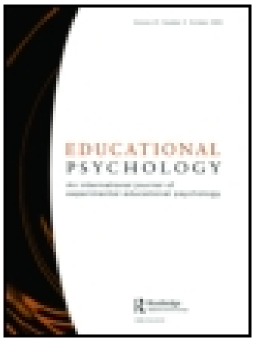

Educational Psychology

An International Journal of Experimental Educational Psychology

ISSN: 0144-3410 (Print) 1469-5820 (Online) Journal homepage: https://www.tandfonline.com/loi/cedp20

Testing the ecological validity of faded worked examples in algebra

Dana Miller-Cotto & Abbey E. Auxter

EDUCATIONAL PSYCHOLOGY
https://doi.org/10.1080/01443410.2019.1646411

Testing the ecological validity of faded worked examples in algebra

Dana Miller-Cotto[a] and Abbey E. Auxter[b]

[a]Learning, Research, and Development Center, University of Pittsburgh, Pittsburgh, PA, USA;
[b]Department of Foundational Mathematics, Community College of Philadelphia, Philadelphia, PA, USA

ABSTRACT

Faded worked examples have been used to promote problem solving performance, such as mathematics performance in many laboratory studies and short-term classroom studies. However, few studies have examined the ways educators may use fading in their own classroom on more accessible platforms that do not require programming experience. Further, few classroom studies have administered fading more than once, limiting the treatment effect. The current study examined whether faded worked examples would promote learning in a classroom. Undergraduates ($N = 135$) completed four homework assignments over the course of one unit in a college semester over the course of two waves of data collection. Using Canvas, homework assignments were deployed once a week for four weeks in the form of (a) faded worked examples, (b) faded worked examples with self-explanations, (c) self-explanations, and (d) business as usual. Results indicated that students in the problem-solving group outperformed those exposed to fading with self-explanation prompts but showed no difference between the fading alone or self-explanation alone condition. Findings are discussed in terms of future research.

ARTICLE HISTORY
Received 6 June 2018
Accepted 17 July 2019

KEYWORDS
Faded worked examples; self-explanation; algebra; higher education

Introduction

Translational research, or using findings from basic laboratory science and attempting to replicate them in real world contexts presents many challenges. Many large scale laboratory experiments attempting to elucidate human learning have produced little to no improvement on student learning and education, though others have produced strong, meaningful outcomes (Booth et al., 2017; Koedinger, Booth, & Klahr, 2013). In particular, many cognitive psychology studies on math performance have been conducted in psychology laboratories, using students enrolled in introductory psychology courses as their participants, often making conclusions about how students learn, but infrequently testing these same studies in classrooms (McCandliss, Kalchman, & Bryant,

CONTACT Dana Miller-Cotto dcotto@pitt.edu Learning, Research, and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, USA

2003). However, the need to translate laboratory work to classrooms has received growing interest in the last 10 years (Begolli & Richland, 2017). Many studies employed cognitive science principles and attempted to replicate findings in a classroom (Booth et al., 2015, 2017). One study (Booth et al., 2015) in particular employed worked examples, which are problems with solution steps included for students to study and practice on their own, with self-explanation prompts in a real world classroom, demonstrating that not only does this work translates well in a classroom, but that students often perform better when studying example problem pairs as opposed to solving problems without the additional support.

A variation of worked examples, faded worked examples, or *fading*, have also shown success in the laboratory, though fewer studies have been tested in real classrooms. Fading involves providing most of the solution steps for a given problem and encouraging students to complete the last step of the problem (i.e. the answer), then students are given a problem where they complete the last two steps of a second problem, and so on until they fully solve the problem on their own. Students are encouraged to complete these steps to scaffold their problem solving (Atkinson, Renkl, & Merrill, 2003). Over time, students who approach mastery level of the content no longer need the fading, and can solve problems on their own without assistance (Schwonke et al., 2009; Schwonke, Renkl, Salden, & Aleven, 2011). This is known as the *guidance fading effect*. Fading has been used to promote mathematical problem solving, and although many cognitive psychologists believe this is an effective strategy, research regarding fading has been executed to a lesser degree than worked examples within classrooms, likely due to the need for computerised tutoring resources. The current study seeks to test the ecological validity of fading in a developmental mathematics classroom.

Theoretical support for worked examples

Sweller (2012) argued that instruction could be more effective when it is aligned with the architecture of human cognition. Human cognitive architecture is defined as the way in which cognitive structures function in organising and processing information (Sweller, 2012). In accordance with our human cognitive architecture, Sweller suggests that when developing instruction, one should consider human cognitive capacity to ensure that cognitive load is reduced. More specifically, the goal is to free up working memory from unnecessary distractions in order to transfer information to long-term memory (Sweller & Sweller, 2006). Taxing working memory resources thus prevents learning from occurring.

Earlier interpretations of Cognitive Load Theory (CLT) suggested that when individuals are solving a problem, much of their cognitive effort is devoted to determining the current problem state (the problem) and the goal state (the solution). This process, particularly for novice learners, imposes greater cognitive load, which prevents the learner from solving the problem or gaining any additional knowledge (Sweller, 2012). In particular, because learners devote most of their cognitive resources to figuring out what steps are needed between the problem and the solution, those lacking experience with the material have greater cognitive load. In an effort to remedy this problem, CLT suggests that instructional designers should focus their efforts on findings ways that students can distribute their cognitive resources during the learning phase and the problem solving phase, but not both simultaneously (Renkl & Atkinson, 2010).

Worked examples have been theorised to reduce cognitive load (Moreno & Mayer, 2010; Sweller, 2012). The *worked example effect* (Sweller & Cooper, 1985) is the opportunity for a learner to study worked out solution steps, rather than having to focus on problem solving which, as previously noted, places high cognitive processing demands on the learner (Sweller & Sweller, 2006). In the later stages, when learners reach mastery, they begin to apply the learned problem-solving procedures and can solve problems quicker, and with greater accuracy, without examples. However, there appears to be a point in which the worked example effect has diminished returns for the learner.

Translating worked examples to real world classrooms

Numerous cognitive science principles have demonstrated successful research results across both laboratory and classroom settings (Booth et al., 2017). Worked examples is one notable example of this successful implementation in both research contexts. Worked examples are a promising method at improving students' performance during the early stages of content learning. Booth et al. (2015) investigated the effect of worked examples, both correct and incorrect, on algebra learning. To do this, students were randomly assigned to either the treatment (i.e. worked example problem pairs) or the control (i.e. business as usual) condition. Results demonstrated that students exposed to worked examples showed improved algebra knowledge, and that this was particularly the case for students with low prior knowledge. Thus, their findings highlight the successful implementation of worked examples within classrooms. Unfortunately, despite this increased interest in classroom-based research, laboratory studies with educational implications still emerge at a faster rate than studies designed for students within classrooms, as such, laboratory studies remain prevalent.

This has certainly been the case in numerous fading studies, although fading with engineering students has been particularly prevalent in the education literature. For instance, Reisslein, Atkinson, Seeling, and Reisslein (2006) compared participants' performance after exposure to fading with worked example problem pairs and problem example pairs, finding that although participants did not demonstrate differences on outcome performance, students with low prior knowledge benefitted most from example problem pairs while high prior knowledge learners benefitted from problem example pairs. It may be the case that learners who are of average prior knowledge may benefit from fading, which was not addressed in the prior study. To address this need, fading is used to remove each of the steps from a worked example for novice learners until they can successfully solve problems without assistance (Sweller, 2012). Thus, within the CLT framework, going from full worked examples to problems alone, with a series of faded steps in between, is believed to reduce cognitive load through a series of problems successively. It is worth noting, however, that these studies included an undergraduate population. Fading can be easily executed in classroom that have the resources to support it. This is potentially one reason that fading has been primarily employed in undergraduate classrooms, although the role of prior knowledge in a typical population, or struggling population like the one in the current study, still remains unclear.

Fading to reduce cognitive load

Many experiments have examined fading in a computerised environment, often with undergraduate participants requiring one session of their time. Salden, Aleven, Renkl, and Schwonke (2010) examined learning gains when adapting worked-examples based on student's explanations of worked-out steps in a laboratory experiment and a separate classroom experiment (i.e. two experiments total), both taking place in one session. Comparisons were made across three conditions: problem solving without assistance, fixed fading (i.e. each student has the same number of steps removed), and adaptive fading (i.e. each student has steps removed based on their prior knowledge level). Results suggested that the students in the adaptive fading condition outperformed those in the other conditions on the posttest and the delayed posttest. The authors concluded that students' knowledge levels increased faster in the adaptive condition in the laboratory, but slower in the classroom, potentially because there is less time available in a classroom. Salden and colleague's (2010) study results highlight that fading, even after one exposure, can be effective at showing learning growth. However, although support for adaptive fading on learning was established, the specialised technology needed to carry out adaptive fading may not be available for students in all classrooms.

Salden and colleagues' (2010) also tested the degree of how well students could identify similarity between problems, or transfer. Their results indicated that transfer was more common in the adaptive fading condition than the problem solving condition at both the immediate posttest and delayed posttest. While the authors' decision to compare results from a lab setting to a classroom setting is a significant contribution to the field, it is unclear why students in the second experiment were measured on different material from students in the lab study due to their familiarity with the intelligent tutoring system program. Thus, the differences in adaptive fading in the classroom setting and lab setting could be due to the difference in problems given to each sample and their experience with intelligent tutoring systems. In other words, condition was confounded with sample setting. It is necessary to use identical studies from laboratory work in classrooms to determine true ecological validity and translate this work to classrooms.

Self-explanations

One of the methods employed to enhance worked examples and their variants (i.e. fading) is self-explanation prompts, or requiring students to explain the reasoning behind their procedures or the procedures presented in a worked example (Booth et al., 2015; Chi, 2000). Self-explanations encourage students to make their knowledge explicit by bringing attention to steps in a worked example or combining this new information with information from their prior knowledge. Earlier work suggests that successful problem-solvers often self-explain without prompting in the absence of worked examples (Chi, 2000), including engaging in self-monitoring strategies during problem solving, such as checking their own work for errors and reviewing steps to solve a problem (VanLehn, Jones, & Chi, 1992). Students with self-explanation prompts performed better than those without prompts for near and far transfer measures (Atkinson and colleague, 2003). Further, Schworm and Renkl (2006) found that self-explanation prompts, regardless of what type of examples they were paired with, provided a positive outcome for learners. And finally, Rittle-Johnson compared groups on instruction versus invention of a procedure and the use of self-explanation. Results indicated that groups who self-explained demonstrated the most learning on outcome measures.

The findings of these previous experiments are supported in numerous studies that also employed fading techniques (Sweller, 2012; Moreno & Mayer, 2009; Reisslein et al., 2006). Notably, despite consistency in the findings of a positive impact of self-explanations on learning outcomes, effect sizes vary depending on a variety of factors: whether the study was conducted in laboratory (Atkinson et al., 2003; Renkl, Atkinson, & Große, 2004) or in a classroom (Atkinson et al., 2003; Flores & Inan, 2014; Hesser & Gregory, 2015), and whether fading was static (i.e. steps removed at the same pace for all students) versus adaptive (i.e. steps removed based on the students' prior knowledge; Salden et al., 2010) and what specific outcomes variables were measured (e.g. overall performance, transfer; Reisslein et al., 2006). Of these moderating variables, prior knowledge has been examined as a potential reason for this discrepancy, suggesting that students with higher prior knowledge do not benefit from this learning technique. It is also worth examining the cognitive principles, like fading, in a real world classroom to understand how potential noise, such as classroom instruction, carry over effects, and factors typical of a classroom experiment, might explain translation research in classrooms.

Thus, the current study sought to determine whether administering fading over several assignments would warrant positive results, relative to those observed in prior laboratory and classrooms that employed fading in one session and explored the effect of instruction by conducting the study in a real world natural setting where instruction could potentially interfere with the effectiveness of the intervention.

The current study

In summary, prior work has established that fading and worked examples with self-explanation prompts often promote learning when examining students' pre- and posttest scores. Unlike worked examples, however, fading has demonstrated less popularity in classrooms. Further, when fading studies were conducted in classrooms, they often took place during one assignment, rather than during multiple assignments much like the research on worked examples.

In the current study, we examined whether fading and self-explanation prompts would improve mathematical performance in a college algebra class over the course of four assignments. Many cognitive science principles, like the *guidance fading effect*, have demonstrated success in laboratories, highly controlled contexts. Given the often robust findings from these prior studies it is imperative that we determine ways to employ these interventions in classroom settings. Further, many randomised control trials based on cognitive science principles have been conducted, and not all produce favourable results (Koedinger et al., 2013). Thus, we aimed to replicate findings from existing laboratory studies and short-term classroom studies in a study with four homework assignments throughout the semester. More concretely, we were interested in whether fading would improve students' final exam scores when integrated into a college algebra unit in a series of four computer based homework assignments. Because fading is often carried out as a step-by-step approach to problem solving, it was anticipated that fading would likely produce gains in overall posttest performance. We outline our research questions below:

1. Are there differences on posttest knowledge when comparing a business as usual condition, fading, fading with self-explanation, and self-explanation alone?
2. Does prior knowledge moderate the effect of condition?

Method

Participants

Participants ($N = 135$) with completed assent forms were drawn from six college developmental algebra classes students at a satellite campus of a larger university in the Mid-Atlantic region of the USA collected during two waves of data collection. That is, the study was conducted over the course of two semesters, three classes per semester. Developmental algebra classes at this specific university implied that students had not entered college with the adequate skills to take college algebra, and instead were urged to take this developmental algebra class that included much of the material they would encounter in the college algebra class. That is, the developmental algebra class was a remedial college algebra course. Two participants failed to return consent forms, requiring us to exclude their data from the final sample. Further, a large proportion of the initial sample ($n = 63$) students did not complete the homework assignments, discouraging us from using their data since they did not technically participate in the intervention. The sample was ethnically and racially diverse, including White (47.4%), African-American/Black (17.0%), Hispanic/Latino (11.9%), Asian (20.0%), as well as students who identified as two or more races (3.7%). More than half of the participants were women (60%). All participants were at least 18 years of age for consent to participate purposes.

The current study employed a pretest administered in Canvas, an online learning management system (LMS) tool used by teachers to post-readings, assignments, and class information in one place, posttest design with the experimental manipulation within the web-based intervention. Students were randomly assigned within classrooms to one of the following conditions: business as usual problem solving ($n = 41$), fading ($n = 33$), worked examples with self-explanation prompts ($n = 36$), and fading with self-explanation prompts ($n = 25$).

Assignments

The intervention portion of this study was assigned as homework assignments. This was done to complement the instruction utilised in the classroom. The intervention was executed within Canvas. Thus, the intervention was done as a course assignment. Students were exposed to a total of four problems for each assignment, but assignments varied based on condition. Students in the business as usual (control) condition simply solved a series of problems assigned for homework without assistance. In the worked example with self-explanation condition, students were instructed to first study the worked example problem and then respond to a prompt 'Why was this step used in the problem?' regarding a specific step. Students were then told to type their responses into a textbox. In the fading condition, students were given the same set of problems in the worked example conditions except the last was removed from the first problem, then the last two steps were removed in the second problem, and so forth. In the fading and self-explanation condition, students were given the same set of problems and missing steps to complete. Once they completed the faded steps, students were prompted to respond to the same self-explanation prompt, 'Why was this step used in the problem?' regarding a specific step in the problem. Assignments included four problems per assignment. There were four total assignments for the entire unit, one assignment per week for four weeks. This was done to remain consistent with the other units in the course. In other words, each of the units in the course took place for four weeks each. Assignments were scored as complete once students submitted them in Canvas so that the course instructor could give each student credit without being made aware that the student chose not to participate in the study. Data were de-identified and not entered nor analysed until course grades were submitted to prevent possible bias. Examples of each condition are illustrated in Figure 1.

Pre- and posttest assessments

The pre- and posttest teacher constructed questions assessed students' ability to perform procedural computations learned in class and the ability to identify similar procedures that are used for problems that are structurally different from problems. In other words, these assessments were used to ascertain whether students could see similarities in problems that were novel in presentation, but similar in concepts they had learned in class (e.g. seeing a word problem instead of an equation to be solved). The test questions were also designed to target students' ability to solve problems from a unit on systems of equations (SOE). Final pre- and posttest scores were computed as the percent of items answered correctly. Pretests included a total of 20 items that assess students' knowledge about solving systems of equations. The posttest included similar items from the pretest, as well as items that were structurally different from the pretest items to assess overall understanding of SOE; internal consistency was considered acceptable for the pretest ($\alpha = 0.743$) and posttest ($\alpha = 0.679$). Sample items are displayed in Figure 2.

Condition	Sample Problem in Canvas
Problem Solving	Prompt: Please complete the following problem
Fading	Please supply the last step: x = -5 9(-5) – 5y = -10 -5y = 35 y = -7
Worked Example + Self-Explanation	Why do we end up solving for x for this problem while in Question 1 we first solved for y? x = -5 9(-5) – 5y = -10 -5y = 35 y = -7 (-5, -7)
Fading + Self-Explanation	Why do we end up solving for x for this problem while in Question 1 we first solved for y? x = -5 9(-5) – 5y = -10 -5y = 35 y = -7

Figure 1. Problem by condition example.

Pre-test Questions	Posttest Questions
Solve system by elimination. Show all Steps. Express answer as an ordered pair. 0.5x – 0.2y = -0.1 0.1x + 0.4y = 3.5	**Solve system by elimination. Show all Steps. Express answer as an ordered pair.** 3x + 2y = -8 2x + 2y = -4
Solve system by elimination. Show all Steps. Express answer as an ordered pair. 0.2x + 0.25y = 3 0.4x + 0.5y = 2	**Solve system by elimination. Show all Steps. Express answer as an ordered pair.** 0.6x – 0.5y = -1.6 -0.2x + 0.6y = 1.4
Solve system by elimination. Show all Steps. Express answer as an ordered pair. 0.6x – 0.5y = -1.6 -0.2x + 0.6y = 1.4	**Solve system by elimination. Show all Steps. Express answer as an ordered pair.** 2x + y = 10 X – 2y = 4
Solve system by elimination. Show all Steps. Express answer as an ordered pair. 6x + 5y = -30 2x + 2y = -12	**Solve system by elimination. Show all Steps. Express answer as an ordered pair.** 0.5x – 0.2y = -0.1 0.1x + 0.4y = 3.5

Figure 2. Pre/posttest sample questions.

Procedure

One instructor taught all six of the courses for the current study. Prior to the field experiment, consent forms were distributed explaining that the instructor was employing a new method to teaching the unit SOE. It was stated that although completion of the assignment was part of their course grade, it was their decision whether to release their data to the researchers, and the decision to release their data would not affect their grade in the class, as consent forms would be reviewed after final grades were entered. After collecting consent forms three weeks later, the instructor administered a pretest to determine students' systems of equations knowledge.

In class meetings prior to teaching SOE the instructor and the students explored the format and meaning of linear equations with one and two variables. This process was used to compare structural differences for problems that were conceptually the same. Questioning was done to promote procedural flexibility and help students realise there are multiple ways of solving the problems; though, in certain instances one method may be easier or quicker than another. Throughout the process students compared the algebraic and graphical representations to each other and applications found in the world. The unit SOE was introduced by reviewing the format of an equation with two variables, its solutions, and what those solutions mean both algebraically and graphically. Students examined graphical representation of SOE and discussed the similarities and differences between the solutions to an equation with two variables versus a system. They compared multiple graphical representations and their algebraic counterparts and students figured out how to find the solution to a system of equations graphically. Students were then introduced to two methods: the substitution method and the elimination method. A discussion as to which method is more beneficial in what circumstances took place and the students were asked prior to every practice problem, 'What method do you suggest? Why?' Additionally, when the substitution method was chosen students were asked, 'For which variable are you substituting? Why?' When the elimination method was chosen student were asked, 'Which variable are you eliminating? Why?' Other questions frequented the solution of systems such as, 'Could you use another method? Could you substitute/eliminate the other variable? Why did you not choose to do that?' After solving a system, the two equations were graphed and the solution was shown both algebraically and graphically in order to promote further understanding of what the solution means. At this point, students were able to transition to applications using SOE as a problem-solving tool. Students and the instructor discussed the identification of two different unknowns to take place of the variables by identifying the unknowns and gave them a name (e.g. hours worked = h), identified the relationships between the pieces (e.g. 15 dollars per hour is a multiplicative relationship between pay and hours worked = 15 h), and created a model based on what we know and do not know. Once the models were created students were able to employ their knowledge of solving systems of equations in two variables. It was the instructor's intention to give students reasoning behind why certain steps were used as they problem solved.

Following each lesson on SOE, students were told to complete the homework assignments available on Canvas. During the initial assignment, students were randomly assigned to one of the experimental conditions within the Canvas platform. They were administered one assignment a week for four weeks in the same condition. This was done to be consistent with the length of the other units in the course (i.e. four week units). After four weeks of exposure to the assignments, students were administered the posttest.

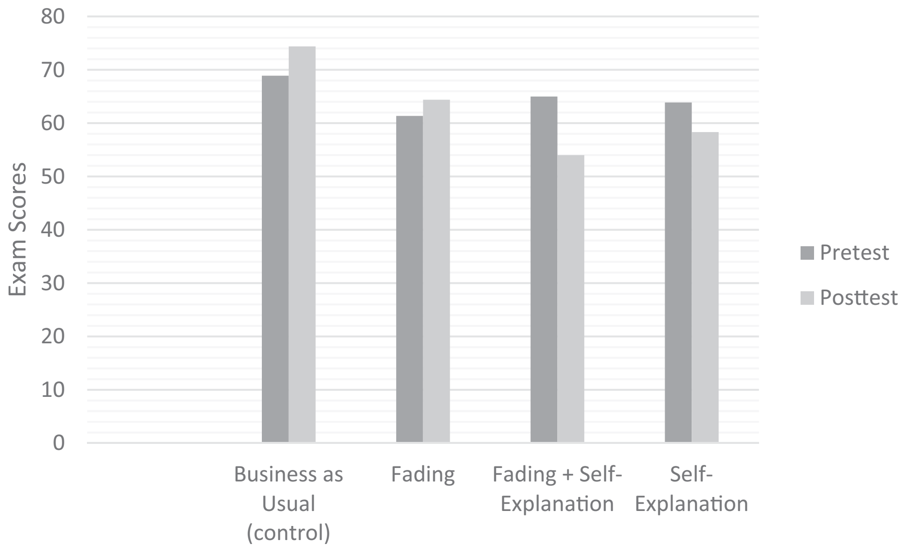

Figure 3. Condition comparisons.

Fidelity of implementation

The first and second author worked together to create a script to ensure instructor fidelity of implementation of instruction. During a practice lesson, the first author attended one session of the instructor's lecture to ensure that this was consistent with the script. These lessons were consistent across all classes. That is, lessons were the same regardless of condition. The first author observed some courses periodically to ensure the instructor was following as close to the script as possible. Students in the experiment were not made aware of the research questions nor were they made aware of the first author's reason for visiting the class.

A priori and post hoc *power analyses*

We used G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) to conduct a priori and *post hoc* power analyses to determine the minimum sample size to detect a significant effect on our outcome measures using $\alpha = 0.05$. Prior analyses suggested a sample size of $N = 120$ would provide a medium effect on learning, however, post hoc analyses revealed that the actual sample of $N = 135$ was slightly underpowered. We elaborate on this limitation in the 'Discussion' section (see Figure 3).

Results

Preliminary analyses were first conducted to screen for potential differences between groups based on the wave of data collection on pre-test measures, differences in conditions on pretest measures, as well as differences between the six classes on pretest measures. A one-way analysis of variance (ANOVA) indicated that there were no significant differences between participants across the two waves, or two semesters, of data collection on pretest scores ($F(1, 134) = 0.041$, $p = .839$). Further, there were no differences between participants between conditions on pretest measures ($F(3,

134) = 0.453, $p = .716$). Finally, we screened for pretest score differences between the six algebra classes that participated in the study. There were no significant differences between the six algebra classes on pretest scores ($F(5, 134) = 1.296$, $p = .270$).

Predicting posttest scores

The current study examined whether students would demonstrate a difference in posttest exam scores after exposure to fading or self-explanation treatments. To answer this, we conducted an analysis of variance (ANOVA). Our analyses revealed a main effect of condition ($F(3, 131) = 3.397$, $p = .015$, $\eta^2 = .076$). Students in the problem-solving condition demonstrated the largest posttest composite scores ($M = 74.4$) as compared to students in fading condition ($M = 64.4$), worked example with self-explanation prompts ($M = 58.3$), and fading with self-explanation prompts ($M = 54$). Tukey's *post hoc* comparisons revealed that the business as usual (control) condition was significantly different from the fading with self-explanation condition ($p = .02$) and self-explanation alone conditions ($p = .05$), with the fading with self-explanation prompt condition performing significantly worse than the problem-solving condition. Thus, the business as usual (control) condition was not significantly different from fading alone, suggesting that there were no differences in posttest scores between these two conditions. There was also a significant main effect of prior knowledge, ($F[1, 130] = 22.76$, $p < .001$, $\eta^2 = 0.14$) with high prior knowledge students outscoring low prior knowledge students on the posttest. The interaction between treatment and prior knowledge was not significant ($F[3, 127] = 1.42$, $p = .241$, $\eta^2 = 0.026$). It is necessary to note, however, that observed power to detect a main effect of condition on learning was not strong enough. We take this to consideration in the 'Discussion' section.

Discussion

Cognitive scientists have asserted for many years that fading is an effective method to improve problem solving for novice learners (Atkinson et al., 2003; Salden et al., 2010). The results of the present analyses indicate that, at least in one classroom study, findings from numerous prior fading studies done in laboratories or as a one-time assignment in undergraduate classrooms do not easily replicate. As the results indicate, fading was no more effective than business as usual. Interaction analysis did suggest that the experimental conditions were all indeed beneficial to students. From a practical standpoint, this may be helpful to educators when designing classroom interventions.

This study, however, supports earlier findings from Salden and colleagues (2010) that indicates that fading is often no better than business as usual. There are numerous reasons for this. First, fading, a component of scaffolding, requires that students perform at a certain level before fading can be effective. If students are far below competency and are exposed to fading, one might expect that they do not have the prior knowledge, or fading is too far outside their zone of proximal development, to have any benefit for them. Alternatively, if students are too advanced, fading is no longer necessary as an intermediary step between worked examples and problem solving alone, and also has no benefit (Renkl, 2011). Second, in a fading environment, if students are exposed to fading where steps are removed at the same rate for everyone, earlier work suggests that business as usual problem solving would outperform those in the fading condition because step removal should be based on the students' performance. Thus, these findings do indeed replicate static fading studies.

However, we did observe a difference between the business as usual and the fading with self-explanation condition. One reason for this might be because of the inclusion of the two elements, fading and self-explanation. It is possible that students may have been distracted by both elements, especially if they found the content particularly difficult. Another reason for this finding could be that the self-explanation prompts were a simple 'why' question, rather than a more structured self-explanation prompt used in earlier work (Chi, Bassok, Lewis, Reimann, & Glaser, 1989) that employed drop down menus (Atkinson et al., 2003), or self-explanation prompts that focused on a particular step or error within the problem (Barbieri & Booth, 2016; Booth, Lange, Koedinger, & Newton, 2013), and may also explain why both self-explanation conditions performed worse at posttest. Unstructured self-explanation prompts, although they do encourage metacognitive monitoring, may be deemed unbeneficial for this reason (see Rittle-Johnson & Loehr, 2017 for a review).

Among the strengths of the study was the field experiment of the current study. As previously noted, many fading studies have been completed in a laboratory with psychology 101 students, limiting the generalizability and practical effect of these findings. Further, when prior studies were conducted in a classroom, students were often exposed to the fading treatment one time. Though it is impressive that students demonstrated an improvement in knowledge after one exposure to the treatment in these earlier studies, this may not be practical for a real classroom where students may need exposure to content more than once given their knowledge or the scheduling of the content. The current study not only executed the same study in a classroom, but presented the treatment to students on more than one occasion, with the same results. However, one limitation that is apparent in the current study is that the study was underpowered. Although our findings replicate prior work, we must acknowledge that including more participants in the study likely would have shown a significant difference between the conditions.

Of course, the current study is not without its limitations. The first consideration is that because this study was conducted in a classroom setting, there were fewer controls available than there would have been in a laboratory setting. It might have been informative to have been controlling for time spent on problems to ensure students were putting in equal effort on the problems within each condition. Along these same lines, while one of the strengths of the study is the dosage, administering four assignments over the course of the experiment, we do not examine the rate of change across assignments between conditions; this is certainly a limitation of the current study. The inclusion of such information could potentially illuminate shortcomings of individual design features of the experiment.

A second consideration when considering the current findings are the unequal groups of students across conditions. With such unequal groups, it is possible that comparisons were difficult to make. It is possible the results might have been different had the groups been even. Finally, the pre- and posttest measures were only deemed acceptable relative to traditional thresholds for the field (Cortina, 1993). This speaks to the importance of strong teacher created measures in classrooms when testing new interventions to improve learning. However, it is also worth noting that many teachers create their own measures for classes, making comparisons between laboratories and classrooms difficult. Indeed, there are many tradeoffs when attempting to conduct cognitive research in classrooms that should be considered when researchers and teachers collaborate.

Though the need for cognitive psychology studies that have implications for education to be replicated in classrooms is clear, there is still a need to understand why such interventions work. Understanding why these interventions work is necessary when designing effective interventions, or when tweaking these interventions for different populations. Future research might address these issues by conducting meta-analyses on earlier fading studies to determine how various moderator variables alter the effect sizes of fading and performance. Alternatively, future fading studies may also include other covariates beyond prior knowledge to determine what types of students benefit from this strategy over others. As previously stated, timing students and measuring their perceived effort on such tasks may illuminate differences in knowledge after the intervention.

Perhaps the most important lesson to be learned from this study is that it is difficult to make conclusions about cognition studies that are relevant to the education community without working directly with the education community. It is necessary for more of these studies to be replicated in classrooms, especially if there if overwhelming confidence in the field about a particular intervention. Intervention design would also benefit from collaboration with classroom teachers to determine what is practical in a classroom, or in our case what self-explanation prompts should focus on, with limited time and other resources for this intervention to benefit their students. Further, teachers hoping to implement fading or an alternative scaffolding strategy in their own classrooms may consider developing a tracking system of students' progress. Doing this, consistent with most fading and scaffolding studies, will allow them to adjust their instructional method and make it consistent with the students' level of understanding as it increases over the course of a unit. An ecological perspective on fading and similar interventions requires that cross collaboration between the fields of cognition, psychology, and education commence to have a better understanding of cognitive science in education. Encouraging such collaborations has the potential to have the much needed impact on education that has been sorely needed.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Dana Miller-Cotto http://orcid.org/0000-0003-0577-4868

References

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. doi:10.1037/0022-0663.95.4.774

Barbieri, C., & Booth, J. L. (2016). Support for struggling students in algebra: Contributions of incorrect worked examples. *Learning and Individual Differences*, *48*, 36–44. doi:10.1016/j.lindif.2016.04.001

Begolli, K. N., & Richland, L. E. (2017). Bridging cognitive science and real classrooms: A video methodology for experimental research in education. *Journal of Experimental Education, 86*, 671–689.

Booth, J. L., Cooper, L. A., 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 (Jespar)*, *20*(1–2), 79–100. doi:10.1080/10824669.2014.986674

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 and Instruction*, *25*, 24–34. doi:10.1016/j.learninstruc.2012.11.002

Booth, J. L., McGinn, K. M., Barbieri, C., Begolli, K. N., Chang, B., Miller-Cotto, D., … Davenport, J. L. (2017). Evidence for cognitive science principles that impact learning in mathematics. In D. C. Geary, D. B. Berch, R. J. Ochsendorf, & K. M. Koepke (Eds.), *Acquisition of complex arithmetic skills and higher-order mathematics concepts* (Vol. 3, pp. 297–325). Oxford, UK: Elsevier.

Booth, J. L., Oyer, M. H., 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. doi:10.1080/19345747.2015.1055636

Chi, M. T. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. *Advances in Instructional Psychology*, *5*, 161–238.

Chi, M. T., 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.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. doi:10.1037/0021-9010.78.1.98

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*(2), 175–191.

Flores, R., & Inan, F. (2014). Examining the impact of adaptively faded worked examples on student learning outcomes. *Journal of Interactive Learning Research*, *25*(4), 467–485.

Hesser, T. L., & Gregory, J. L. (2015). Exploring the use of faded worked examples as a problem solving approach for underprepared students. *Higher Education Studies*, *5*(6), 36. doi:10.5539/hes.v5n6p36

Koedinger, K. R., Booth, J. L., & Klahr, D. (2013). Instructional complexity and the science to constrain it. *Science*, *342*(6161), 935–937. doi:10.1126/science.1238056

McCandliss, B. D., Kalchman, M., & Bryant, P. (2003). Design experiments and laboratory approaches to learning: Steps toward collaborative exchange. *Educational Researcher*, *32*(1), 14–16. doi:10.3102/0013189X032001014

Moreno, R., & Mayer, R. E. (2010). Techniques that reduce extraneous cognitive load and manage intrinsic cognitive load during multimedia learning. In J. L. Plass, R. Moreno, & R. Brunken (Eds.), *Cognitive load theory*. Cambridge: Cambridge University Press.

Reisslein, J., Atkinson, R. K., Seeling, P., & Reisslein, M. (2006). Encountering the expertise reversal effect with a computer-based environment on electrical circuit analysis. *Learning and Instruction*, *16*(2), 92–103. doi:10.1016/j.learninstruc.2006.02.008

Renkl, A. (2011). Instruction based on examples. *Handbook of research on learning and instruction*, 272–295.

Renkl, A., & Atkinson, R. K. (2010). Learning from worked examples and problem solving. In J. L. Plass, R. Moreno, & R. Brunken (Eds.), *Cognitive load theory*. Cambridge: Cambridge University Press.

Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading worked solution steps works – A cognitive load perspective. *Instructional Science*, *32*(1/2), 59–82. doi:10.1023/B:TRUC.0000021815.74806.f6

Rittle-Johnson, B., & Loehr, A. M. (2017). Eliciting explanations: Constraints on when self-explanation aids learning. *Psychonomic Bulletin & Review*, *24*(5), 1501–1510. doi:10.3758/s13423-016-1079-5

Salden, R. J. C. M., Aleven, V., Schwonke, R., & Renkl, A. (2010). The expertise reversal effect and worked examples in tutored problem solving. *Instructional Science*, *38*(3), 289–307. doi:10.1007/s11251-009-9107-8

Schwonke, R., Renkl, A., Krieg, C., Wittwer, J., Aleven, V., & Salden, R. (2009). The worked-example effect: Not an artifact of lousy control conditions. *Computers in Human Behavior*, *25*(2), 258–266. doi:10.1016/j.chb.2008.12.011

Schwonke, R., Renkl, A., Salden, R., & Aleven, V. (2011). Effects of different ratios of worked solution steps and problem-solving opportunities on cognitive load and learning outcomes. *Computers in Human Behavior*, *27*(1), 58–62. doi:10.1016/j.chb.2010.03.037

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education*, *46*(4), 426–445. doi:10.1016/j.compedu.2004.08.011

Sweller, J., & Sweller, S. (2006). Natural information processing systems. *Evolutionary Psychology*, *4*, 434–458.

Sweller, J. (2012). Human cognitive architecture. In K. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook* (pp. 295–395). Washington, DC: American Psychological Association.

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. doi:10.1207/s1532690xci0201_3

VanLehn, K., Jones, R. M., & Chi, M. T. (1992). A model of the self-explanation effect. *The Journal of the Learning Sciences*, *2*(1), 1–59. doi:10.1207/s15327809jls0201_1