COLLEGE STUDENTS' INPUT ON THE DESIGN OF WORKED EXAMPLES FOR ONLINE ENVIRONMENTS

Avery H. Closser
Worcester Polytechnic Institute
ahclosser@wpi.edu

Jenny Yun-Chen Chan
Worcester Polytechnic Institute
jchan2@wpi.edu

Hannah Smith
Worcester Polytechnic Institute
hsmith2@wpi.edu

Erin Ottmar
Worcester Polytechnic Institute
erottmar@wpi.edu

Worked examples have been shown to improve student learning in algebra. However, less is known about how to design worked examples to support student learning in online settings. We explore how college students react to worked examples that vary in their degree of extensiveness and dynamicness. In an online, within-subjects study, 109 college students viewed six worked example presentations: 1) static concise, 2) static extended, 3) sequential concise, 4) sequential extended, 5) dynamic history, and 6) dynamic no history. Students rated the helpfulness of each worked example and explained their rating. We found that students rated the static concise presentation as the most helpful and the dynamic no history presentation as the least helpful example. Responses were coded by researchers for common themes and revealed insights that may inform how researchers and teachers design worked examples for online environments.

Keywords: Learning Theory, Instructional Activities and Practices, Metacognition, Technology.

Purpose of the Study

Worked examples are an effective means of instructional support for math education (e.g., Booth et al., 2015; Carroll, 1994; Foster et al., 2018). However, worked examples have primarily been presented the same way, with research focusing on how static images of worked examples that display major derivations align with students' underlying cognitive mechanisms and in turn support learning (e.g., Sweller, 2006, 2020). Recently, Scheiter (2020) called for broadening the research on example-based learning beyond pure cognitive mechanisms to incorporate non-cognitive theories. While some affective factors have been investigated in relation to example-based learning (Hartmann et al., 2020; Tempelaar et al., 2020), we propose that analyzing students' perceptions of online worked examples may also a) advance our understanding of how worked examples impact student thinking and b) inform future iterations of worked examples that optimize support for student learning in online settings.

Previously, we tested the effectiveness of six worked example formats that varied in presentation but matched in content and found that grade-school Algebra I students improved in their ability to simplify equations after completing instructional practice with any of the six worked example formats (Smith et al., 2022). Since all of the worked example formats led to comparable learning gains among Algebra I students, we now aim to understand how different features of worked examples affect students beyond learning gains alone. In this study, we assess college students' perceptions of the six worked example formats on simplifying equations to gather student feedback that might inform future designs of worked examples for online math learning environments. Specifically, we ask: 1) *Which worked examples are rated as the most vs.*

least helpful? 2) *What themes emerge in students' explanations for their rating?* 3) *How do students' explanations for the most helpful and least helpful worked example presentations provide further insights into the features to which students attend?*

Theoretical Framework

Frameworks of self-regulated learning (SRL; e.g., Bjork et al., 2013; Dunlosky & Ariel, 2011; Nelson & Narens, 1990) broadly model learning as a cycle between students monitoring and controlling learning tactics based on self-assessments of their progress in developing content knowledge or skills. Recent evidence showed that students were able to self-regulate learning by deciding if and when to use worked examples or practice problem solving (Foster et al., 2018). However, students tended to underutilize worked examples, and were more likely to study worked examples *after* problem solving. These student behaviors contrast with evidence that studying worked examples is more effective *prior to* problem solving (Leppink et al., 2014; Van Gog et al, 2011). Although students are able to self-regulate learning, Foster and colleagues (2018) call for more research on how students' self-regulation decisions are related to learning outcomes. More broadly, we posit that students' perception of worked examples may influence the ways in which they regulate learning and impact learning outcomes.

Here, we compare students' reactions to six different formats of worked examples varying in their extensiveness and degree of dynamicness based on cognitive load theory (Sweller, 2006) and perceptual learning theory (Gibson, 1969; Goldstone et al., 2017), respectively. We reason that there are two primary competing hypotheses as to which worked example format students will perceive as most helpful. First, based on cognitive load theory, students may prefer the *concise static* worked example as it provides only the major derivation steps to a problem without splitting their attention across multiple sources of information in worked examples. Second, based on perceptual learning theory, students may prefer the *dynamic* worked examples that show fluid transformations between each derivation step. Aligned with perceptual learning theory, the additional visual cues may help direct students' attention towards relevant actions for problem solving and raise awareness of perceptual cues to aid with problem solving. However, it is unclear how much detail and animation may be helpful, prompting us to also explore the effect of studying *dynamic* compared to *static* and *sequential* (an intermediate version that presents each equation line as a new animation) worked examples that vary in their amount of detail. As a first step towards understanding students' perception of the worked examples, we asked college students to rate the helpfulness of the worked examples as well as to explain their ratings.

Methods

A total of 109 students from a private university in the Northeastern U.S. participated for partial course credit. The sample afforded 80% power to detect the effect of $f > 0.35$ at $p < .05$.

Students completed a 30-minute online, within-subjects study designed as an assignment in ASSISTments, an online homework and research platform (Heffernan & Heffernan, 2014). Within the assignment, students completed six pairs of worked examples and practice problems, with each pair followed by two survey items, in a randomized order. For each worked example, students were instructed to study the worked example then enter the solution as an answer. On the following page, students completed a practice problem that matched the equation structure of the worked example without any instructional support or feedback. Immediately following each pair, students were asked how much they agree with the following statement: "*The worked examples were helpful for learning how to solve equations*". Students rated their perceived helpfulness on a 6-point Likert scale (1=Strongly Disagree; 6=Strongly Agree). They were then prompted to explain their rating in an open-response textbox that appeared below on the same screen which said, “Please use the open response to explain your answer.”

We designed six worked example presentations that varied in their visual features but not in content. Specifically, we manipulated the degree of *dynamicness* (static, sequential, or dynamic) and *extensiveness* (concise or extended) of each worked example. We adapted six worked examples that were designed for seventh-graders from Rittle-Johnson and Star (2007), so they were similar in content and difficulty. Previously, we also used these worked example presentations to investigate their differential effects on learning among Algebra I students; we found that students did improve from pretest to posttest on simplifying equations after completing instructional practice with any of the six worked example presentations (Smith et al., 2022). In this study, students viewed all six of the following formats in a randomized order.

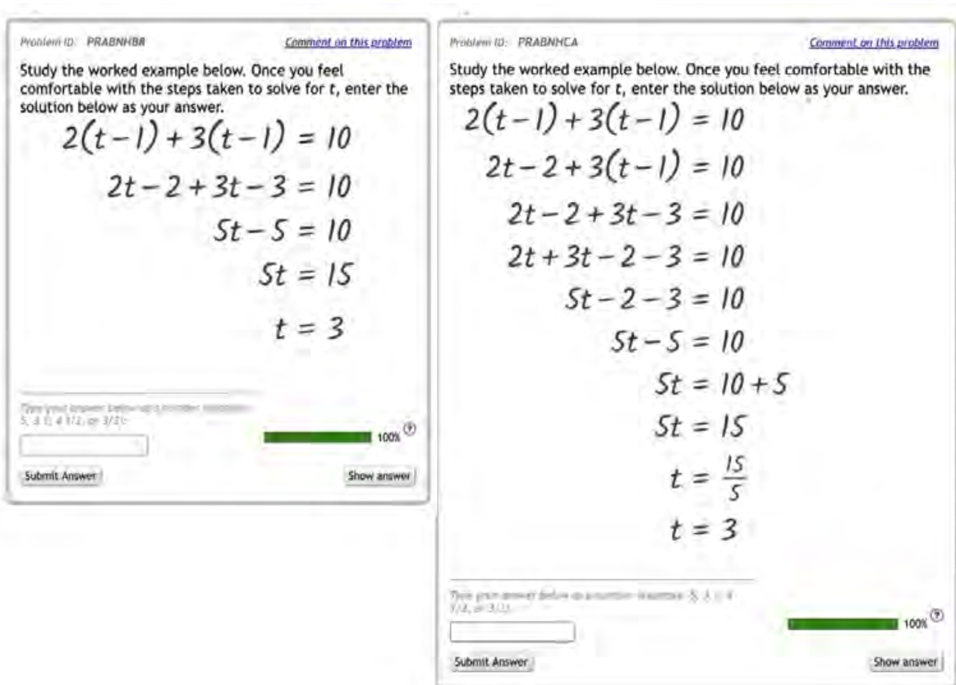

Figure 1: Static Concise Presentation (Left) and Static Extended Presentation (Right)

The first four conditions were based on a 2×2 design, differing in (a) length (i.e., concise or extended), and (b) the dynamicness of the worked example, (i.e., static or sequential). The *concise* worked examples displayed only the major steps in the derivation while the *extended* worked examples showed every step in the derivation. Static presentations mirrored the most commonly used design of worked examples in math (e.g., Rittle-Johnson & Star, 2007), presenting each worked example as images (Figure 1). The sequential worked examples were displayed as a looping GIF video that presented the static worked examples line-by-line in approximately 3-second intervals, creating a step-by-step history of the derivation over time. By varying the length and dynamicness, we created four different worked examples: *static concise*, *static extended*, *sequential concise*, and *sequential extended*.

Additionally, we created two presentations which displayed dynamic transformations of the expression and varied in whether the history of the derivation was displayed or not. The *dynamic history* worked examples showed looping videos of the transformation process through a screen recording. For example, to transform $2(t-1)+3(t-1)=10$ to $2t-2+3(t-1)=10$, students watched as the 2 was dragged over the parentheses to enact the transformation and record the result of the action on a new line (Figure 2). Each step of the problem was shown on the following line and a history of the steps taken was displayed sequentially. The transformations in the *dynamic no history* presentation were identical to those in the *dynamic history* presentation, but all occurred on one line of the equation without creating a derivation history.

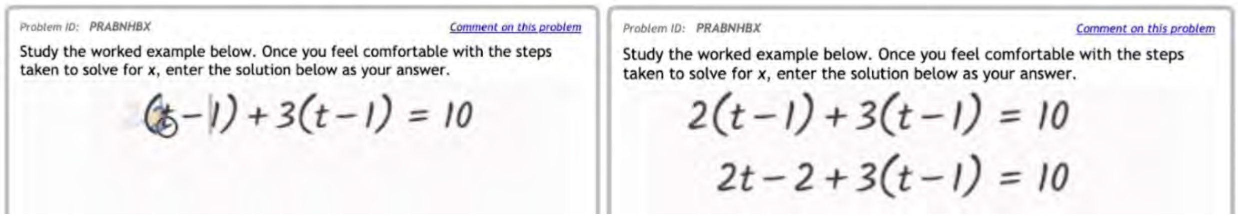

Figure 2: The *Dynamic History* Presentation Shows the Process of Each Transformation

Approach to Analysis and Results

Research Question 1: Perceived Helpfulness Ratings of Worked Example Presentations

To first compare how students rated the different worked example formats, we treated students' ratings (on a 6-point scale) as a continuous variable (Robitzsch, 2020). A one-way ANOVA revealed a significant effect of worked example presentation on students' rating, $F(5, 648) = 5.76$, $p < .001$. Post hoc comparisons using Bonferroni correction indicated that the mean rating for the *static concise* presentation ($M = 5.17$, $SD = 1.21$) was significantly higher than the *dynamic no history* ($M = 4.22$, $SD = 1.56$), *sequential extended* ($M = 4.59$, $SD = 1.44$), and *dynamic history* ($M = 4.60$, $SD = 1.53$) presentations, $ps < .05$. Further, the mean rating for the *static extended* presentation ($M = 4.92$, $SD = 1.31$) was significantly higher than the *dynamic no history* condition ($M = 4.22$, $SD = 1.56$), $p < .01$ (Figure 3). There were no significant differences in students' rating between the *sequential concise* presentation ($M = 4.61$, $SD = 1.38$) and any of the other presentations.

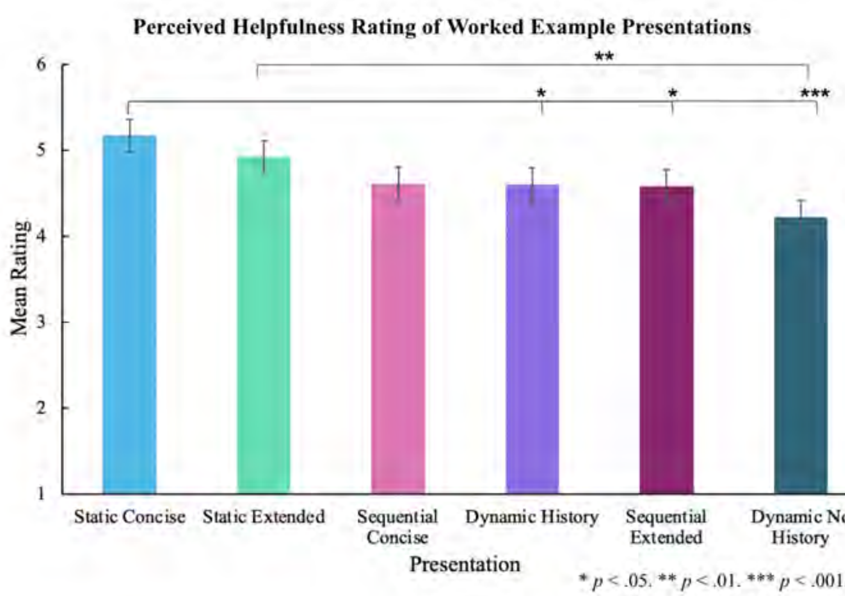

Figure 3: Average Perceived Helpfulness Rating of Each Worked Example Presentation

Research Question 2: Themes Among Students' Explanations

Next, we identified common themes in students' explanations of why they found each worked example helpful or unhelpful. Two researchers independently coded the 654 explanations with 79% initial agreement and then resolved discrepancies together to create the final codes. The following themes emerged: amount of content, video speed, dynamic coherence, and visual features (Table 1). Importantly, students' explanations were coded as whether they included a reference to one or more of the themes rather than whether students found each feature to be a positive or negative feature of the worked example presentation.

Table 1: Themes Identified Across Students' Explanations

Theme	Definition	Student Example
Amount of Content	Comments on the quantity of information presented in the worked example	*"I like seeing each step in the problem (regardless of how "trivial" the step is). It helps me see the train of thought that I'm supposed to be having while solving the problem."*
Dynamic Coherence	Comments on the presence or absence of fluid transformations and transitions that connect steps and actions.	*"The transitions may have helped but made it more visually confusing"*
Speed	Comments on the speed of the animations	*"It was very slow and hard to focus on"*
Visual Features	Comments on visual features such as font, spacing, styling, and format	*"…The font is very clear…"*

The themes identified were not exhaustive or mutually exclusive: students' explanations could be coded as pertaining to none, one, or more than one theme. As a result, 303 of the 654 explanations were not labeled as matching any of the above themes. For example, one student explained the problem-solving process rather than describing why they found the example to be helpful or unhelpful: "*By using the distributive property, the 3 was multiplied to the h and the −2, and the 5 was multiplied to the h and the −2. Like terms were added together, and h was isolated by simple algebra*." Conversely, another student's reaction to the *sequential concise* worked example referenced amount of content, video speed, and dynamic coherence:

> *While this animation laid out all the necessary steps to solve the equation, the steps were revealed somewhat spastically and it was almost startling to the eye to watch the equation being solved. Additionally, this animation left the viewer with little time to slowly work through the problem at their own pace and go back to a previous step if they were confused. This is because the format of this animation was a bit stressful.*

Table 2 presents the frequency of explanations that were labeled as containing one or more of the described themes by condition and overall. Students commented most often on the *amount of content* presented followed by the *dynamic coherence* of the worked examples and *speed* of the worked example. Miscellaneous visual features such as font and format were less often noted.

Table 2: Frequency of Themes Across Conditions

Worked Example Condition	Theme			
	Amount of Content	Dynamic Coherence	Speed	Visual Features
Static Concise	15	20	18	6
Static Extended	30	11	5	6
Sequential Concise	33	17	30	9
Dynamic History	16	36	12	9
Sequential Extended	16	24	19	6
Dynamic No History	31	31	21	7
Total	141	139	105	43

Research Question 3: Students' Explanations for Most and Least Helpful Presentations

Finally, to delve deeper into how students attend to features of worked examples and to contribute findings to cognitive load and perceptual learning theories, we explored how students' reactions to the *static concise* and *dynamic no history* presentations may explain their highest and lowest helpfulness ratings, respectively (Table 3). A closer inspection revealed that students found the *static concise* worked examples familiar and helpful because students could self-pace how they studied them. Twelve students praised the opportunity to self-pace themselves; further, four students noted that they would have preferred to also have instructional explanations accompany the *static concise* worked example. On the other hand, nine students positively noted that the *dynamic no history* presentations were helpful by providing explicit transitions between derivation steps while 38 students noted that the format caused confusion, and the derivation disappearing as each new line appeared required more effort and time to process.

Table 3: Sample Explanations for Static Concise and Dynamic No History Presentations

Static Concise	Dynamic No History
"the worked example was very helpful because steps were concise and understandable and the numbers didn't move around or disappear."	"Having all of the steps on the same line during the animation can be cumbersome if a student trying to learn the process has a specific question about a particular set and has to wait for the full animation to finish for it to restart."
"The example showed the steps, so I could understand the process. However, if I did not know how to do it already I may have been confused because there were no	"This worked example really helped as it would physically move each number around so you would physically see what was happening and then show what it would produced so you would

written instructions of what the person did.”	know where each number was going. This helped with visualizing what was taking place.”
“With few steps, this is not too hard to follow. But with larger numbers or more written steps, this would easily become overwhelming!”	“I think it would have been more helpful to see the steps written out statically rather than in a gif. It could be difficult to keep up, especially if you didn't understand a step right away.”
“I could look at the whole problem at my own pace.”	“Was way harder to follow what was happening without everything written out; once the directions disappeared I couldn't follow the solution anymore”

Discussion

This study is a foundational effort to explore how students respond to worked examples in ways that may not be reflected through learning gains alone. To integrate students’ perspectives on worked example designs for algebra, we asked college students to rate the helpfulness of six worked example presentations and to explain their ratings. We found that students rated the *static concise* presentation as the most helpful and the *dynamic no history* presentation as the least helpful. Students’ explanations for their ratings revealed that they commonly attended to the following elements of the worked examples: *amount of content*, *dynamic coherence*, *speed*, and other *visual features*. On the one hand, students revealed that the *static concise* presentation was familiar to previous classroom experiences and allowed for self-pacing, potentially supporting students’ chances to self-regulate their learning. On the other hand, the *dynamic no history* presentation required more effort to process, rendering it less helpful. These findings advance cognitive theories of mathematics learning and provide implications for researchers and teachers.

Our finding that students rated the *static concise* presentation as the most helpful aligns with cognitive load theory and explanations for the worked example effect. Specifically, students preferred the presentation with the least amount of information and content which likely prevented students from feeling as though they were studying redundant information or splitting their attention between different areas of the worked example (Sweller, 2020). This presentation is also a commonly used format for algebra worked examples (e.g., Rittle-Johnson & Star, 2007) and multiple students commented that they found it familiar with their past experiences.

Further, the finding that students found the *dynamic no history* presentation the least helpful aligned with prior work comparing different formats of worked examples. In particular, Lusk and Atkinson (2007) exposed college students to worked examples on proportional word problems that were: *fully-embodied* with an animated parrot that used gesture and gaze, *minimally embodied* with a static parrot who could only talk, or *voice-only* worked examples that provided only an audio description for the worked example. They found that while learning gains were the highest in the *fully-embodied* condition, students who studied the *minimally embodied* worked examples had the lowest levels of cognitive load. Lusk and Atkinson’s (2007) results suggest that extra information presented through animation and dynamic videos may support learning but may also present more challenges to students as they study worked examples with animations. Although ample research has demonstrated the benefits of perceptual scaffolding in math to direct students’ attention towards important cues in notation (e.g., Goldstone et al., 2017), perhaps the fluid transformations displayed in the *dynamic no history* worked example presentation provide extraneous distractions that increase demands on, rather than offload, students' working memory.

Students' explanations for the static concise and dynamic no history presentations indicated important visual features to consider when modifying these worked example presentations in the future. Namely, students indicated that they would find the worked examples more helpful if accompanied by written explanations for each step in the derivation. Further, if researchers or instructors choose to use a dynamic version of worked examples in classroom instruction or online practice, slowing the video speed or providing students with the autonomy to pause videos may reduce the cognitive effort and time that students indicated needing to process the *dynamic no history* presentations.

Importantly, we acknowledge that students' ratings and explanations for the helpfulness of the worked examples may be influenced by their own content knowledge. Specifically, the study materials were developed based on middle-school math content and might have been too easy for our sample of college students. However, by conducting this study with college students, we received thoughtful reactions to the worked example formats that still provide insights that may inform future iterations of worked examples used for online instructional practice.

Looking ahead, conducting comprehensive studies with algebra students may help delineate cognitive and non-cognitive mechanisms of learning and inform the design of worked examples for online learning environments. For instance, the six worked example presentations here were modeled after principles of cognitive load theory and perceptual learning theory; future studies should measure how students' cognitive load is impacted and how students may be impacted by studying different worked example presentations at different stages of learning. Further, it may be worthwhile to investigate the relation between students' learning gains, cognitive load, and perceptions of helpfulness. For instance, Whitehill et al. (2019) found that learners' subjective ratings of various tutorial videos (i.e., how helpful the videos would be for others) were highly correlated with their later learning gains. Participants who found the videos helpful learned more than those who did not find the videos helpful, confirming their initial bias towards the videos. These findings suggest that learners might be able to gauge the effectiveness of instructional support or that learning may be modulated by learners' perceived helpfulness. Regardless of the directionality of the influences, students' perceived helpfulness of the instructional materials may be closely related to learning outcomes, warranting further investigations of how this effect may impact the effectiveness of worked examples.

Analyzing students' perceptions of six worked example formats revealed that students prefer worked examples with minimal instructional content in a static image rather than worked examples with animations and fluid transformations between steps. These findings provide insights for researchers and teachers as they design algebra worked examples for online settings.

Acknowledgments

We thank Nellie Rushton and Lilly-Beth Linnell for coding students' responses. Additionally, the research and development reported here was 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 Institute or the U.S. Department of Education.

References

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444.

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.

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

Dunlosky, J., & Ariel, R. (2011). Self-regulated learning and the allocation of study time. B. Ross (Ed.). *Psychology of Learning and Motivation, 54*, 103–140.

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.

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.

Hartmann, C., Van Gog, T., & Rummel, N. (2020). Do examples of failure effectively prepare students for learning from subsequent instruction? *Applied Cognitive Psychology, 34*, 879–889. https://doi.org/10.1002/acp.3651

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building 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.

Leppink, J., Paas, F., Van Gog, T., Van Der Vleuten, C. P. M., & Van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction, 30*, 32–42.

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: The Official Journal of the Society for Applied Research in Memory and Cognition, 21*(6), 747-764.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe, & A. P. Shimamura (Eds.). *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

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.

Robitzsch, A. (2020, October). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. In *Frontiers in Education* (Vol. 5, p. 177). Frontiers.

Scheiter, K. (2020). Embracing complexity in research on learning from examples and from problem solving. *Applied Cognitive Psychology*, *34*(4), 906-911.

Smith, H., Closser, A. H., Ottmar, E. R., & Chan, J. Y. C. (2022). The impact of algebra worked example presentations on student learning. *Applied Cognitive Psychology, 36*(2), 363-377.

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.

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*, 890–905. https://doi.org/10.1002/acp.3652

van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology, 36*, 212–218.

Whitehill, J., Aguerrebere, C., & Hylak, B. (2019). Do Learners Know What's Good for Them? Crowdsourcing Subjective Ratings of OERs to Predict Learning Gains. *International Educational Data Mining Society*.