Smith Hannah (Orcid ID: [0000-0001-8979-3315)](http://orcid.org/0000-0001-8979-3315)

Closser Avery H (Orcid ID: [0000-0003-2712-8509)](http://orcid.org/0000-0003-2712-8509)

Chan Jenny Yun-Chen (Orcid ID: [0000-0003-1479-5935)](http://orcid.org/0000-0003-1479-5935)

IMPACT OF WORKED EXAMPLES ON LEARNING

# The Impact of Algebra Worked Example Presentations on Student Learning

Hannah Smith\*, Avery H. Closser, Erin Ottmar, and Jenny Yun-Chen Chan

Social Science and Policy Studies, Worcester Polytechnic Institute 100 Institute Road, Worcester, MA 01609 USA

\* Corresponding author: Hannah Smith, hsmith2@wpi.edu, Orcid: 0000-0001-8979-3315

Avery H. Closser, ahclosser@wpi.edu, Orcid: 0000-0003-2712-8509

Erin R. Ottmar, erottmar@wpi.edu, Orcid: 0000-0002-9487-7967

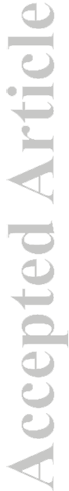
Jenny Yun-Chen Chan, jchan2@wpi.edu, Orcid: 0000-0003-1479-5935

**Conflict of Interest Statement.**

Erin Ottmar is a designer of Graspable Math and owns 10% equity in Graspable Inc. This has been disclosed to WPI’s Conflict Management Committee, and a conflict management plan has been implemented. All other authors have no competing interests to disclose.

**Data Availability Statement.**

The final dataset for the current study is available on the Open Science Framework at <https://osf.io/uv48t/?view_only=b23d407d698f423b9d037fec646ca98b>

**Acknowledgments.**

We thank the ASSISTments Foundation and WPI for hosting the www.etrialstestbed.org that made this study possible and that is funded by NSF grants (1724889 & 1931523) and Schmidt Futures. The research and development reported here was also supported by the Institute of

Education Sciences, U.S. Department of Education through the Small Business Innovation Research (SBIR) program contracts 91990019C0034 and 91990018C0032 to Graspable Inc. The opinions expressed are those of the authors and do not represent views of the National Science

Foundation, the Institute or the U.S. Department of Education. We thank Anthony Botelho,

Cindy Trac, Hailey Anderson, and members of the Math Abstraction Play Learning and Embodiment Lab for their work on this project, and the participating teachers and students for their support with this study.

**Author Contribution Statement.**

**Hannah Smith:** Materials, Methodology, Analysis, Investigation, Writing, Visualization, Project Leader. **Avery H. Closser:** Methodology,Investigation, Materials, Writing. **Erin Ottmar:** Writing, Funding Acquisition. **Jenny Yun-Chen-Chan:** Conceptualization, Methodology, Project Administration, Writing, Supervision.

**Ethics Approval Statement:** The data reported in this manuscript were collected in compliance with a protocol approved by the Institutional Review Board at Worcester Polytechnic Institute. This manuscript has not been previously published, nor has it been submitted to any other journal for review. No copyrighted material has been used.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the [Version of Record](http://dx.doi.org/10.1002/acp.3925). Please cite this article as doi: [10.1002/acp.3925](http://dx.doi.org/10.1002/acp.3925)

This article is protected by copyright. All rights reserved.

# Abstract

Worked examples are effective learning tools for algebraic equation-solving. However, they are typically presented in a static concise format which only displays the major derivation steps in one static image. The current work explores how worked examples that vary in their *extensiveness* (i.e., detail) and degree of *dynamic presentation* (i.e., static vs. sequential line-byline vs. dynamic format that demonstrates the problem-solving process) impact learning. We conducted an online experiment in which 230 algebra students completed a pretest, studied worked examples in one of six presentation conditions, and completed a posttest. We found that overall, students improved from pretest to posttest after viewing the worked examples; we did not find significant differences on posttest performance between worked example presentations. These results have implications for the design of worked examples in online tutoring systems as well as for cognitive load theory and perceptual learning theory in the design of worked examples.

**Keywords:** Worked Examples, Algebra, Online Learning, Cognitive Load Theory, Perceptual Learning

# The Impact of Algebra Worked Example Presentations on Student Learning

When students learn new procedures in mathematics, they may struggle to apply these procedures correctly during equation-solving. For instance, students may still struggle to correctly distribute to all terms within parentheses even after being taught the procedure for distribution. In cases like these, worked examples can be used as a tool for learning by providing the derivations to mathematics problems; they also explicitly break down a given problem stepby-step to show one way of correctly solving the problem. Worked examples have proven to be effective instructional support in a broad range of subjects (e.g., *language*: Lu et al., 2020; *chemistry*: McLaren et al., 2016; *statistics*: Tempelaar et al., 2020) and specifically in algebra

(Atkinson et al., 2000; Barbieri & Booth, 2020; Booth et al., 2013; Carroll, 1994; Renkl, 2014).

Prior research on worked examples has mostly presented *static concise* derivations of problems in which worked examples are displayed as an image of the major derivation steps (Booth et al., 2013; Rittle-Johnson et al., 2009; Star et al., 2015). However, it is unclear whether *static* *concise* worked examples are the most effective presentation to support student learning since limited research has investigated the impact of different worked example presentations. Further, with the introduction of new technologies, we can now dynamically interact with algebraic symbols, and alter the visual format and perceptual features of worked examples effortlessly at scale. As more teachers incorporate educational technologies in their classroom instruction, it is important to explore how these technologies provide affordances for presenting worked examples, and how different presentations of equation-solving processes may help students learn and practice algebra in online learning platforms.

We aim to add to existing research on worked examples by investigating the effects of different features of the worked example on algebra learning in an online environment. In particular, we explore whether students benefit more from viewing *concise* or *extended* worked examples, varying in the length and detail of problem derivation. We also explore whether students benefit more from viewing worked examples that are *static* images, worked examples that provide *sequential* presentations of derivation steps in a looping video, or worked examples showing the *dynamic* process of solving an equation with an online algebra notation tool in a screen-recorded video. Building on cognitive load theory (Sweller, 1994) and perceptual learning theory (Gibson, 1969), this study aims to contribute a richer understanding of how viewing different worked example presentations impacts student learning and to provide recommendations for online tutoring systems which use worked examples for algebra instruction.

# Traditional Worked Examples

The benefits of using worked examples in algebra have been examined over the past few decades (e.g., Sweller & Cooper, 1985; see Atkinson et al., 2000, for a review) with much of the research focusing on the impact of worked examples on learning and the methods for implementing worked examples. Carroll (1994) found that algebra students who were given worked examples with practice paired problems learned quicker, with less instruction, and made fewer errors during practice compared to their counterparts who practiced solving problems without worked examples. In another study, Booth and colleagues (2013) explored Algebra I student performance in an online learning platform. Students either viewed worked examples and solved practice problems or only solved problems in the program. The students who viewed worked examples outperformed the students who only practiced solving problems on measures of conceptual knowledge of algebra. Similarly, Foster and colleagues (2018) confirmed that undergraduate students who received worked examples followed by problem-solving outperformed those who only practiced problem-solving. Together, these studies show that worked examples lead to more efficient and effective student learning than solving problems alone.

However, it is important to note that the worked example materials in these studies are all static images that display the major derivations of each problem solution. Extending prior work and drawing from the cognitive load theory, we examine how worked example presentations— specifically how *extensive* and *dynamic* they are—may impact algebra learning in an online environment among middle and high school students.

# Cognitive Load Theory

Worked examples are considered to be an effective tool for learning because viewing worked examples reduces the cognitive load that is placed on students when problem-solving (Sweller, 1994). Providing a step-by-step example to reference frees up working memory and gives students more cognitive space for learning. As a result, students are more likely to learn from worked examples than from problem-solving alone, which is known as the worked example effect (Sweller, 2006).

Research has shown that the effect of worked examples on learning varies by students' prior knowledge (Renkl, 2014). For example, Kalyuga and colleagues (2001) found that students with lower prior knowledge benefited more from worked examples whereas students with higher prior knowledge benefited more from problem-solving. However, the relation between prior knowledge and the use of worked examples varies by the worked example format. For example, when worked examples contain errors, only students with sufficient prior knowledge benefit from them (Große & Renkl, 2007). Together, these findings suggest that learners with different levels of prior knowledge may benefit differently from differing formats of worked examples, and that learners with lower prior knowledge may benefit more from worked examples with more explicit instruction.

In order for worked examples to be most effective for learning, it is important to consider whether and how the presentation format and visual features of worked examples increase or decrease learners’ cognitive loads. Although worked examples are an effective instructional tool, the presentation of worked examples and instructional materials has been shown to impact learning gains and cognitive load (Sweller et al., 2019; Sweller 2020). For instance, students who viewed on-screen text that duplicated an audio explanation of how lightning forms scored lower on retention and knowledge transfer compared to their counterparts who did not view extra onscreen text, likely due to cognitive overload (Mayer et al., 2001). These findings have led to guidelines for creating worked examples, such as minimizing extraneous complexity, avoiding stimuli that split students’ attention or provide redundant information, and drawing attention to the subgoals of the problem (Schwartz, Tsang, & Blair, 2016; Sweller et al., 2019; Sweller, 2020).

While worked examples may reduce students’ cognitive load, it remains unclear which features of worked examples impact students’ cognitive load during algebra learning. We examine how extensive (i.e., how detailed) worked examples should be to maximize algebra learning. Research on element interactivity—number of elements that are simultaneously processed in working memory—in worked examples has shown that students learn more from scaffolded instruction over time. Specifically, students benefit from studying worked examples that present problem elements sequentially instead of simultaneously, which have higher element interactivity and task students to make connections while studying (Lu et al., 2020). Prior work with college students has also demonstrated that students outperform their peers on near and far transfer items after viewing worked examples which present information sequentially rather than simultaneously (Lusk & Atkinson, 2007). Further, novice learners benefit from exposure to worked examples that show extensive details between each step before studying worked examples that only show the major steps in an industrial skills course (Pollock et al., 2002), although less is known about how algebra students may benefit from different degrees of detail in worked examples. Because prior research has only examined the effects of concise worked examples on learning, it is unclear how much detail is ideal to present in a worked example. On one hand, concise worked examples may help focus students’ attention and avoid cognitive overload; on the other hand, extensive worked examples may help students offload the steps between major derivations onto the screen, and explicitly connect these derivations in the worked examples.

# Worked Examples in Online Learning Environments

Prior work has shown that worked examples in online learning environments, such as tutoring systems, are effective at decreasing instructional time and increasing learning (Salden et al., 2010). Therefore, it is important to consider how to effectively present worked examples to students as well as leverage the affordances of technologies for learning in online environments. Different from traditional learning with pencil and paper, one unique affordance provided by educational technologies is the ability to create worked examples with dynamic perceptual features for online viewing.

With the affordances of technologies, it may be worthwhile to consider designing worked examples from cognitive perspectives other than cognitive load. Prior research on perceptual learning has shown that individuals rely on visual cues in mathematics materials; consequently, using visual cues to direct students’ attention to relevant information can support their cognitive processes and increase learning (Gibson, 1969; Goldstone et al., 2017; Kirshner, 1989). For example, Harrison et al. (2020) showed that subtle manipulations to the spacing between terms in a mathematics problem can help guide learners to the correct calculations while problem-solving. Similar research has demonstrated the impact of spatial grouping on students’ problem-solving performance, providing evidence that minute visual features in instructional materials can impact student performance (Braithwaite et al., 2016; Landy & Goldstone, 2007, 2010). Extending beyond altering visual features in static equations, educational technologies allow us to guide students’ attention to important information through *motion* as they study worked examples with dynamic features.

Dynamic educational technologies such as Desmos (Ebert, 2014) and Graspable Math (Weitnauer et al., 2016) allow users to manipulate linear equations, graphs, and expressions, and see the outcomes of their actions in real-time on their computer screens. As interactive educational technologies like these become increasingly common, it is also possible to provide students with worked examples that demonstrate the dynamic process of solving algebraic equations (e.g., in 3*x* = 6, dragging 3 across the equal sign initiates the inverse operations and divides both sides by 3). Previous studies have shown the effectiveness of watching videos or animations of experts solving problems and providing explanations in an online environment (e.g., Wouters et al., [2008)](https://link.springer.com/article/10.1007/s10648-010-9145-4?shared-article-renderer#ref-CR25). These animations provide attentional cues, such as highlighting and motion, for students so that they can attend to the right information at the right times (Ayres &

Paas, 2007; De Koning et al., [2009)](https://link.springer.com/article/10.1007/s10648-010-9145-4?shared-article-renderer#ref-CR7).

# Current Study

Most prior research presents worked examples to students in a static fashion that displays the major steps taken to solve a problem. However, with new technology and interactive tools, we posit that there may be ways to leverage the affordances of technology in order to present worked examples that help students learn effectively in online environments. We explore these different presentations of worked examples through the lenses of cognitive load and perceptual learning theory by manipulating features of worked examples. We aim to determine whether cognitive load theory, perceptual learning, or a combination of both theories will provide the best explanation for how different presentation formats of worked examples impact student learning in algebra.

In this study, we vary worked example presentations to experimentally test their effects on learning and to identify elements of worked examples that are effective for learning (preregistered with SREE under Registry ID #1905.1v1). To account for the potential effect of students’ prior knowledge when exploring how different levels of detail and dynamicness impact learning, we recruited Algebra I students who were still learning to solve equations and controlled for students’ prior knowledge in the analyses. We compare posttest performance among middle and high school algebra students who complete an online problem set in one of six conditions that vary in the *extensiveness* and *dynamicness* of worked example presentations.

In the current study, we utilize images or videos of worked examples made using Graspable Math (GM), a freely available interactive algebra notation tool, to explore the potential benefits of showing the process of equation-solving through dynamic worked examples. Because dynamic worked examples animate actions involved in each step of the derivation, they may provide perceptual cues to the relevant information at the right time as students view the worked examples. In a randomized controlled study conducted in the ASSISTments platform, we compare the effect of traditional *static* worked examples to *sequential* and *dynamic* worked examples among middle and high school students. We include worked examples that present the derivation steps line-by-line in *sequential* order as an intermediate comparison between fully *static* and fully *dynamic* worked examples. We hypothesize that the *dynamic* worked examples may provide learners with additional perceptual cues above and beyond static and sequential worked examples, allowing them to clearly follow the connections between derivation steps..

Additionally, we compare *concise* and *extended* versions of traditional static worked examples (displaying complete derivations in one image) andsequential worked examples (showing derivations line by line in a looping GIF video), as well as two versions of dynamic worked examples (where screen casts were made as equations were manipulated and transformed in GM and displayed as a looping GIF video in ASSISTments) that vary in the amount of information presented on screen. We hypothesize that while extensive worked examples may provide more information, concise worked examples may be more beneficial for learning by minimizing the content on the screen.

Based on both cognitive load theory and perceptual learning theory, we hypothesize that students who view the *concise* and *dynamic* worked examples may demonstrate higher performance on the posttest compared to students in other conditions. Based on cognitive load theory, concise worked examples that only present the major derivation steps may reduce students’ cognitive load as they learn from worked examples. Based on perceptual learning theory, dynamic worked examples that animate problem-solving processes may provide the most perceptual support for students as they view worked examples.

**Methods**

# Participants

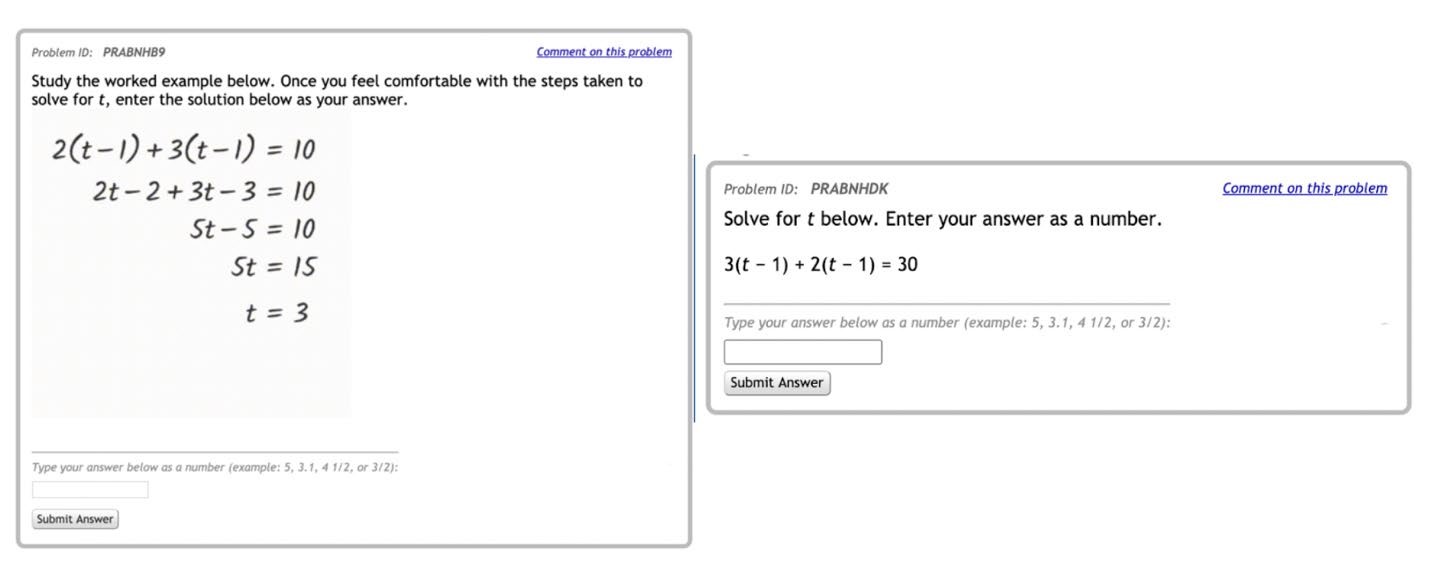
Participants were recruited from the ASSISTments teacher community through a monthly newsletter for teachers interested in supporting research through the ASSISTments platform. We provided a brief description of the study as well as a link to the problem set for teachers to assign to their Algebra I students in seventh to twelfth grade. A total of 25 teachers located in North America or Central America expressed interest and assigned the problem set to their students between October 2019 to March 2021. A total of 454 students started the assignment while only 300 completed the pretest and 230 completed the entire problem set. We note that 29 data points were excluded from analyses due to data logging errors, thus the final analytic sample comprised 230 students who completed the problem set and did not have data logging errors. Of the students who started the problem set, only 51% completed the problem set. The completion rate was relatively low, but comparable to that of other research studies conducted in ASSISTments. Specifically, in Feng and colleagues’ study (2021), only 49% of their participating students completed the problem set before the COVID-19 restrictions began in March 2020, and the completion rate decreased to 28% after March 2020. We conducted a chi-square test comparing the completion rate across the six conditions, and the differences between condition was not significant, 2(5, *N* = 454) = 4.23, *p* = .52.

This research was approved by the Institutional Review Board at a University in Northeastern United States. This research involved typical educational practices and did not require parental consent or student assent.

# Procedure

This study was conducted within ASSISTments, an online tutoring system (Heffernan & Heffernan, 2014; www.assistments.org). In ASSISTments, teachers can assign problem sets to students, and students can receive hints and correctness feedback during problem-solving. ASSISTments includes the technical infrastructure for conducting randomized controlled trials and is actively used by 500,000 students and 20,000 teachers around the world (ASSISTments, n.d.). The platform thus provides a convenient and ecologically valid context for researchers to study student learning and problem-solving in an online environment. We utilized ASSISTments as an online study platform in which students received the pretest, worked examples with paired practiced problems, and the posttest in one problem set. The problem set was designed to be completed online in one 60-minute session. Teachers assigned the online problem set for their students to complete in mathematics classrooms during instructional periods, and students worked individually at their own pace using their own device. If additional time was needed, students completed the assignment as homework or in a subsequent class period at the discretion of the teacher.

As students opened the problem set, they were randomly assigned to one of six worked example conditions. All students first completed an eight-item pretest on algebraic equationsolving. Next, they completed six problem pairs, each consisting of one worked example and one practice problem. The six pairs of worked examples and practice problems were presented in the same order across all conditions. Within each pair of worked examples and practice problems, students first studied the worked example for as long as they needed, then entered the answer of the worked example on the screen (e.g., 3 in Figure 1 Left). We asked students to enter the answer of each worked example to ensure that they looked at the worked example prior to solving the practice problem. Next, students solved a paired practice problem mirroring the structure of the worked example without the worked example in view (Figure 1 Right). All six practice problems were identical across conditions (Appendix A). Students then finished the problem set by completing a posttest mirroring the pretest. Students did not receive any accuracy feedback while completing the assignment.



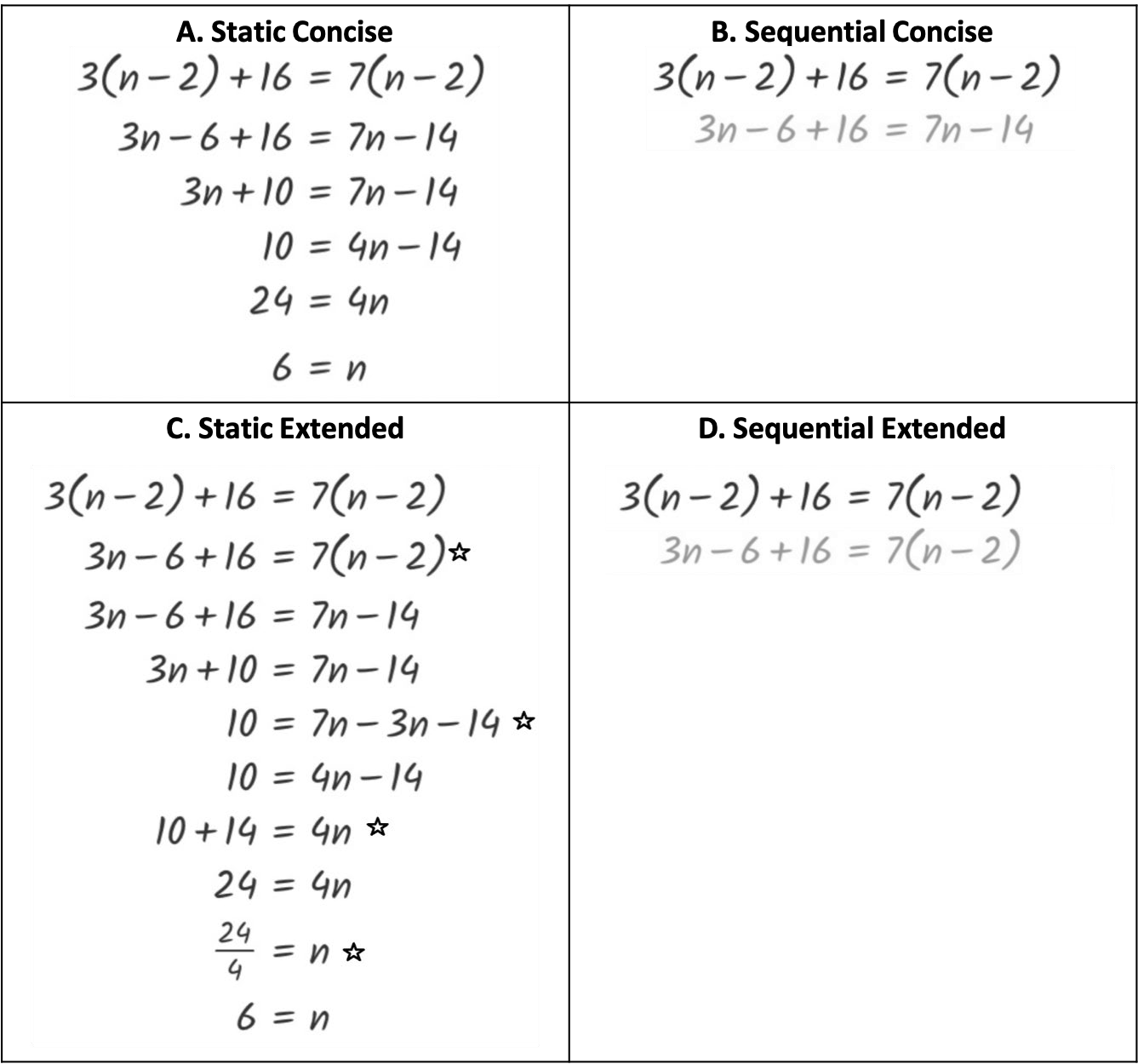
**Figure 1.** *Left:* Example of a static concise worked example as seen by a participant in the online tutoring system. *Right:* The following paired problem to be completed for practice.

# Pretest and Posttest

The eight-item pretest was constructed with six problems adapted from two open-source curricula, Engage New York (2014) and Utah Math Project (2016), and two problems designed by the authors (see Appendix B for all items). We selected and adapted six algebra equationsolving problems, then designed the remaining two problems following the structure of the worked examples. Of the eight problems, four had similar equation structures as the worked examples (items 1, 4, 7, 8) and the remaining four did not (items 2, 3, 5, 6). This design ensured that the pretest aligned with the content presented in our worked example study yet was representative of the problems students might encounter in classrooms. The eight-item posttest mirroring the pretest was then created by substituting numbers of similar magnitudes and maintaining the equation structures in the pretest problems. Each item was scored as correct (1) or incorrect (0), and the reliability of these eight items was KR-20 = 0.86 at pretest and KR-20 = 0.89 at posttest. The percent correct on pretest and posttest were included as the covariate and dependent variable, respectively, in the primary analyses.

# Experimental Conditions

We designed six conditions varying how *extended* and *dynamic* the worked examples were presented. The worked examples were adapted from Rittle-Johnson and Star (2007), a project aimed to improve middle-school students’ equation-solving performance. The starting equation of each worked example was identical for all students, and the presentation of the derivation steps varied across conditions. The first four conditions were based on a 2 (static vs. sequential) × 2 (concise vs. extended) design in which either the worked examples were presented fully in a static image, or each line of the worked examples was presented sequentially over time. Within static and sequential conditions, the worked examples showed either only the major derivation steps (concise) or all of the steps (extended; See Figure 2). Different from the first four conditions, the remaining two conditions involved dynamic presentations of worked examples. In both dynamic conditions, students saw all derivation steps for each worked example as well as the animation of the equation transformation processes. The two conditions varied in whether the history of the derivation remained on the screen or not (Figure 3). We describe each condition in detail below.



**Figure 2.** A worked example in the (A) static concise, (B) sequential concise, (C) static extended, and (D) sequential extended condition. In the sequential conditions, each step of the derivation is revealed over time. The stars in panel C indicate the additional derivation steps

displayed in the extended, but not concise, conditions.

**Static Concise.** A static concise worked example was presented as a static image that displayed the *major* steps in the derivation (Figure 2 A). The worked examples in this condition were identical to those used by Rittle-Johnson and Star (2007), except for the font in which the worked examples were presented. Rittle-Johnson and Star presented the worked examples in

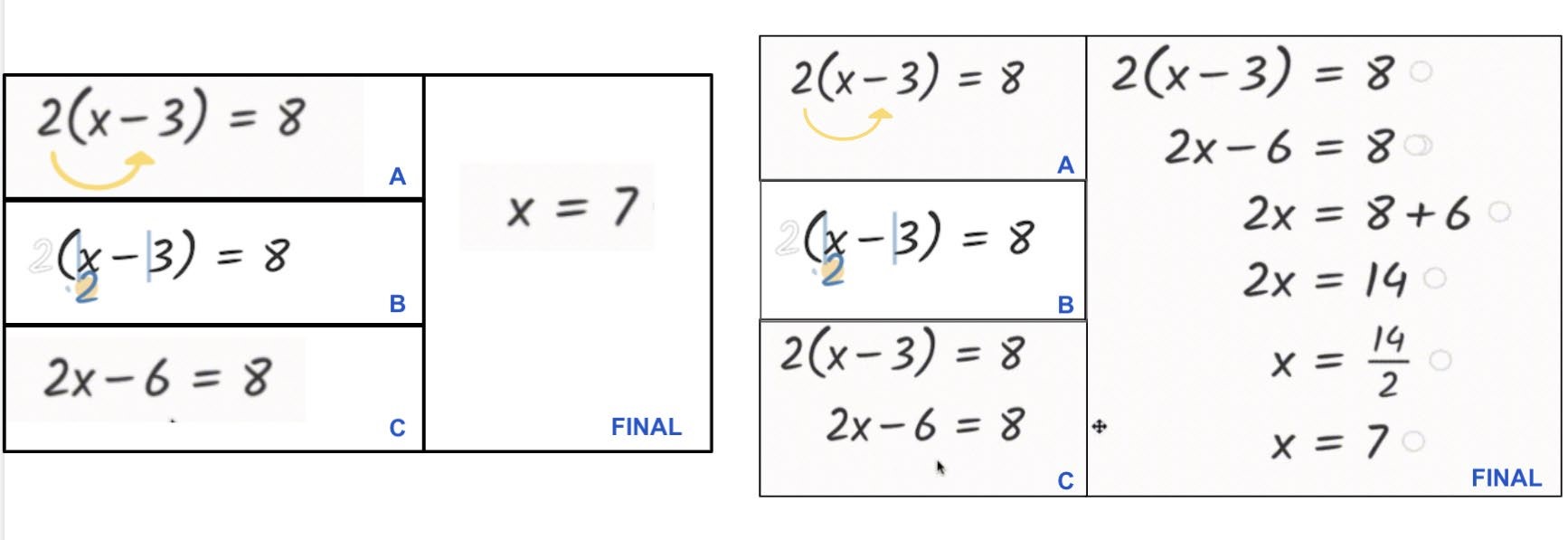
Times New Roman, whereas we presented the worked example in Kalam (the font type used in GM). We made this modification to match the font type of the worked examples across conditions. The static concise presentation of worked examples aligned with those used in textbooks (Engage New York, 2014) and other research studies (e.g., Sweller & Cooper, 1985).

**Static Extended.** A static extended worked example was presented as a static image that displayed *all* the steps in thederivation (Figure 2 Cx). Similar to the static concise worked example, each static extended worked example presented the derivation steps simultaneously in one static image. Unlike the static concise worked example, the static extended worked examples explicitly displayed each and every step of the derivation, extending beyond the scope of the concise conditions.

**Sequential Concise.** A sequential concise worked example was presented as a GIF video that displayed the *major* steps in the derivation one step at a time. The steps of the derivation presented in the sequential concise condition were identical to those in the static concise condition, but they appeared on the screen one line at a time in two- to three-second intervals, creating a history of the derivation over time. When the last step of the worked example (e.g., 6 =

*n*) was presented, the complete derivation remained on screen for five to seven seconds, allowing time for students to view the completed example. After which, the video automatically repeated from the beginning so that students could watch the video as many times as they wished.

**Sequential Extended.** A sequential extended worked example was presented as a GIF video that displayed *all* the steps in thederivation one step at a time. Each sequential extended worked example was identical to that in the sequential static condition, but the steps were added on the screen one at a time in two- to three-second intervals. Similar to the worked examples in the sequential concise condition, the complete derivation remained on-screen for five to seven seconds, then the video automatically repeated from the beginning.



**Figure 3.** *Left*: Dynamic no history condition in which all transformations occur on one line. *Right*: Dynamic history condition in which the result of each transformation is added as a new

line creating a sequential derivation of the worked example.

*Note*: Panels in both figures illustrate (A) the intended action of dragging 2 inside the parentheses, (B) the fluid transformation of distributing 2 over *x* and -3, (C) the result of the transformation, and (Final) the end result with the solution of the worked example.

**Dynamic History.** A dynamic history worked example was presented as a GIF video that displayed the process of solving an equation using the GM tool. For example, in 2(*x* − 3) = 8, students watched as the 2 was dragged into the parentheses and the equation was transformed into 2*x* − 6 = 8. The result of the transformation (e.g., 2*x* − 6 = 8) was then added as a new line in the derivation, and the next action was made on this new line of the derivation (Figure 3 Right).

As each line of the derivation was added, a history of the equation-solving process was created.

The complete derivations were identical to those in the two extended conditions (static extended and sequential extended), but actions and fluid transformations between each step of the derivation were demonstrated in the video. As the video reached the end of the equation-solving process, the complete derivation remained on-screen for five to seven seconds, then the video automatically repeated from the beginning.

**Dynamic No History.** Similar to the dynamic history worked examples, the dynamic no history worked examples were presented as a GIF video that displayed each step of the derivation that were created using the GM tool. However, instead of displaying the history of the derivation by adding a new line after each action, all the actions were made on the initial equation. Students watched the initial equation (e.g., 2(*x* − 3) = 8) being transformed into the answer (e.g., *x* = 7) over time in one line without the step-by-step history of the complete derivation. As the video reached the end of the equation-solving process, the solution remained on screen for five to seven seconds, then the video automatically repeated from the beginning.

# Approach to Analysis

Prior to testing our primary hypotheses, we first reported descriptive statistics of pretest and posttest scores by worked example conditions. We also conducted a paired-sample t-test to examine overall learning from worked examples regardless of the presentation conditions. Next, we conducted a one-way ANOVA to examine whether students’ pretest scores were comparable across conditions and to inform our primary analyses.

To test our hypotheses, we first conducted a one-way ANCOVA comparing students’ posttest scores across the six worked example conditions while controlling for their pretest scores. The post-hoc power analysis revealed that the sample size afforded 90% power to detect a moderate to large effect of worked example format (*f* ≥ .27) in the current study. The conventional cut-offs for small, medium, and large effects are 0.10, 0.25, and 0.40, respectively (Cohen, 1992). Next, to compare the effects of concise vs. extended, as well as static vs.

sequential formats, and to test the potential interaction between these two factors, we conducted a 2 (concise vs. extended) × 2 (static vs. sequential) ANCOVA controlling for pretest scores. This ANCOVA did not include either dynamic condition because they did not align with the concise vs. extended, or static vs. sequential, formats. Instead, to investigate how the extensiveness in dynamic worked examples may impact learning, we conducted a one-way ANCOVA comparing students’ posttest scores between the dynamic history and dynamic no history conditions while controlling for pretest score.

Finally, informed by the findings on the extensiveness in worked examples, we collapsed across concise and extended as well as history and no history conditions to investigate how the level of dynamicness impacts student learning from worked examples. We conducted a one-way ANCOVA comparing students’ posttest scores after viewing worked examples in static (concise and extended), sequential (concise and extended) and dynamic (history and no history) formats while controlling pretest scores. Collapsing across conditions allowed us to specifically examine the effect of the level of dynamicness in the worked example above and beyond the extensiveness.

To further explore the effects of worked example formats and to extend beyond aforementioned frequentist analyses, we included Bayesian statistics for all analyses (Faulkenberry et al., 2020). Bayesian approach allows us to go beyond the null results in frequentist analysis, and we are able to detect evidence in favor of the null hypothesis by comparing the strength and likelihood of both the alternative and null hypothesis models. In this way, we are able to determine whether the alternative or null hypothesis is more likely (Lakens et al., 2020). All analyses were conducted with JASP (JASP Team, 2020; Wagenmakers et al., 2018). We used the default, non-informative prior specifications in JASP for all Bayesian analyses as recommended by Faulkenberry and colleagues (2020). The default specification uses a JZS (multivariate Cauchy) prior on the effect scales with a default scale of 0.5. Per GossSampson and colleagues (2020), we used the scale of strength of evidence to interpret the Bayes factor (𝐵𝐵𝐵𝐵10). A value of 0 to 1 provides no evidence for either the alternative or the null hypothesis. The cutoffs of anecdotal, moderate, strong, very strong, and decisive evidence for the alternative hypothesis are 3, 10, 30, 100, and over 100, respectively. Similarly, the cutoffs of anecdotal, moderate, strong, very strong, and decisive evidence for the null hypothesis are 0.33,

0.10, 0.033, 0.01, and less than 0.01, respectively.

**Results**

# Preliminary Analysis

The descriptive statistics revealed that the pretest and posttest scores were not subject to ceiling or floor effects (Table 1). On average, students scored 41% (*SD* = 34%) on the pretest and 48% (*SD* = 37%) on the posttest. The scores were widely distributed, as indicated by the standard deviation, suggesting that the pretest and posttest captured the variability in the students’ equation-solving performance within the sample.

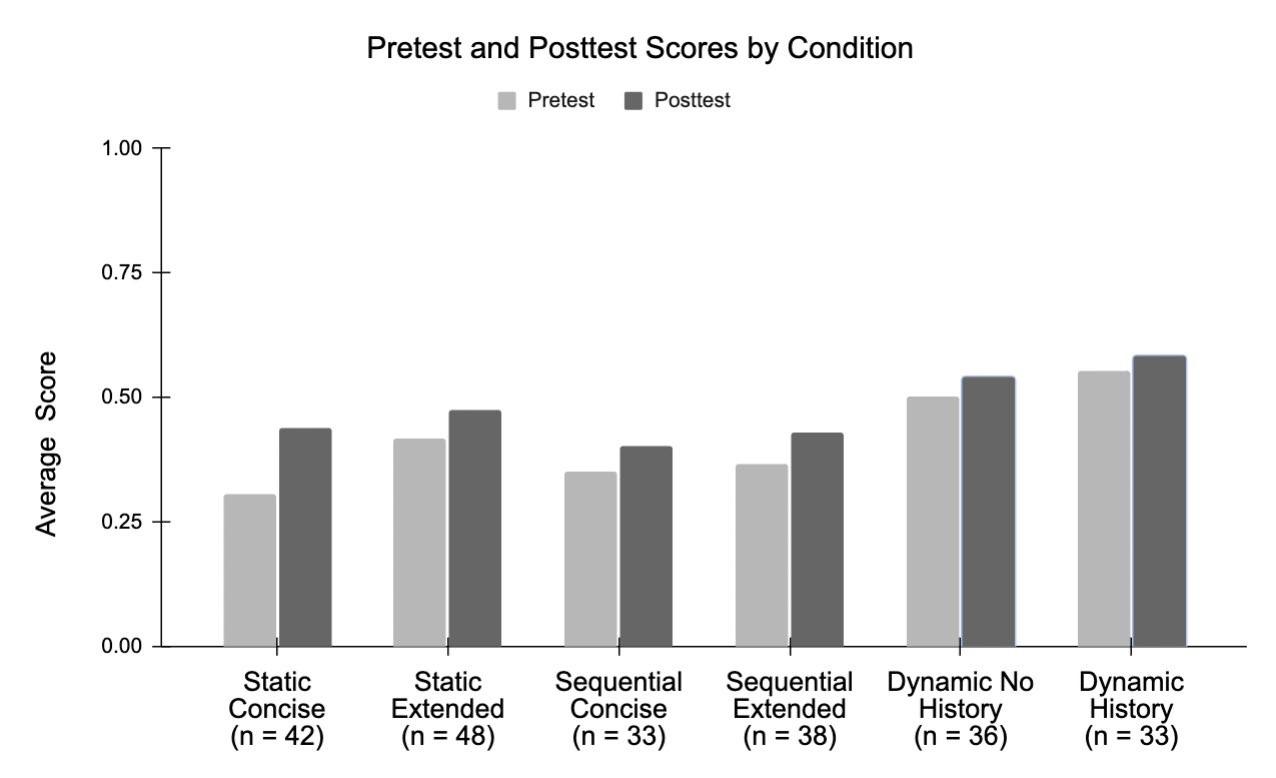
A paired-sample t-test revealed that, regardless of conditions, students significantly improved from pretest to posttest, *t*(229) *=* 3.69 *, p <* .01, *d* = .24. In addition, a Bayesian paired samples t-test revealed a Bayes factor of 𝐵𝐵𝐵𝐵10= 50.99, providing very strong evidence that students scored higher on the posttest compared to pretest. Prior to conducting our primary analyses, we first examined whether students’ pretest scores were comparable across the six conditions. A one-way ANOVA revealed that students’ pretest scores did significantly differ by condition, *F*(5,224) = 2.91, *p* = .015. Further, the data provided anecdotal evidence for the differences in pretest scores by condition, 𝐵𝐵𝐵𝐵10= 1.62. The post-hoc pairwise comparisons with Bonferroni correction revealed that the difference between static concise and dynamic history conditions were significant, *p* = .02. The remaining pairwise comparisons were not significant, *p*s > .10. Due to the differences in pretest scores across conditions, we controlled for pretest scores to test the effect of work example conditions on the posttest performance in the following analyses.

Table 1. Means and standard deviations of pretest and posttest scores by condition

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | Pretes | t | P | osttest |  |
| Condition | *N* | *M* | *SD* | *M* | *SD* |
| Static Concise | 42 | 0.307 | 0.313 0.440  0.318 0.474 | | 0.333 |
| Static Extended | 48 | 0.417 | 0.376 |
| Sequential Concise | 33 | 0.352  0.366 | 0.354 0.402 | | 0.395 |
| Sequential Extended | 38 | 0.314 0.431  0.337 0.583 | | 0.370 |
| Dynamic History | 33 | 0.553 | 0.342 |
| Dynamic No History | 36 | 0.503 | 0.376 0.542 | | 0.380 |
|  |  |  |  | |  | |

# Condition Effects

A one-way ANCOVA comparing students’ posttest scores by the six worked example conditions while controlling for their pretest scores revealed that there was no main effect of condition, *F*(5,223) = .38, *p* = .86. 𝜂𝜂2= .004. Students’ posttest performance did not significantly differ by worked example conditions (Figure 4). Further, a Bayesian ANCOVA revealed strong evidence that there was no effect of condition on posttest scores, 𝐵𝐵𝐵𝐵10= .076. As expected, students’ pretest score was a significant and positive covariate of their posttest score, *F*(1,223) = 234.56, *p* < .01. 𝜂𝜂2= .51. Bayesian analysis confirmed this result with decisive evidence that pretest score was associated with posttest score, (𝐵𝐵𝐵𝐵10>100).



**Figure 4.** Pretest and posttest scores by worked example presentation conditions.

# Extensiveness: Concise vs. Extended and History vs. No History

Next, a 2 (extensiveness: concise vs. extended) × 2 (dynamicness: static vs. sequential)

ANCOVA controlling for pretest score revealed that there was no main effect of extensiveness, *F*(1, 156) = .148, *p* = .701, 𝜂𝜂2< .001. A Bayesian ANCOVA revealed moderate evidence that extensiveness had no effect on posttest scores, 𝐵𝐵𝐵𝐵10= .196. There was no main effect of dynamicness, *F*(1, 156) = .831, *p* = .363, 𝜂𝜂2= .003, with moderate evidence supporting this null finding, 𝐵𝐵𝐵𝐵10= .216. There was also no interaction effect, *F*(1, 156) = .668, *p* = .415, 𝜂𝜂2= .002, with very strong evidence supporting this null finding, 𝐵𝐵𝐵𝐵10= .01. As expected, students’ pretest score was a significant and positive covariate of their posttest score, *F*(1,156) = 135.216, *p* < .01, 𝜂𝜂2= .462, which was confirmed by the Bayes factor indicating decisive evidence, 𝐵𝐵𝐵𝐵10>

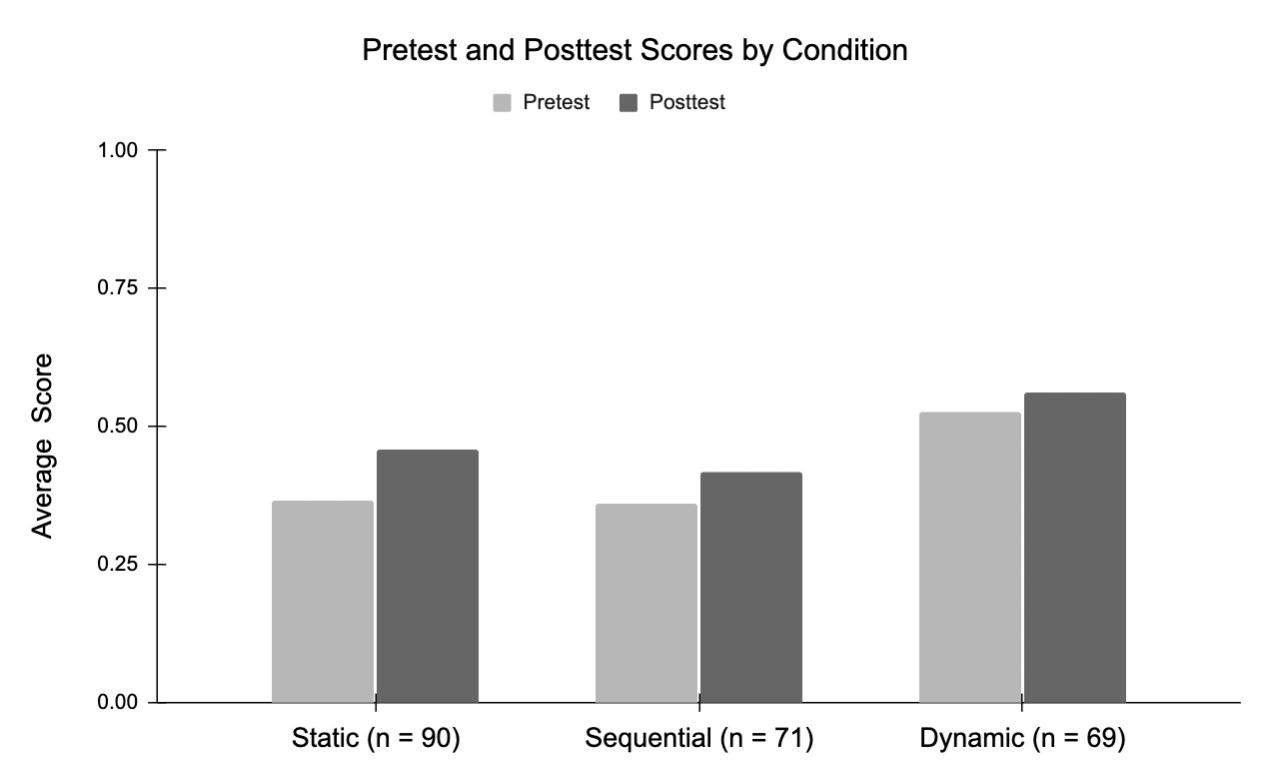
100.

We also conducted a one-way ANCOVA comparing students’ posttest scores in dynamic no history and dynamic history conditions while controlling for their pretest scores. The

ANCOVA revealed that the effect of condition was not significant, *F*(1, 66) = .001, *p* = .97, 𝜂𝜂2< .01. Students in the dynamic no history condition did not outperform students in the dynamic history condition. The Bayes factor indicated moderate evidence for this null finding, 𝐵𝐵𝐵𝐵10= .273. As expected, students’ pretest score was a significant and positive covariate of their posttest score, *F*(1,66) = 112.26, *p* < .01, 𝜂𝜂2< .63, which was confirmed by the Bayes factor indicating decisive evidence, 𝐵𝐵𝐵𝐵10> 100.

# Dynamicness: Static vs. Sequential vs. Dynamic

Finally, a one-way ANCOVA comparing students’ posttest scores in static, sequential, and dynamic presentation formats while controlling for their pretest scores revealed that there was no main effect of presentation format, *F*(2,226) = .43, *p* = .65, 𝜂𝜂2= .002. Students’ posttest performance did not significantly differ based on how dynamic the worked examples were presented (Figure 5). A Bayesian ANCOVA revealed anecdotal evidence (𝐵𝐵𝐵𝐵10= .607) in favor of this null finding. As expected, students’ pretest score was a significant and positive covariate of their posttest score, *F*(1,226) = 236.64, *p* < .01, 𝜂𝜂2= .51, which was confirmed by the Bayes factor indicating decisive evidence for this finding, 𝐵𝐵𝐵𝐵10> 100.



**Figure 5.** Pretest and posttest scores by static, sequential and dynamic conditions.

# Discussion

In this study, we set out to identify which presentation of worked examples was most beneficial for students to view and to test how different presentation features impact learning from worked examples in an online platform. We found that on average, students improved their algebraic equation-solving performance from pretest to posttest after completing a brief activity with worked examples and paired practice problems. Extending the literature on worked examples, we found that, after controlling for pretest scores, students’ posttest scores did not significantly differ between (a) the six worked example conditions, (b) the concise vs. extended or history vs. no history presentation formats, or (c) the static vs. sequential vs. dynamic presentation formats. These findings were further strengthened by the Bayesian analyses providing strong evidence that (a) students’ posttest scores did not differ across six conditions, and moderate evidence that (b) the extensiveness and (c) the level of dynamicness did not differentially impact students’ posttest scores. Students across six conditions improved their performance from pretest to posttest, suggesting that the worked examples, regardless of their formats, were effective. These results contribute findings to cognitive load theory and perceptual learning theory that may guide future designs of worked examples. We discuss these findings and their implications in detail below.

# Worked Example Effect

We found that students, averaging across six conditions, experienced learning gains from pretest to posttest which aligns with prior research on the worked example effect (Booth et al., 2013; Carroll, 1994; Foster et al., 2018). Given that the worked example effect is well substantiated in the literature, the current experiment explored the differential effects of worked example presentations with one brief activity that included an immediate pretest and posttest. Our effect size for the pretest vs. posttest comparison across six conditions (*d* = .24) was comparable to that of a prior study with Algebra I students in an online intervention study (*d* = .19; Booth et al., 2013), aligning with prior literature on the worked example effect. Together, our finding suggests that the learning gains may be attributed to students’ participation in the worked example activity regardless of the worked example formats.

Given that the results show ubiquitous learning gains without reliable differences by condition and strong evidence suggesting the lack of condition effect, it was possible that different presentations of the same worked example may be equally beneficial for student learning. This finding aligns with prior research demonstrating that students benefit from studying different worked example presentations and that there are reliable differences between conditions at posttest or a delayed posttest (Reed et al., 2013). One potential implication of this finding is that teachers and content developers may have the flexibility to utilize various technologies and formats, including dynamic educational technologies, to create different presentations of worked examples for online settings. This flexibility may also allow students more autonomy in their choice of worked example presentations and benefit from the additional level of choice during online practice. Prior work has demonstrated that students who were allowed to choose their feedback format (text or video) after mathematics problem-solving in an online problem set outperformed their peers who were randomly assigned to a feedback format (Ostrow & Heffernan, 2015). Similar to Ostrow and Heffernan’s finding that feedback formats, when randomly assigned, did not impact students’ learning, we found that worked example formats did not impact student learning of algebraic equation-solving. A future direction is to conceptually replicate this research by providing students with choices of worked example formats and examining the impacts of choice on student learning.

**Comparing Extensiveness and Dynamic Presentations of Worked Examples**  We hypothesized that concise worked examples may be more beneficial for student learning compared to extended worked examples. However, when comparing worked examples that presented the concise derivation vs. complete derivation of an algebra problem, we did not find differences in students’ posttest scores and there was moderate evidence in support of this null finding. We also did not find an effect of showing vs. not showing the history of dynamic worked examples on students’ posttest scores and there was moderate evidence in support of this finding. The lack of an effect combined with the moderate evidence of the null findings on concise vs. extended and history vs. no history worked examples suggest that perhaps the extensiveness of a worked example may not impact student learning.

Alternatively, the extensiveness of a worked example may differentially impact student learning at different stages. Pollock et al. (2002) found that novice learners in a high school industrial skills course benefited from exposure to worked examples with extensive details before studying worked examples without extensive details. Perhaps, novice learners may benefit from first studying the extended worked examples, then gradually transitioning to concise worked examples. However, among our sample of Algebra I students who were still learning algebraic equation-solving, we did not find that students benefited more from the extended worked examples compared to the concise worked examples. Because our sample was algebra learners and the number of students within each condition was relatively small, we did not further explore the effect of concise vs. extended worked examples on students with varying levels of prior knowledge. Future studies should recruit a larger sample to examine the effect of concise vs. extended worked examples among students at different phases of algebra learning. If the later findings on algebra learning replicate Pollock and colleagues’ work, teachers and content designers may consider varying the extensiveness of the worked examples based on students’ knowledge level in order to better support learning through individualized practices. Based on perceptual learning theory (Gibson, 1969), we initially hypothesized that worked examples with dynamic presentation may direct students’ attention to important pieces of notation and support learning beyond viewing static or sequential worked examples. Contrary to our hypothesis, we found that there were no significant differences between worked examples presented in static, sequential, or dynamic formats and that there was anecdotal evidence in support of no differences between conditions. One possible interpretation is that these nuanced variations of worked example formats may be equally effective for student learning. If so, students, teachers, and content designers may have the flexibility to choose their preferred worked example when teaching or learning. Alternatively, different factors may contribute to the seemingly comparable effects of worked example formats. In particular, based on anecdotal feedback from participating teachers, videos of the dynamic worked examples might be too fast for students to follow the derivation steps. Perhaps, dynamic worked examples do direct students’ attention to the important problem-solving procedures; however, without a pause button to control the speed of the video, the dynamic worked examples may inadvertently increase students’ cognitive load, preventing the dynamic, fluid transformations from being more helpful for learning beyond viewing static examples. Another explanation may be that students are used to static worked examples in their textbooks and curricular materials, and the novelty of the videos may have decreased the initial impact of the dynamic worked examples. If this is the case, providing students with more experience viewing the dynamic worked examples by increasing the number of worked examples and the learning sessions in the study should improve student learning above and beyond viewing static worked examples. With the anecdotal evidence for the null finding, the current result is inconclusive. With a larger sample of students, future studies should further investigate the potential affordances of dynamic worked examples, and how to maximize their effectiveness for student learning.

# Limitations and Future Directions

Several limitations warrant mentioning. First, approximately half of the students enrolled in the study did not complete the pretest and were excluded from the analyses. The majority

(69%) of these students excluded from the analyses dropped out of the study prior to viewing the worked examples, suggesting that the attrition may not be associated with the experimental conditions. Even though we asked teachers to dedicate instructional time for students to complete the study, teachers were likely to assign the study problem set as a supplemental practice or ungraded homework assignment within ASSISTments, providing little motivation for students to complete the entire problem set. Further, this study was conducted in schools during the COVID19 pandemic where there was a lot of disruption to typical everyday practice; therefore, the disruption related to COVID-19 may also contribute to the high attrition rate. Given that the completion rate of our study was comparable to that of a prior study using ASSISTments (e.g., Feng et al., 2021), we posit that the relatively low completion rate might not be specific to our study but more general to the platform.

Next, after accounting for attrition, the sample size within each worked example condition was relatively small, and the experimental manipulation on the worked example format was modest, with overlapping manipulations to each condition. It also is possible that some idiosyncratic aspects of the worked example formats, such as the speed of the videos, or students’ individual characteristics, such as their learning phases and preferences of worked examples, contributed to how much students learned from the worked examples in the current study. Although a power analysis suggested that we had adequate power to detect a moderate to large effect of the worked example formats, we might still be underpowered to detect the nuanced effects from a brief worked example session. Further, the worked examples in the current study only focused on exemplifying the equation-solving process within algebra. Therefore, the findings cannot be generalized to other types of worked examples that focused on learning and strategies, or other non-algorithmic domains, such as argumentation (e.g., Schworm

& Renkl, 2007) and medical diagnostics (e.g., Stark et al., 2011). Future studies should investigate the impact of various types of worked examples, content areas, as well as students’ prior knowledge, choice, feedback, and perceived helpfulness of the worked examples to further investigate how these factors together impact student learning from worked examples.

Finally, our findings suggest that the extensiveness or dynamicness of worked examples may not significantly impact students’ algebra learning. However, a main limitation is that this study was conducted as a brief online intervention within the context of algebra worked examples. The current study provides a starting point for future work to further investigate how these factors impact student learning from worked examples. Future studies should replicate current findings with a larger sample, a longer-term intervention over multiple class periods, and delayed posttests, comparable to related research on the effects of worked examples in online settings (e.g., Booth et al., 2013; McLaren et al., 2016; Reed et al., 2013). Doing so will provide stronger and more conclusive evidence of how the extensiveness or dynamicness of worked examples may impact students’ algebra learning.

Future work should examine other outcomes, such as learning efficiency, in addition to posttest performance, as a result of participating in worked example practice (McLaren et al., 2016). Prior studies have found that students achieved the same level of performance in less amount of time when viewing worked examples compared to solving practice problems (Salden et al., 2010). Although we did not find an effect of worked example formats on posttest performance, there may be an effect of formats on learning efficiency such that some presentations may require less study time than others to be beneficial. Identifying the unique affordances of different worked example presentations will further inform the design of worked examples in online learning environments and better support student learning. Future work should also examine the mechanisms through which these worked example formats impact learning. Another future direction is to measure students’ cognitive load or eye gaze in order to further investigate the relations between cognitive load theory, perceptual learning theory, and the worked example effect.

# Conclusion

We found that, on average, students improved from pretest to posttest after completing a problem set with one of six different presentations of worked examples, suggesting that regardless of worked example format, students improved their algebra performance through this online activity. In addition, we found strong evidence for the result that student learning did not significantly differ between experimental conditions, leading us to conclude that worked examples, regardless of their formats, may be effective learning tools.

These results have implications for designing worked examples based on principles of cognitive load theory and perceptual learning theory, as both theories have informed teaching and learning. It seems that including more details in worked examples may not decrease student performance compared to viewing worked examples with less details. Further, including perceptual features to guide students’ attention in worked examples may not lead to additional gains beyond the traditional static worked examples. These results have implications for classrooms and the design of online worked examples in tutoring systems. Specifically, teachers and content creators may be able to design different presentations of worked examples that will still be effective for learning.

# References

ASSISTments. (n.d.). ASSISTments. Retrieved May 5, 2021, from https://new.assistments.org/

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples:

Instructional principles from the worked examples research. *Review of Educational*

*Research, 70*(2), 181-214. doi:10.3102/00346543070002181

Ayres, P., & Paas, F. (2007). Making instructional animations more effective: A cognitive load approach. *Applied Cognitive Psychology*, *21*(6), 695-700.

https://doi.org/10.1002/acp.1343

Barbieri, C. A., & Booth, J. L. (2020). Mistakes on display: Incorrect examples refine equation solving and algebraic feature knowledge. *Applied Cognitive Psychology*, *34*(4), 862-878. https://doi.org/10.1002/acp.3663

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.

https://doi.org/10.1016/j.learninstruc.2012.11.002

Braithwaite, D. W., Goldstone, R. L., van der Maas, H. L. J., & Landy, D. (2016). Non-formal mechanisms in mathematical cognitive development: The case of arithmetic. *Cognition*,

*149*, 40–55. https://doi.org/10.1016/j.cognition.2016.01.004

Carroll, W. M. (1994). Using worked examples as an instructional support in the algebra classroom. *Journal of Educational Psychology*, *86*(3), 360–367.

<https://doi.org/10.1037/0022-0663.86.3.360>

Cohen, J. (1992). Statistical power analysis. Current directions in psychological science, 1(3),

98-101.

De Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design.

*Educational Psychology Review*, *21*, 113-140. https://doi.org/10.1007/s10648-009-9098-

7

Ebert, D. (2014). Graphing projects with Desmos. *The Mathematics Teacher, 108*(5), 388-391.

https://www.jstor.org/stable/10.5951/mathteacher.108.5.0388

Engage New York (2014). New York State Education Department. Retrieved from https://www.engageny.org

Faulkenberry, T. J., Ly, A., & Wagenmakers, E. J. (2020). Bayesian inference in numerical cognition: A tutorial using JASP. *Journal of Numerical Cognition*, *6*(2), 231-259.

Feng, M., Heffernan, N., Wang, H., & Collins, K. (April, 2021) Boosting middle school math learning with online homework. In Li, L. & Matlen, B.(Chair), How can children optimize their online learning experience with educational technology? *Symposium presented at the Society for Research in Child Development*. Minneapolis, MN.

Foster, N. L., Rawson, K. A., & Dunlosky, J. (2018). Self-regulated learning of principle-based concepts: Do students prefer worked examples, faded examples, or problem solving?

*Learning and Instruction*, *55*, 124–138. https://doi.org/10.1016/j.learninstruc.2017.10.002

Gibson, E. J. (1969). *Principles of perceptual learning and development*. Englewood Cliffs, NJ:

Prentice Hall.

Goldstone, R. L., Marghetis, T., Weitnauer, E., Ottmar, E. R., & Landy, D. (2017). Adapting perception, action, and technology for mathematical reasoning. *Current Directions in Psychological Science,* 26(5), 434–441. <https://doi.org/10.1177/0963721417704888>

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

Goss-Sampson, M. A., van Doorn, J., & Wagenmakers, E. J. (2020). Bayesian inference in

JASP: A guide for students. *University of Amsterdam: JASP team*.

Harrison, A., Smith, H., Hulse, T., & Ottmar, E. R. (2020). Spacing out! manipulating spatial features in mathematical expressions affects performance. *Journal of Numerical*

*Cognition*, *6(2)*, 186–203. https://doi.org/10.5964/jnc.v6i2.243

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education*, *24*,

470–497. https://doi.org/10.1007/s40593-014-0024-x

JASP Team (2020). JASP (Version 0.14.1)[Computer software].

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, *93*(3), 579.

Kirshner, D. (1989). The visual syntax of algebra. *Journal for Research in Mathematics*

*Education*, *20*(3), 274–287. <https://doi.org/10.2307/749516>

Lakens, D., McLatchie, N., Isager, P. M., Scheel, A. M., & Dienes, Z. (2020). Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of*

*Gerontology: Series B*, 75(1), 45-57.

Landy, D., & Goldstone, R. L. (2007). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning Memory and Cognition*, *33*(4), 720–733.

https://doi.org/10.1037/0278-7393.33.4.720

Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly*

*Journal of Experimental Psychology*, *63*(10), 1953–1968.

https://doi.org/10.1080/17470211003787619

Lu, J., Kalyuga, S., & Sweller, J. (2020). Altering element interactivity and variability in example‐practice sequences to enhance learning to write Chinese characters. *Applied*

*Cognitive Psychology*, *34*(4), 837-843. https://doi.org/10.1002/acp.3668

Lusk, M. M., & Atkinson, R. K. (2007). Animated pedagogical agents: Does their degree of embodiment impact learning from static or animated worked examples? *Applied*

*Cognitive Psychology*, *21*(6), 747–764. https://doi.org/10.1002/acp.1347

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational*

*Psychology*, *93*(1), 187-198. doi:10.1037/0022-0663.93.1.187

McLaren, B. M., van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human*

*Behavior*, *55*, 87-99. https://doi.org/10.1016/j.chb.2015.08.038

Ostrow, K. S. & Heffernan, N. T. (2015). The role of student choice within adaptive tutoring. In

C. Conati, N. Heffernan, A. Mitrovic & M. Verdejo (Eds.) *Proceedings of the 17th International Conference on Artificial Intelligence in Education* (AIED 2015). Springer

International Publishing. Madrid, Spain. June 22-26. pp. 752- 755.

https://doi.org/10.1007/978-3-319-19773-9\_108

Pollock, E., Chandler, P., & Sweller, J. (2002). Assimilating complex information. *Learning and*

*Instruction*, *12*(1), 61-86. https://doi.org/10.1016/S0959-4752(01)00016-0

Reed, S. K., Corbett, A., Hoffman, B., Wagner, A., & MacLaren, B. (2013). Effect of worked examples and Cognitive Tutor training on constructing equations. *Instructional Science*,

*41*(1), 1–24. https://doi.org/10.1007/s11251-012-9205-x

Renkl, A. (2014). Learning from worked examples: How to prepare students for meaningful problem solving. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum* (p. 118–130). Society for the Teaching of Psychology.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations.

*Journal of Educational Psychology*, *99*(3), 561–574. https://doi.org/10.1037/0022-

0663.99.3.561

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

https://doi.org/10.1037/a0016026

Salden, R. J., Koedinger, K. R., Renkl, A., Aleven, V., & McLaren, B. M. (2010). Accounting for beneficial effects of worked examples in tutored problem solving. *Educational*

*Psychology Review*, *22*(4), 379-392. <https://doi.org/10.1007/s10648-010->

Schworm, S., & Renkl, A. (2007). Learning argumentation skills through the use of prompts for self-explaining examples. *Journal of Educational Psychology*, 99(2), 285.

Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. WW Norton & Company.

Star, J. R., Pollack, C., Durkin, K., Rittle-Johnson, B., Lynch, K., Newton, K., & Gogolen, C.

(2015). Learning from comparison in algebra. *Contemporary Educational Psychology*,

*40*, 41–54. <https://doi.org/10.1016/j.cedpsych.2014.05.005>

Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education.

*Learning and instruction*, 21(1), 22-33.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*, 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2006). The worked example effect and human cognition. *Learning and Instruction*,

*16*(2), 165–169. https://doi.org/10.1016/j.learninstruc.2006.02.005

Sweller, J. (2020). Cognitive load theory and educational technology. *Educational Technology*

*Research and Development*, *68*(1), 1-16. https://doi.org/10.1007/s11423-019-09701-3

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.

https://doi.org/10.1207/s1532690xci0201\_3

Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, *31*(2), 261-292.

https://doi.org/10.1007/s10648-019-09465-5

Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2020). Individual differences in the preference for worked examples: Lessons from an application of dispositional learning analytics.

*Applied Cognitive Psychology*, *34*(4), 890-905. https://doi.org/10.1002/acp.3652

Utah Math Project (2016). *The Utah Middle School Math Project.* Retrieved from http://utahmiddleschoolmath.org/%0A

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Morey, R. D.

(2018). Bayesian inference for psychology. Part II: Example applications with JASP.

*Psychonomic Bulletin & Review*, *25*(1), 58-76.

Weitnauer, E., Landy, D., & Ottmar, E. R. (2016). Graspable Math: Towards dynamic algebra notations that support learners better than paper. In Future Technologies Conference (pp.

406–414). San Francisco.

Wouters, P., Paas, F., & van Merriënboer, J. J. (2008). How to optimize learning from animated models: A review of guidelines based on cognitive load. *Review of Educational*

*Research*, *78*(3), 645-675. https://doi.org/10.3102/0034654308320320

**Appendix A.** Algebraic equations used in worked examples across all conditions and the paired problem shown after each worked example.

|  |  |
| --- | --- |
| Worked Example (concise format) | Paired Practice Problem |
| 2(*x* − 3) = 8  2*x* − 6 = 8 2*x* = 14 *x* = 7 | − 3(*y* − 4) = 18  − 3*y* + 12 = 18 − 3*y* = 6 *y* = − 2 |
| 2(*t* − 1) + 3(*t* − 1) = 10  2*t* − 2 + 3*t* − 3 = 10  5*t* - 5 = 10 5*t* = 15 *t* = 5 | 3(*t* − 1) + 3(*t* − 1) = 30  3*t* − 3 + 3*t* − 3 = 30  6*t* - 6 = 30 6*t* = 36 *t* = 6 |
| 5(*y* + 1) = 3(*y* + 1) + 8  5*y* + 5 = 3*y* + 3 + 8  5*y* + 5 = 3*y* + 11  2*y* + 5 = 11 2*y* = 6 *y* = 3 | 5(*m* + 4) = 2(*m* + 4) + 15  5*m* +20 = 2*m* + 8 + 15  5*m* + 20 = 2*m* + 23  3*m* + 20 = 23 3*m* = 3 *m* = 1 |
| 3(*n* − 2) + 16 = 7(*n* − 2)  3*n* − 6 + 16 = 7*n* − 14  3*n* + 10 = 7*n* − 14  10 = 4*n* − 14  24 = 4*n*  6 = *n* | 3(*n* − 2) + 12 = 6(*n* − 2)  3n − 6 + 12 = 6*n* − 12  3*n* + 6 = 6*n* − 12  6 = 3*n* − 12  18 = 3*n*  6 = *n* |
| 9 = 5(*m* + 2) + 4(*m* + 2)  9 = 5*m* + 10 + 4*m* + 8  9 = 9*m* + 18  − 9 = 9*m*  − 1 = *m* | 9 = 3(*y* + 5) + 6(*y* +5)  9 = 3*y* + 15 + 6*y* + 30  9 = 9*y* + 45  − 36 = 9*y*  − 4 = *y* |
| 3(*h* − 2) + 5(*h* − 2) = 24  3*h* − 6 + 5*h* -10 = 24  8h − 16 = 24 8*h* = 40 *h* = 5 | 6(*w* − 4) + 7(*w* − 4) = 26  6*w* − 24 + 7*w* − 28 = 26  13*w* − 52 = 26 13*w* = 78 *w* = 6 |
| Note: The worked example derivations are those used by Rittle-Johnson and Star (2007) and do | |

not include the extra steps included in the derivations for the extended worked example conditions.

**Appendix B.** Pretest and Posttest Items.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pretest Items | Source | Posttest Items | Open Source | |
| 8(2𝑥𝑥 + 9) = 56 | Engage NY | 11(𝑥𝑥 + 10) = 132 | Engage NY | |
| −(*x* − 5) + 2 − *x* = 3 | Project Utah | −(4*x* − 10) + 4 − 4*x* = 6 | -- | |
| 5 − 4(2*b* − 5) + 3*b* = 15 | Project Utah | 30 − 4(*b* − 5) + 1*b* = 20 | -- | |
| 10 = 3(*x* − 2) − 2(5*x* − 1) | Project Utah | 20 = 3(2*x* − 2) – 2(5*x* − 1) | -- | |
| 3(2𝑥𝑥 − 14) + 𝑥𝑥 = 15− (−9𝑥𝑥 − 5) | Engage NY | 6 (4*x −* 28) + 2*x* = 30 – ( –18*x* – 10) | -- | |
| −4𝑥𝑥 − 2(8𝑥𝑥 + 1) = −(−2𝑥𝑥 −10) | Engage NY | −6*x −* 4(3*x* + 2) = − ( −1*x* − 2) | -- | |
| 5(*y* – 12) = 3(*y* – 12) + 20 | Authors | 2(*y* – 4) = (*y* – 4) + 6 | -- | |
| 3(*h* + 2) + 4(*h* + 2) = 35 | Authors | 2(*h* + 1) + 4(*h* + 1) = 12 | -- | |
|  | | | |  |