useful errors are for their own learning which may influence the effectiveness of an intervention that relies on error reflection. This may be a fruitful area for future research.

The current study has certain methodological strengths. The random within-classroom assignment to conditions is a design feature that is not commonly possible in education research but that serves to reduce potential teacher influences. Other strengths include the lack of differences in demographic, motivation, or prior knowledge by classroom that are often found in educational research. However, there are certain methodological limitations to the current study. As we did not have access to classroom learning measures, we do not currently know whether the learning experienced over the course of the study transferred to classroom learning.

However, classroom teachers deemed the assessment as well aligned with what is normally used to measure learning of solving systems of equations in Algebra I. Also, due to the limited sample size, we could not account for the nested structure of the data with multilevel modeling; although no differences were found by classroom on key study variables, further research is needed with a larger sample size that can account for nesting in the data and extend the findings of the current study. Lastly, future research may explore whether a combination of incorrect and correct worked examples would be superior to either type of example alone in terms of supporting motivation or learning. .

Our finding that being in the correct or incorrect worked examples groups did not have an overall influence on learning seems inconsistent with previous worked examples studies conducted in laboratory settings (e.g., Sweller & Cooper, 1985; Booth et al., 2013). Though a priori power analyses suggested our sample size would be sufficient to detect significant findings with a small-medium effect size, observed power was lower than expected. A larger sample size may have indeed yielded significant effects of condition.

Nevertheless, the lack of significant effect of worked examples *is* consistent with findings that laboratory-proven principles do not always lead to learning in traditional classrooms (Booth, Oyer, Paré-Blagoev, Elliot, Barbieri, Augustine, & Koedinger, 2015; Davenport, Klahr, & Koedinger, 2007). There are a number of methodological differences between much of the prior research and the current study which may have contributed to this discrepancy. Firstly, the present study was conducted in traditional classrooms, with teachers administering the intervention. Thus, the level of control over order and time of completion of the worked examples was not as high as prior laboratory research (e.g., Paas & Van Merriënboer, 1994), researcher-implemented interventions (Durkin & Rittle-Johnson, 2012) or classroom research using computer-based interventions (Adams, et al., 2014; Booth et al., 2013). Another consideration is that potentially influential features of worked examples may be difficult to ascertain when paired with regular classroom instruction (Van Loon-Hillen, van Gog & Brand-Gruwel, 2012). Finally, the current study did not employ a delayed post-test. Recent research reveals benefits of incorrect worked examples only at a delay (Adams et al., 2014; McLaren, Adams, & Mayer, 2015). Further, Adams and colleagues have recently suggested that a vital component of incorrect worked examples is asking students to correct the errors displayed prior to moving on to a practice problem (Adams, McLaren, Mayer, Goguadze, & Isotani, 2013); this was not required in the current study. As worked examples are an intervention recommended by the What Works Clearinghouse based on strong laboratory findings (Pashler et al., 2007), further research is needed on how to improved design and classroom implementation.

5.2 Conclusion

Findings from the present study support the use of incorrect worked examples by students with low prior knowledge in algebra. Exposure to incorrect methods of solving a problem, displayed to point out common student errors, does not hamper learning in algebra. In contrast, they may provide learning benefits for students who need the most support. Studying common errors may allow students to encode the relevant features of the problem at the same time as target relevant misconceptions that students in the early stages of skill acquisition are likely to have.

Acknowledgements

The authors would like to thank Laura K. Young and Kelly M. McGinn for comments on earlier versions of this manuscript. Thanks are due to Ann Arbor Public School District. A portion of this work was presented at a meeting of the American Educational Research

Association. Research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through Grant R305A100150 to Strategic Education Research Partnership Institute. Continued work was supported by Grant R305B130012 from the Institute of Education Sciences, U.S. Department of Education to the University of Delaware. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Adams, D.M., McLaren, B.M., Durkin, K., Mayer, R.E., Rittle-Johnson, B., Isotani, S., & van Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a web-based tutoring system. *Computers in Human Behavior, 36*, 401-411.

Adams, D.M., McLaren, B.M., Mayer, R.E., Goguadze, G., Isotani, S. (2013, January). Erroneous examples as desirable difficulty. *Paper presented at the International Conference on Artificial Intelligence in Education*, Memphis, USA.

Anderman, L.H. (2003). Academic and social perceptions as predictors of change in middle school students' sense of school belonging. *The Journal of Experimental Education, 72*(1), 5-22.

Bong, M. & Skaalvik, E.M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1-40.

Booth, J.L., Cooper, L., Donovan, M.S., Huyghe, A., Koedinger, K.R., & Paré-Blagoev, E.J. (2015). Design-based research within the constraints of practice: AlgebraByExample. *Journal of Education for Students Placed at Risk, 20*(1-2), 79-100.

Booth, J.L., Oyer, M., Paré-Blagoev, E.J., Elliot, A.J., Barbieri, C., Augustine, A., & Koedinger, K.R. (2015). Learning algebra by example in real-world classrooms. *Journal of Research on Educational Effectiveness, 8*(4), 530-551.

Booth, J.L. & Davenport, J.L. (2013). The role of problem representation and feature knowledge in algebraic equation-solving. *The Journal of Mathematical Behavior, 32*, 415-423.

Booth, J.L., Lange, K.E., Koedinger, K.R., & Newton, K.J. (2013). Using example problems to improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning & Instruction*, *25*, 24-34.

Chen, P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh grade mathematics students. *Learning and Individual Differences*, *14*, 79-92.

Chen, X. (2009). Students who study Science, Technology, Engineering, and Mathematics (STEM) in postsecondary education (NCES 2009-161). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Cho, Y., Weinstein, C.E., & Wicker, F. (2011). Perceived competence and autonomy as moderators of the effects of achievement goal orientations. *Educational Psychology*, 31(4), 393-411.

Davenport, J. L., Klahr, D., & Koedinger, K. R. (2007). *The influence of diagrams on chemistry learning*. Paper presented at the Biennial Conference of the European Association for Research on Learning and Instruction, Hungary, Budapest.

Department of Education. (1997). *Mathematics equals opportunity*. White Paper prepared for U.S. Secretary of Education Richard W. Riley.

Durkin, K. & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, *22*, 206-214.

Dweck, C. S. (1986) Motivational processes affecting learning. *American Psychologist*, *41*, 1040-1048.

Eccles (Parsons), J., Adler, T.F., Futterman, R., Goff, S.B., Kaczala, C.M., Meece, J.L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J.T. Spence (Ed)., *Achievement and achievement motivation* (pp. 75-146). San Francsico: W.H. Freeman.

Eccles, J.S. & Roeser, R.W. (2009). Schools, Academic Motivation, and Stage-Environment Fit. In R. Learner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology,* (pp. 404-434). New York: Wiley.

Eccles, J.S. & Wigfield, A. (1995). In the mind of the achiever: The structure of adolescents' academic achievement-related beliefs and self-perceptions. *Personality and Social Psychology Bulletin, 21*, 215-225.

Eccles, J.S., Wigfield, A., Harold, R., & Blumenfeld, P.B. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*, 830-847.

Elliot, A.J. (2005). A conceptual history of the achievement goal construct. In Elliot, A.J. & Dweck, C.S. (Eds.), *Handbook of competence and motivation*, (pp. 52-72). New York: Guilford.

Elliot, A.J. & Church, M.A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 72*(1), 218-232.

Elliot, E.S., & Dweck, C.S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54,* 5-12.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A., (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Ferla, J., Valcke, M., Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences, 19,* 499-505.

García, T., Rodríguez, C., González-Castro, P., González-Pienda, J.A., & Torrance, M. (2015). *Metacognition and Learning, May 2015.*

Good, C., Rattan, A. & Dweck, C.S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology, 102*(4), 700-717.

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

Heemsoth, T. & Heinze, A. (2014). The impact of incorrect examples on learning fractions: A field experiment with 6th grade students. *Instructional Science, 42*(4), 639-657.

Kelly, D., Xie, H., Nord, C.W., Jenkins, F., Chan, J.Y., and Kastberg, D. (2013). Performance of U.S. 15-Year-Old Students in Mathematics, Science, and Reading Literacy in an International Context: First Look at PISA 2012 (NCES 2014-024). U.S. Department of Education. Washington, DC: National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubsearch.

Matthews, M.S. & Farmer, J.L. (2008). Factors affecting the Algebra I achievement of academically talented learners. *Journal of Advanced Academics, 19*(3): 472-501.

McLaren, B.M., Adams, D.M., & Mayer, R.E. (2015). Delayed learning effects with erroneous examples: A study of learning decimals with a web-based tutor, *International Journal of Artificial Intelligence in Education, 25*, 520-542.

Muller, D., Judd, C.M., & Yzerbyt, V.Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*(6), 852-863.

Osterman, K.F. (2000). Students' need for belonging in the school community. *Review of Educational Research, 70*(3), 323-367.

Paas, F.G.W.C. & van Merrienboer, J.J.G. (1994). Variability of worked examples and transfer of geometrical problem solving skills: A cognitive load approach. *Journal of*

Educational Psychology, 86, 122-133.

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., and Metcalfe, J. (2007). *Organizing Instruction and Study to Improve Student Learning* (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Pintrich, P. & Schunk, D. (1996). *Motivation in Education: Theory, Research, & Applications,* Ch. 3. Englewood Cliffs, NJ: Prentice-Hall.

Rattan, A., Good, C., & Dweck, C.S. (2012). "It's ok - Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, *48*, 731-737.

Santagata, R. (2004). "Are you joking or are you sleeping?" Cultural beliefs and practices in Italian and U.S. teachers' mistake-handling strategies. *Linguistics and Education*, 15, 141-164.

Schleppenbach, M., Flevares, L. M., Sims, L. M., & Perry, M. (2007). Teachers' responses to student mistakes in Chinese and US mathematics classrooms. *Elementary School Journal*, *108*(2), 131-147.

Schneider, B., Swanson, C.B., & Riegle-Crumb, C. (1998). Opportunities for learning: Course sequences and positional advantages. *Social Psychology of Education, 2*(1), 25-53.

Siegler, R. S. (2002). Microgenetic studies of self-explanations. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31-58). New York: Cambridge University Press.

Stigler, J.W. & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.

Stigler, J.W. & Perry, M. (1988). Mathematics learning in Japanese, Chinese, and American classrooms. In G.B. Saxe & M. Gearhart (Eds.), *Children's mathematics: New directions for child development*, No. 41 (pp. 27 – 54.) San Francisco: Jossey-Bass.

Sweller, J. & Cooper, G.A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59-89.

Sweller, J. (2012). Human cognitive architecture: Why some instructional procedures work and others do not. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook; Volume 1: Theories, Constructs, and Critical Issues*. Washington DC: American Psychological Association.

Tulis, M. & Ainley, M. (2011). Interest, enjoyment and pride after failure experiences? Predictors of students' state-emotions after success and failure during learning in mathematics. *Educational Psychology, 31*(7), 779-807.

Van Loon-Hillen, N., van Gog, T., & Brand-Gruwel, S. (2012). Effects of worked examples in a primary school mathematics curriculum. *Interactive Learning Environments, 20*(1), 89-99.

Wang, Q., & Pomerantz, E. (2009). The motivational landscape of early adolescence in the United States and China: A longitudinal investigation. *Child Development, 80*, 1272-1287.

Wigfield, A., Eccles, J., Mac Iver, D., Reuman, D., & Midgley, C. (1991). Transitions at early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology, 27,* 552-565.

Wigfield, A. & Eccles, J.S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J.S. Eccles (Eds.), *Development of achievement motivation* (pp. 92-121). New York: Academic Press.

Zhu, X. & Simon, H.A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction, 4*, 137-166.