

# Promoting Students' Self-Regulated Choices in Learning with Visual Representations

**Tomohiro Nagashima**

CMU-HCII-22-102

August 2022

Human-Computer Interaction Institute  
School of Computer Science  
Carnegie Mellon University  
Pittsburgh, Pennsylvania 15213

**Thesis Committee:**

Vincent Aleven, Chair

Geoff Kaufman

Amy Ogan

Martha W. Alibali (University of Wisconsin-Madison)

Timothy Nokes-Malach (University of Pittsburgh)

*Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy*

Copyright © 2022 Tomohiro Nagashima. All rights reserved.

This work is supported in part by National Science Foundation grant (#1760922) and a Presidential Fellowship by the School of Computer Science at Carnegie Mellon University. All opinions, findings, conclusions, or recommendations expressed in this material are those of the author and do not necessarily reflect those of the funding agencies.

## Acknowledgements

I am grateful for all the support that I have been given by my colleagues and friends. My PhD would have been undoubtedly impossible without the support I have received during my time in Pittsburgh and even before that. Thank you very much, Vincent Aleven, for your amazing mentorship and a countless number of our intellectual discussions. Thank you, Martha Alibali, for helping me see education research from a different perspective. Thank you very much, my dissertation committee, which includes Geoff Kaufman, Amy Ogan, and Timothy Nokes-Malach, for their insightful comments, encouragement, and support. I also appreciate my first-year PhD advisor, John Stamper, for letting me freely pursue my own interest. Thank you also to other faculty members at CMU who have helped me grow and learn in various ways: Zach Branson, Motahhare Eslami, Ken Holstein, Ken Koedinger, Bruce McLaren, Raelin Musuraca, Adam Perer, Stacie Rohrbach, and Franceska Xhakaj.

I appreciate the support I have received from other CMU colleagues. Thank you, Jeff Bigham and Queenie Kravitz, for making my PhD at HCII a pleasant experience. Thanks to the members of PIER for including me in the wonderful network of education researchers. Thank you, Diana Rotondo and Jo Bodnar, for their daily support. Thank you, Gail Kubit and Susan Bruner for their support through IRB. And of course, big thanks to Max Benson, Octav Popescu, Jonathan Sewall, and Cindy Tipper! No way I could have done my studies without your dedicated support. I so much appreciate it.

I am lucky that I have been supported by many friends. Thank you so much, Anna Bartel, Elena Silla, and Nicholas Vest, for having been my best friends through our wonderful CMU-UWMadison collaboration. I learned a lot from you. Thank you, Vanessa Echeverria, Yun Huang, Shamyia Chodumada Karumbaiah, LuEttaMae Lawrence, Naomie Williams, Meng Xia, and Kexin Yang, for having been my great lab mates. Thank you so much to fellow PhD students, in particular, Julia Cambre, Vikram Kamath, Lynn Kirabo, Samantha Reig, Stephanie Valencia Valencia, and Nur Yildirim for sharing many great experiences together throughout the journey. Also, thank you my awesome research assistants: Marcus Artigue, John Britti, Dreami Chambers, Jeff Chen, Evan Fang, Xinying Hou, Hailey Jeong, Ruitao Li, Elizabeth Ling, Cindy Liu, Jordan Love, Michelle Ma, Xiaoying Meng, Trula Rael, Yuling Sun, Stephanie Tseng, Xiran Wang, Sihan Wu, Gautam Yadav, Junhui Yao, Alan Zhao and Bin Zheng – it has been wonderful working with you all.

I appreciate the support from the community outside CMU that helped me and my family during my PhD. Thank you Kristen Carey, Aksel Casson, Shivaun Curry, Jesus Silva Elizalde, Leah Jacobs, and Tetsuya Yamada for having been the nicest friends to my family. Thank you, Jenny Yun-Chen Chan, for all the exciting conversations and encouragements. Thank you, Steffi Bruninghaus, for many fun chats. Thank you, Karin Forssell, Roy Pea, and Candace Thille for supporting me during my first year in US at Stanford. Also, thank you very much, Katsusuke Shigeta, Toru Iiyoshi, and Tomoaki Watanabe for your continuous support for me and my family from Japan. Thank you very much, teachers, students, and their parents, who kindly helped me do my research.

And of course, Rumi, Aoi, and Sui – thank you for everything we have been through together. My PhD journey has truly been a joint accomplishment with my family, especially with Rumi.

## Abstract

An ultimate goal of education is to help learners become self-regulated, autonomous learners who can continue to learn beyond formal school education by strategically choosing to use available resources around them. My dissertation approaches this grand question with a specific focus on the topics of choosing to use visual representations strategically and effectively learning with them during algebra problem solving. Visual representations are a type of instructional aids that help learners' sense-making processes during problem solving and learning. I have used classroom experiments, user-centered design, and educational data mining approaches to investigate and support students' learning with visual representations and their learning of strategic use of visual representations.

In my dissertation, I studied two key questions: 1) *how might we support students' effective and efficient algebra learning with visual representations?* and 2) *how might we support students' self-regulated, strategic use of visual representations during algebra learning?* In earlier work, I conducted five studies involving user-centered research and classroom experiments to design and establish novel interactive visual scaffolding called "diagrammatic self-explanation" in the context of intelligent tutoring software for algebra learning. In my later work, I conducted three design and experimental studies to design and evaluate an adaptive metacognitive intervention that supports students' self-regulated use of visual scaffold in the intelligent tutoring software in classroom.

My work with about 20 middle-school teachers and 500 students in the U.S. provides several new contributions in the field of the learning sciences. Among others, my dissertation shows that novel, interactive diagrammatic self-explanation activities for algebra designed with teachers supported effective and efficient learning. My work also shows that a metacognitive intervention designed with middle-school students facilitated students' strategic choices in using visual scaffolding in an interactive learning environment and enhanced both conceptual and procedural learning in early algebra, a challenging dual goal in the field.

### **Keywords**

Algebra, Instructional design, Intelligent Tutoring Systems, Metacognition, Self-explanation, Self-regulated learning, Student choice, Tape diagrams, and Visual representations

## TABLE OF CONTENTS

<b>CHAPTER 1: INTRODUCTION .....</b>	<b>6</b>
<b>CHAPTER 2: THEORETICAL BACKGROUND .....</b>	<b>14</b>
2.1: <i>Learning with Visual Representations .....</i>	<i>14</i>
2.2: <i>Scaffolding Learning with Visual Representations by Leveraging Pedagogical Affordances .....</i>	<i>17</i>
2.3: <i>Learning to Self-Regulate the Use of Visual Representations .....</i>	<i>22</i>
2.4: <i>Promoting Strategic Use of Visual Representations .....</i>	<i>23</i>
<b>CHAPTER 3: LEARNING WITH VISUAL REPRESENTATIONS .....</b>	<b>32</b>
3.1: <i>Study 1: Pedagogical Affordance Analysis .....</i>	<i>32</i>
3.2: <i>Study 2: Supporting Conceptual Understanding through Confirmatory Diagrammatic Self-Explanation .....</i>	<i>40</i>
3.3: <i>Study 3: Supporting Learning and Performance through Anticipatory Diagrammatic Self-Explanation .....</i>	<i>53</i>
3.4: <i>Study 4: Testing the Effectiveness of Anticipatory Diagrammatic Self-Explanation on Students with Different Prior Knowledge Levels .....</i>	<i>60</i>
3.5: <i>Study 5: Reducing Scaffolding in Anticipatory Diagrammatic Self-Explanation through Interleaved Practice .....</i>	<i>67</i>
<b>CHAPTER 4: SUPPORTING SELF-REGULATED USE OF VISUAL REPRESENTATIONS .....</b>	<b>81</b>
4.1: <i>Motivation .....</i>	<i>81</i>
4.2: <i>Study 6: Choice-based Anticipatory Diagrammatic Self-Explanation .....</i>	<i>82</i>
4.3: <i>Study 7: Generating Ideas for Metacognitive Interventions with Children .....</i>	<i>90</i>
4.4: <i>Study 8: Supporting Self-Regulated Use of Diagrams and Math Learning .....</i>	<i>98</i>
<b>CHAPTER 5: CONCLUSION .....</b>	<b>118</b>
5.1: <i>Concluding Thoughts .....</i>	<i>118</i>
5.2: <i>Limitations .....</i>	<i>120</i>
5.3: <i>Contributions .....</i>	<i>120</i>
<i>References .....</i>	<i>129</i>
<i>Appendix 1: Tests and surveys used in experiments .....</i>	<i>148</i>
<i>Appendix 2: Permission letter from the International Society of the Learning Sciences .....</i>	<i>176</i>

## 1 Chapter 1: Introduction

An ultimate goal of education is to help learners become self-regulated, autonomous learners who can continue to learn by themselves beyond formal school education (Chin et al., 2019; Cutumisu et al., 2015; Schwartz & Arena, 2013). In today's world where one's surrounding environment is changing rapidly, one cannot follow a designed career path. Rather, life in the 21<sup>st</sup>-century is like a "kayaker navigating white water"; learners need to read the currents, assess benefits and risks, choose to use tools and resources strategically, and exercise skills to quickly adapt to changes in society (Pendleton-Jullian & Brown, 2018). This rapidly-changing, unknown future makes it more important than ever to design and study instruction that helps learners become self-regulated and autonomous. *How might we foster learners of the 21<sup>st</sup> century who can autonomously learn by effectively and efficiently using available resources?*

My dissertation addresses this grand question with a focus on the specific topics of strategically choosing to use visual representations and learning with them during math problem solving. Visual representations are a type of instructional aid that help learners' sense-making processes during problem solving and learning (Rau, 2017). By visualizing information that would be difficult to convey with textual or verbal information, visual representations help learners notice and understand important features of the target content (e.g., scientific concepts) (Larkin & Simon, 1987; Rau, 2017). Choosing to use visual representations (when it is optional to use visual representations) is an important self-regulated learning skill (Clarebout & Elen, 2006; Zimmerman, 2008; Zimmerman & Pons, 1986); using visual representations is a strategy that learners can choose to adopt or not (e.g., learners might not use visual representations for the sake of faster problem solving) but using them can effectively aid problem solving and learning when learners actively engage with visuals in a meaningful way (Booth & Koedinger, 2010; Uesaka et al., 2007). Past studies have examined only one aspect of visual use (i.e., if students spontaneously use diagrams, Uesaka & Manalo, 2012). My dissertation research uses classroom experiments and fine-grained data from a tutoring system to investigate students' detailed choice behaviors involving the use of visual representations (e.g., "when and how do students choose to and not to use visuals?"). This is an important yet understudied area of self-regulated use of visual representations.

If choices matter in learning, researchers need to appropriately measure learner choices (Cutumisu et al., 2015; Schwartz & Arena, 2013). Advanced learning technologies such as Intelligent Tutoring Systems (ITSs) provide an opportunity to design a wide range of learning environments and to assess learner choices appropriately (Kulik & Fletcher, 2016). Research with ITSs to date has largely focused on supporting learners' domain-level knowledge and skills (e.g., mathematics, Alevan &

Koedinger, 2002) but recent efforts have also shown the promise and evidence of intelligent systems in supporting self-regulated learning skills, including the skill of choosing to seek help (Roll et al., 2011), choosing to ask for feedback (Tan & Biswas, 2006), choosing to use self-regulated learning strategies during problem solving (Azevedo et al., 2009), and choices involving problem selections in an intelligent system (Long & Alevan, 2017). Among many advantages of ITSs, a particular advantage in the context of choice-based learning is that ITSs provide fine-grained, step-by-step learning and performance data that allow researchers to understand students' learning processes and choice behaviors in the designed learning environment, which is critical in assessing students' learning processes (Ben-Eliyahu & Bernacki, 2015, Roll & Winne, 2015). My work leverages fine-grained ITS log data to uncover students' learning processes and patterns involving the use of visual scaffolding, aiming to contribute to better understanding of students' self-regulated use of visual representations in problem solving.

My entire work in the dissertation, consisting of eight studies, uses various methodological approaches to investigate the following open questions in the field of the learning sciences:

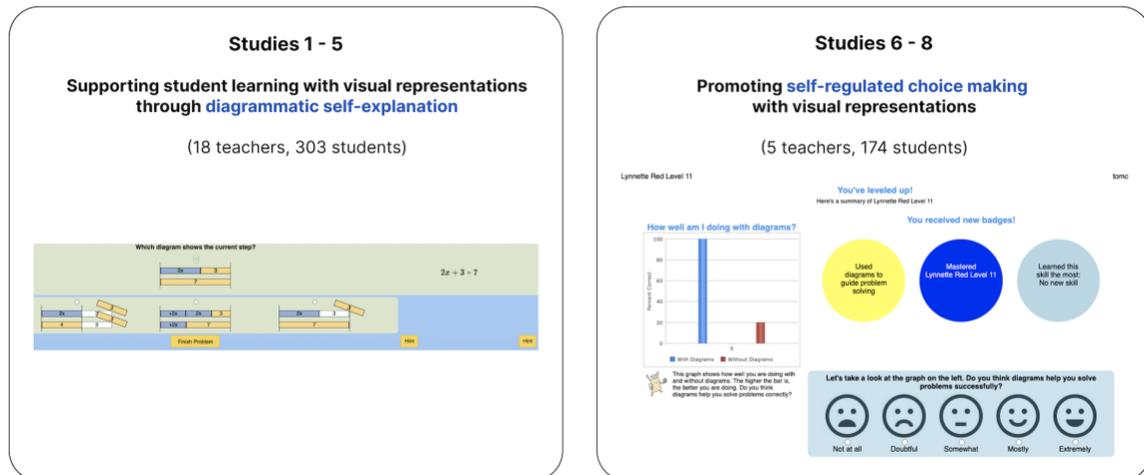
1. How might we design a learning environment to support students' algebra learning and problem-solving processes with visual representations (Studies 1-5)? and
2. How might we design a learning environment to support students' self-regulated use of visual representations (Studies 6-8)?

Exploring these open questions makes important theoretical and practical contributions: theoretically, my series of studies advances scientific understanding of how students learn with visual representations and how to support their learning and performance using various forms of an instructional scaffold called "diagrammatic self-explanation." It also helps understand how to support students' effective strategies involving the use of visual representations. Practically, my work offers knowledge on how researchers and designers can use a user-centered approach to design instructional tools, and how educators and designers can structure their learning activities using visual representations to promote effective, efficient, and meaningful learning.

Throughout my eight studies, I have closely worked with more than 20 K-12 math teachers at 15 schools and about 500 students in researching how visual representations as a form of scaffolding could best be leveraged to support learners. Specifically, I designed and evaluated a novel form of scaffolding student learning with visual representations called diagrammatic self-explanation. I use

user-centered design, classroom experiments, and data mining techniques to uncover the complex nature of domain-specific learning (i.e., use visual representations to learn math concepts and skills), self-regulated learning (i.e., choose to and not to use visual representations in a meaningful way), and the relations between the two. My dissertation work can be categorized into two bigger topics, illustrated in Figure 1. To be more specific, I have:

- **Designed and tested, in a series of classroom and design studies, different ways of using visual scaffolding (called “diagrammatic self-explanation”) to support student learning and problem-solving performance in early algebra** (Studies 1-5, with 303 students, see also Table 1)
- **Designed and tested, in a series of classroom and design studies, metacognitive support to help students strategically choose to and not to use visual scaffolding in algebra learning** (Studies 6-8, with 203 students, see also Table 2)



**Fig. 1.** My thesis studied two topics involving the use of visual representations in algebra learning.

My studies make the following specific contributions to the broad field of the learning sciences (Tables 1 and 2). Tables 3 and 4 provide detailed descriptions of my studies and what research questions were addressed in each study.

**Table 1.** Major contributions of my studies (Studies 1-5).

Supporting student learning with visual representations in algebra (Studies 1-5)		
Contribution area	What was known already	What my research adds
Instructional design	Many instructional design models are used to prescribe what to do to design instruction (designers' "over-reliance" on models).	Provides a user-centered, systematic method ("Pedagogical Affordance Analysis") to identify pedagogical affordances and constraints in designing instructional tools by working with educators
Self-explanation	It is challenging to design a self-explanation activity that effectively and efficiently supports both student learning and performance during learning.	Designed "confirmatory diagrammatic self-explanation" and "anticipatory diagrammatic self-explanation", two novel interactive self-explanation strategies ( <a href="#">URL</a> ) Classroom studies showed that these strategies support both learning and problem-solving performance.
Mathematical cognition	Tape diagrams, a popular visual representation in algebra instruction were considered a useful tool to support student learning but few studies have achieved their effect on student learning.	Designed and tested visual scaffold using tape diagrams (confirmatory and anticipatory diagrammatic self-explanation) Classroom studies show their benefits on student learning of algebra concepts and strategies.
Educational practice	Despite an increasing interest among educators in using tape diagrams, few useful resources existed.	Designed a digital template of various types of tape diagrams and distributed it as an Open Educational Resource ( <a href="#">URL</a> )

**Table 2.** Major contributions of my studies (Studies 6-8).

Promoting self-regulated, strategic choices in using visual representations in algebra (Studies 6-8)		
Contribution area	What was known already	What my research adds
Diagram research	Studies have mostly looked at students' spontaneous diagram use when diagram use is investigated	Using fine-grained temporal data from an intelligent tutor, I found that students' use patterns and how they change over time are also important in studying diagram use during learning
Self-regulated learning	Interventions aimed at supporting self-regulated learning have not always been successful at improving domain-level knowledge and skills	Designed a metacognitive intervention with students, which helped students learn better conceptual and procedural knowledge than students without the metacognitive intervention
Behavior change research	Many behavior change models assume straightforward target behaviors, reducing its applicability to complex human activities	Developed the Metacognitive Choice Behavior Model, which can be applied to behavior adoptions that involve complicated choice-making behaviors and processes
Mathematical cognition	In algebra instruction, few studies have shown instructional strategies that support both conceptual and procedural learning	Designed a metacognitive intervention that helped students learn both conceptual and procedural knowledge in algebra
Intelligent Tutoring Systems (ITSs)	There are not many ITSs that are designed with users (educators and learners), especially those aimed at improving metacognition and self-regulated learning	Designed features of an algebra ITS to support self-regulated learning with users (middle-school students), which helped students learn effectively in the classroom environment
Choice making in learning	Studies on students' choice behaviors typically examine aggregated choice measures (e.g., sum of the number of certain behaviors), making it hard to understand how choice behaviors change over time	Investigated students' fine-grained choice behaviors in an algebra intelligent tutoring system and found that students' strategic choice behaviors seemed to affect effective learning, which would not have been observed with aggregated data only

**Table 3.** Studies included in this dissertation (Studies 1-5).

Supporting student learning with visual representations in algebra (Studies 1-5)				
Study #	Date	Research Question(s)	Participants	Outcomes/ Publications
Study 1	Fall 2018 - Spring 2019	“How might we re-design an existing visual representation (tape diagrams) to help struggling middle-school students learn conceptual knowledge in algebra?”	8 middle-school teachers	Nagashima, Yang et al., 2020
Study 1	Spring 2019	“How do tape diagrams help learners think conceptually in early algebra?”	7 middle-school students	Nagashima, Yang et al., 2020
Study 2	Spring 2019	“Will the use of tape diagrams in early algebra through a scaffolded sense-making activity called “diagrammatic self-explanation” help students learn conceptual knowledge	41 middle-school students, 2 teachers, 2 classes, 1 school	Nagashima, Bartel et al., 2020
Study 3	Spring 2020	“How might we support both learning and performance in early algebra using tape diagrams	108 middle-school students, 5 teachers, 9 classes, 2 schools	Nagashima, Bartel, Yadav et al., 2021
Study 4	Spring 2020	“Will scaffolded use of tape diagrams in an Intelligent Tutoring System help student learning and performance regardless of their prior knowledge?” “How do learner interactions with visual representations influence their symbolic equation solving in an Intelligent Tutoring System?”	84 middle-school and elementary-school students, 2 teachers, 3 classes, 1 school	Nagashima, Bartel, Tseng et al., 2021
Study 5	Spring 2021	“Will interleaving scaffolded problems and un-scaffolded problems enhance better learning than giving scaffolded problems all the time?”	63 middle-school students, 1 teacher, 5 classes, 1 school	Nagashima, Ling et al., 2022

**Table 4.** Studies included in this dissertation (Studies 6-8).

Promoting self-regulated, strategic choices in using visual representations in algebra (Studies 6-8)				
Study #	Date	Research Questions	Participants	Outcomes/ Publications
Study 6	Fall 2021	<p>“How will students choose to use tape diagrams during problem solving in an Intelligent Tutoring System?”</p> <p>“Will students’ choices in using tape diagrams predict learning outcomes?”</p>	26 middle-school students, 2 teachers, 2 classes, 1 school	Nagashima, Tseng et al., 2022
Study 7	Spring 2022	“How might we design a metacognitive intervention that helps students think deeply and strategically about using tape diagrams during algebra problem solving?”	8 middle-school students	
Study 8	Spring 2022	“Will students who receive a metacognitive intervention on diagram use demonstrate strategic use of tape diagrams, better learning performance and learning outcomes?”	140 middle-school students, 2 teachers, 2 schools	

**Studies reported in this dissertation have been published in conferences (see below). For papers published at the International Conference of the Learning Sciences and at the Annual Meetings of the Learning Sciences, the International Society of the Learning Sciences owns the copyright and the permission for the reuse of published texts, tables, and figures has been obtained, only for the use in this document. For the papers published at the Annual Meetings of the Cognitive Science Society, the Cognitive Science Society holds the copyright and the use of published texts, tables, and figures is allowed under a Creative Commons Attribution 4.0 International License (CC-BY).**

**Nagashima, T.**, Yang, K., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2020). Pedagogical Affordance Analysis: Leveraging teachers' pedagogical knowledge for eliciting pedagogical affordances and constraints of instructional tools. In M. Gresalfi & I. S. Horn (Eds.), *14th International Conference of the Learning Sciences* (pp. 1561-1564). Nashville, TN: International Society of the Learning Sciences. [\[URL\]](#)

**Nagashima, T.**, Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2020). Enhancing conceptual knowledge in early algebra through scaffolding diagrammatic self-explanation. In M. Gresalfi & I. S. Horn (Eds.), *14th International Conference of the Learning Sciences* (pp. 35-43). Nashville, TN: International Society of the Learning Sciences. [\[URL\]](#)

**Nagashima, T.**, Bartel, A. N., Yadav, G., Tseng, S., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevan, V. (2021). Using anticipatory diagrammatic self-explanation to support learning and performance in early algebra. In E. de Vries, J. Ahn, & Y. Hod (Eds.), *15th International Conference of the Learning Sciences* (pp. 474-481). International Society of the Learning Sciences. [\[URL\]](#)

**Nagashima, T.**, Bartel, A. N., Tseng, S., Vest, N.A., Silla, E. M., Alibali, M. W., & Alevan, V. (2021). Scaffolded self-explanation with visual representations promotes efficient learning in early algebra. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society* (pp. 1858-1864). Cognitive Science Society. [\[URL\]](#)

**Nagashima, T.**, Ling, E., Zheng, B., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2022). How does sustaining and interleaving visual scaffolding help learners? A classroom study with an Intelligent Tutoring System. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*. Cognitive Science Society. [\[URL\]](#)

**Nagashima, T.**, Tseng, S., Ling, E., Bartel, A. N., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevan, V. (2022). Students' self-regulated use of diagrams in a choice-based intelligent tutoring system. In *Proceedings of the Annual Meeting for the International Society of the Learning Sciences*, Hiroshima, Japan [\[proceedings\]](#).

## **2 Chapter 2: Theoretical Background**

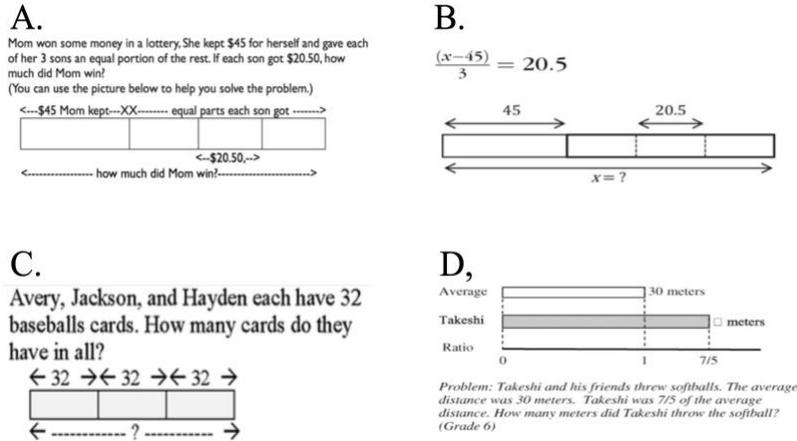
### **2.1 Learning with Visual Representations**

We use visual representations to navigate everyday problem-solving and decision-making situations (Lurie & Mason, 2007). For example, when driving, we perceptually recognize road signs and quickly take actions based on what the signs tell us. Visual representations are also critical in teaching and learning. As seen in many textbooks, visual representations play an important role in scaffolding learning of scientific and mathematical concepts in a variety of study domains, especially science, technology, engineering, and mathematics (STEM) domains (Ainsworth & Loizou, 2003; Fukuda et al., 2021; Rau, 2017).

Many types of visual representations are used in learning research and instructional practice. A shared principle among many types of visual representations is that visual representations make the “unseen” visible (Arcavi, 2003). By visualizing what would otherwise not be seen (or not be as easy to see), visual representations benefit human sense making (Arcavi, 2003; Larkin & Simon, 1987; Lurie & Mason, 2007). For instance, we use visual data representations (e.g., bar charts) to see patterns and get insights in data, which would be hard to quickly learn if there was no use of such data visualizations.

A number of studies have demonstrated the effect of using visual representations on human learning (e.g., Rau et al., 2015; Suyatna et al., 2017; Yun & Paas, 2015). Visual representations benefit problem solving and learning as they illustrate complex concepts and situations visible, accessible, and comprehensive (Larkin & Simon, 1987; Rau, 2017). However, it has also been found that the benefits of visual representations are not universal. Some prior studies have found detrimental or mixed effects of visual representations (e.g., Magner et al., 2014) and other studies have found that visual representations are not effective for all subgroups of students (e.g., Cooper et al., 2018; van Garderen & Montague, 2003). Many factors moderate the effects of learning with visual representations, including prior knowledge, ability, age, as well as how the learning with visual representations is supported (e.g., Booth & Koedinger, 2012). Thus, careful design of visual representations and the accompanying instruction may be essential to meaningfully support student interactions with visual representations, especially for younger, lower-achieving students, and those who have lower prior knowledge (Davenport et al., 2008). Because these groups of students are the ones who need the most instructional support to achieve their learning goals, it is important that visual representations, as instructional scaffolding, are designed to help these groups of students.

One of the many domains in which visual representations are frequently used is early algebra. Among many visual representations used for algebra teaching, one specific type called “tape diagrams” is used frequently in practice (Fukuda et al., 2021; Murata, 2008) and in research (Booth & Koedinger, 2012; Chu et al., 2017). Most commonly, tape diagrams are used during instruction in Japan and Singapore, two countries where mathematics performance is high compared to the performance of many other developed countries (Fukuda et al., 2021; Lee et al., 2013; Murata, 2008). For instance, Singapore and Japan are ranked #1 and #4, respectively, for the 8<sup>th</sup> grade mathematics achievement ranking in the 2019 Trends in International Mathematics and Society Study (TIMSS) (Mullis et al., 2019). As shown in Figure 2, tape diagrams use bar-type representations to show how the different quantities are related in an equation (Murata, 2008). Despite its prevalence in instructional practice in Asia and a growing interest in other countries (e.g., in the US, see Common Core State Standards Initiative, 2010), however, little evidence is available regarding the effectiveness of tape diagrams for helping students learn in algebra. A small number of prior studies have investigated the effect, on student’s problem-solving performance, of providing tape diagrams alongside the corresponding equation (Booth & Koedinger, 2012; Chu et al., 2017). These studies showed that tape diagrams can lead to increased accuracy in problem solving and can reduce conceptual errors. Yet, it was also found that tape diagrams failed to help lower-grade and low-achieving students in enhancing problem solving accuracy (Booth & Koedinger, 2012) or in understanding the meaning of tape diagrams (Chu et al., 2017). Moreover, despite the belief that tape diagrams help students with conceptual understanding (Murata, 2008), whether and how tape diagrams could support student learning had never been explored, prior to the work presented in this thesis (the studies mentioned above investigated effects on *performance*, not *learning*). In sum, while tape diagrams appear to have the potential to influence students’ conceptual and procedural learning in equation solving, it is unclear if and how they could be used to do so.



**Fig. 2.** Example tape diagrams found in prior research. (A) Booth & Koedinger, 2012, (B) Chu et al., 2017, (C) Morin et al., 2017, and (D) Murata, 2008. The figure itself was obtained from Anna Bartel, with permission.

The mixed effects mentioned above, especially the different effects for subgroups of students, raise the question of how visual representations, including tape diagrams, might support learning for a wide range of students. Such differences in the benefits suggest that visual representations may need to be re-designed to scaffold students’ sense-making of the target concepts and content well (Booth & Koedinger, 2010; Davenport et al., 2008). In particular, it is important that visual representations are designed and used based on an understanding of their “pedagogical affordances” (i.e., action possibilities of an instructional tool that would help achieve the instructional goal, Mawasi et al., 2021; Nagashima, Yang et al., 2020). Pedagogical affordances offer implications for how to effectively (re-)design instructional tools. The earlier part of my dissertation takes a lens of pedagogical affordances in designing the visual representation and learning technologies. *“How might we re-design tape diagrams so that students, especially low-performing students, could benefit from using the visual representation in algebra problem solving?”*

## 2.2 Scaffolding Learning with Visual Representations by Leveraging Pedagogical Affordances

Visual representations are considered as a form of instructional scaffolding (or instructional aids) (Hubber et al., 2010; Tippett, 2016). One role of instructional scaffolding is to support learners' sense-making of the target content and thereby facilitate effective and efficient problem solving and learning (Pea, 2004). However, designing effective visual representations for a specific learning context, learning goal, and learners requires a clear understanding of what the visual representation communicates to learners, what it adds to what learners know already, and how it will guide learners towards a desired instructional goal. In other words, understanding what affordances exist between the visual representation and target learners would help design an effective instructional scaffold.

For the above-mentioned question, *how might we re-design tape diagrams so that students, especially low-performing students, could benefit from using the visual representation in algebra problem solving?*, I used pedagogical affordances as an approach to identify how to re-design a visual representation to better scaffold learners. Pedagogical affordances refer to action possibilities of an instructional tool that could help educators and learners achieve instructional goals (Nagashima, Yang et al., 2020; c.f., Pea, 1993; Wu & Puntambekar, 2012). Identifying pedagogical affordances will help increase the likelihood that the tool will benefit student learning while helping to avoid a situation in which the tool will affect learning in an undesired way (Martin, et al., 2018; Nagashima, Yang et al., 2020).

Prior studies have investigated what pedagogical affordances are embedded in various instructional tools (Airey & Erikson, 2019; Bano et al., 2017; Foster, et al., 2011, Krauskopf et al., 2012; Wu & Puntambekar, 2012), but the pedagogical affordances are usually examined without carefully considering how they interact with actual teaching and learning contexts. These past studies on pedagogical affordances typically take a *top-down* approach, in which pedagogical affordances are identified through literature review, examinations of technological features, or inferences based on theoretical assumptions. Such a top-down approach for understanding pedagogical affordances, however, does not sufficiently inform real-world pedagogical practices, where the use of instructional tools involves a dynamic decision-making process where stakeholders define an instructional goal and examine the context in which teaching and learning occur (Lim, 2007). Gresalfi (2013) argues that affordances in a practical context of technology use need to be examined and understood in relation to whether and how individuals realize affordances. Affordances are inherently a *relational* concept (Faraj & Azed, 2012; Gibson, 1997; Greeno, 1994); that is, affordances of an object can best be analyzed by examining the interactions between the object, the

environment, and the individual(s) interacting with the object (Gibson, 1977; Gresalfi et al., 2012). Depending on the interactions, affordances may or may not be realized.

In particular, when it comes to pedagogical affordances, instructional goals play an important role in that certain affordances will become more relevant than others depending on the goal defined (Bower, 2008; Dickey, 2003; Foster et al., 2012; Mawasi et al., 2021). Furthermore, *top-down* methods may overlook affordances that are less obvious (Bower, 2008). Hence, existing approaches are not necessarily adequate for capturing pedagogical affordances and constraints that offer pragmatic implications, making them less useful for dynamic instructional contexts. There is a growing need for addressing complex instructional contexts that involve different stakeholders' goals when designing and adopting tools (e.g., Ahn et al., 2019; Holstein et al., 2019). Therefore, it is increasingly important that the pedagogical affordances of an instructional tool are identified and understood in relation to a targeted instructional goal.

A grounded, human-centered approach may be effective for identifying pedagogical affordances and constraints that consider instructional practices. Particularly, an approach that incorporates educators' pedagogical knowledge would be effective, given that educators play an essential role when instructional tools are adopted in classrooms and help students realize affordances (Gresalfi et al., 2012; Koehler & Mishra, 2009; Krauskopf et al., 2012; Putnam & Borko, 2000; Tan et al., 2012). Educators have professional pedagogical knowledge known as Technological Pedagogical Content Knowledge (TPACK) (Foster et al., 2012; Kohler & Mishra, 2009; Krauskopf et al., 2012). Their ability to identify and react to these affordances allows them to leverage instructional tools and support students as they interact during learning. It is, therefore, worth exploring how educators' pedagogical knowledge could be leveraged in analyzing pedagogical affordances of instructional tools. For example, Dickey (2003) explored pedagogical affordances of a 3D virtual learning environment for the goal of promoting constructivist learning by observing and interviewing a university instructor. Their *bottom-up* approach investigated the interaction between the tool, the instructor, and students to understand affordances. However, the process of identifying affordances was rather unclear, making the method unreproducible. To date, no systematic, goal-oriented method for eliciting pedagogical affordances and constraints using educators' pedagogical knowledge is available, despite its potential usefulness for understanding pedagogical affordances and constraints. In Study 1, described below, I developed *Pedagogical Affordance Analysis*, a novel, systematic, user-centered method for identifying pedagogical affordances. I applied the approach to create a version of tape diagrams that is well-suited to the goal of enhancing conceptual knowledge among middle-school students.

Pedagogical Affordance Analysis (Study 1) informed an instructional idea for a scaffolded form of diagrammatic self-explanation, a strategy in which students would *explain* target learning content by using diagrams (Ainsworth & Iacovides, 2005). I designed, tested, and established two different ways in which diagrammatic self-explanation can be embedded as an interactive scaffold in an Intelligent Tutoring System: “confirmatory diagrammatic self-explanation” (i.e., students use diagrams to *explain* their problem-solving steps that they have already solved in a conventional, non-diagrammatical notation, in Study 2) and “anticipatory diagrammatic self-explanation” (i.e., students use diagrams to *explain* their future problem-solving steps before attempting to solve the step in a conventional, non-diagrammatical notation, in Studies 3-5). Figures 3 and 4 show how the team and I designed and implemented confirmatory and anticipatory diagrammatic self-explanation in the domain of early algebra.

A series of classroom studies (described more in detail in Chapter 3) have shown that these diagrammatic self-explanation scaffolds, added to the conventional problem-solving notation, support students’ problem-solving performance (during the learning activity) and their learning outcomes. Across four classroom experiments (Studies 2-5), I found that 1) confirmatory diagrammatic self-explanation supports students’ conceptual learning, 2) anticipatory diagrammatic self-explanation (Figure 4) enhances students’ problem-solving performance during the learning activity, and that 3) anticipatory diagrammatic self-explanation helps students makes an effective transition from using informal problem-solving strategies (e.g., guessing) to the formal problem-solving strategy in algebra (Koedinger et al., 2008). Please read more on my past studies in Chapter 3.

(a) **Which diagram shows the current step?**

$19 = 3x + 4$

Hint Finish Problem

(b) **Please solve for x**

$19 = 3x + 4$

$15 = 3x$

Hint Finish Problem

(c) **Which diagram shows the current step?**

$19 = 3x + 4$

$15 = 3x$  ✓

Hint Finish Problem

(d) **Please solve for x**

$19 = 3x + 4$

$15 = 3x$  ✓

$x =$

Hint Finish Problem

**Fig. 3.** Confirmatory diagrammatic self-explanation. (a) Students first select a diagram that represents the given equation (i.e.,  $19 = 3x + 4$  in this example). (b) Then they will solve the first step in the symbolic form (i.e., subtracting 4 from both sides of the equation). (c) After that, students will select a diagram that matches what they just did (i.e., subtracting 4 from both sides). (d) This interaction continues until students solve the problem.

(a) Which diagram shows the current step?  $7x = 4x + 9$

(b) Which diagram shows the current step?  $7x = 4x + 9$

(c) Please solve for x  $7x = 4x + 9$

(d) Please solve for x  $7x = 4x + 9$

(e) Which diagram shows the current step?  $7x = 4x + 9$

3x = 9 ✓

**Fig. 4.** Anticipatory diagrammatic self-explanation. (a) Students start by selecting a diagram that matches the given equation (i.e.,  $7x = 4x + 9$  in this example). (b) Then they will select a diagram that shows a strategic next step to do. (c) Students, once they select a correct representation (*explain* through selecting a diagram), they will be prompted to enter the step in the symbolic representation. (d) Students enter the step in symbols. (e) This interaction continues until solving the equation. [[Try the tutor here](#)]

### 2.3 Learning to Self-Regulate the Use of Visual Representations

Although the question of *how students learn with visual representations* has been researched extensively, including in my past studies, there is scarce research on how students *learn to self-regulate the use of visual representations* (i.e., being able to strategically choose whether, when, and how to use visual representations). As illustrated in Chapter 1, it is critical that learners acquire the self-regulated skill of appropriately choosing to use external resources (e.g., hints in Cognitive Tutors, Roll et al., 2011), including visual representations (Uesaka et al., 2007). According to Pintrich (1995, 1999) and Zimmerman (1990), self-regulated learning involves learner control over their own behaviors, motivation, and cognition. Using visual representations strategically is an important self-regulation skill; based on theories of self-regulation, it can be derived that learners would attempt to control whether they will use a visual representation or not during problem solving (behavior) by assessing if they understand the visual and how beneficial it would be to use the visual (cognition) and by developing the confidence in using the visual (motivation/self-efficacy).

Without such a self-regulatory skill of using visual representations, learners would not be able to use visual representations effectively and efficiently in future problem-solving situations, both in and out of school (Uesaka & Manalo, 2006, 2012; Uesaka et al., 2010). In fact, it is known in prior research that learners do not spontaneously use visual representations when there is no support for using visual representations (Uesaka & Manalo, 2006). Few studies have been conducted to investigate when and how learners spontaneously choose to use visual representations during problem solving. For example, Uesaka et al. (2010) tested if verbally encouraging students to use visual representations and assigning practice in diagram construction tasks would facilitate spontaneous use of diagrams with 86 Japanese 8th graders. They found that the combination of verbal encouragement and practice facilitated spontaneous diagram use in a subsequent practice session, but spontaneous use of diagrams did not increase when only one of those strategies was provided (Uesaka et al., 2010). Another study has found that peer instruction on using visual representations (Uesaka & Manalo, 2007) is effective in promoting spontaneous visual use (Uesaka & Manalo, 2012). More broadly, there are several other prior studies that examined students' spontaneous use of available resources other than visual representations (e.g., Duffy & Azevedo, 2015; van Harsel et al., 2022). Some studies also demonstrated that metacognitive interventions could help students to be self-regulated (e.g., Roll et al., 2011; Schwonke et al., 2013). For instance, interventions such as explicit instruction on a targeted self-regulatory skill can help learners become more self-regulated (also see Zepeda et al., 2015).

Despite the growing body of studies on self-regulated learning skills, none of the prior studies on the use of visual representations investigated whether students' choices in using visual representations result in better performance (during the learning process) and learning outcomes, and what interventions could help students be strategic in using visual representations (e.g., Bielaczyc et al., 1995). In other words, it is not yet known if any intervention could support learners in strategically choosing to use visual representations *and* how it may or may not improve performance and learning (Wu et al., 2020). It is vital to investigate whether students' self-regulated use of visual representations can result in better performance and learning. Therefore, research is needed to understand what support would help learners strategically choose to use visual representations and help them achieve better performance and learning (Chin et al., 2019; Bransford & Schwartz, 1999; Nagashima, Tseng et al., 2022; Schwartz & Arena, 2013).

What is a strategic use of visual representations (tape diagrams, in my studies)? Is more use always better than any other use patterns? Given a lack of past studies on strategic use of visual representations, it is hard to define what exactly constitutes a strategic use of visual representations. However, with regard to this question, my dissertation suggests some principles that help us understand what a strategic use of visual representations would look like, at least in the context of using tape diagrams during algebra problem solving.

## 2.4 Promoting Strategic Use of Visual Representations

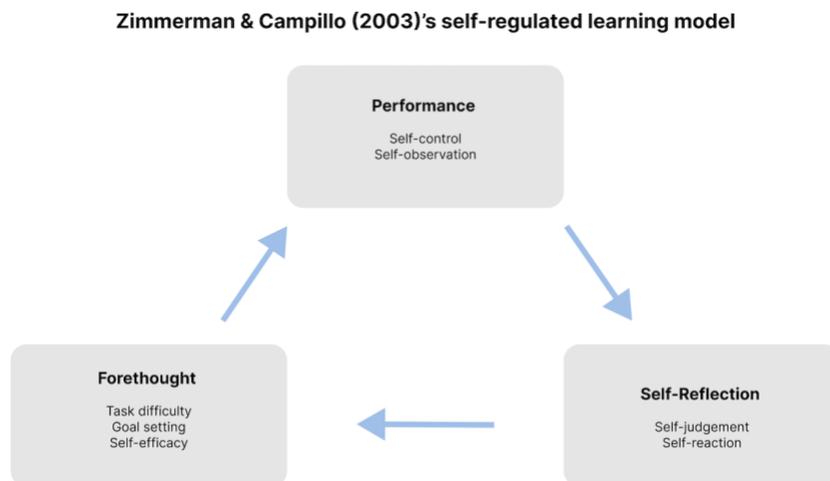
*How might we facilitate self-regulatory use of visual representations during problem solving that would help students perform and learn better?* To approach this question, I have explored relevant literature in the areas of self-regulated learning and behavior change, two fields that can inform effective approaches for my research question. In this section, I review key relevant theoretical views, existing models, and relevant prior work in these areas.

### Self-regulated learning perspective

Self-regulated learning can be defined as “the process whereby students activate and sustain cognition, behaviors, and affects, which are systematically oriented toward attainment of their goal” (Sabourin et al., 2013; Schunk, 2008). While various other definitions of self-regulated learning exist in the literature (Pintrich, 1995; Zimmerman, 2002, 1999, 1990, 1986), self-regulation of learning can be characterized by its two key features and processes. First, self-regulated learners *proactively* seek and choose to adopt learning strategies (that are often effortful) and are aware of the processes

and outcomes associated with the adoption of certain learning strategies (Cleary & Zimmerman, 2004; Zimmerman, 2013). Second, self-regulated learners carefully observe and examine their proactive adoption of certain learning strategies, often based on feedback they receive (e.g., on their own learning outcomes, processes, and skill acquisition). In other words, self-regulated learners do not only choose to use effortful-yet-effective strategies but also attend to the effectiveness of using the strategies in relation to their own knowledge and skills, self-efficacy, motivation, and actual benefits that the adoption of the strategy produce (Cleary & Zimmerman, 2004). Learners may adjust or quit the use of the strategies if they (metacognitively) do not see benefits and usefulness for their learning and own goals.

Researchers have investigated processes involving self-regulated learning to uncover metacognitive, cognitive, and motivational processes that play a role in self-regulated learning (Ben-Eliyahu & Bernacki, 2015). They have proposed several different models and theoretical frameworks that attempt to describe the processes involving self-regulation. For example, Zimmerman and Campillo (2003) describe self-regulated learning processes through a cyclical model that is composed of “forethought”, “performance”, and “self-reflection” phases (Figure 5). In the *forethought* phase, students analyze the presented task (e.g., its difficulty and familiarity) and plan a learning goal (Cleary & Zimmerman, 2013). Another core characteristic of the *forethought* phase is its emphasis on students’ motivational aspects; students analyze tasks and establish goals based on their self-efficacy beliefs, interests, and utility value they see in the task (i.e., task utility) (Panadero & Alonso-Tapia, 2014). In the *performance* phase, students carry out the learning task, observe their own learning processes, and control their behaviors based on the needs they see during performance. For example, students might change the strategy they use based on what they observe in their own task performance. Finally, in the *self-reflection* phase, students evaluate their own performance, and react to it to improve their performance. These sub-components within the self-reflection phase are called *self-judgement* and *self-reaction*, respectively (Panadero & Alonso-Tapia, 2014) and they are dependent on each other. For instance, if learners realize that there is room for improvement in their own work (*self-judgement*), they may adaptively adjust their strategy use to improve their performance (*self-reaction*).



**Fig. 5.** Zimmerman & Campillo (2003)'s model of self-regulated learning. It consists of three stages in a cyclic form (“forethought”, “performance”, and “self-reflection” stages)

Another model that is widely used in education research, especially studies of self-regulated learning in computer-supported learning environments (Matcha et al., 2019), is Winnie and Hadwin (1998)'s COPES model that involves four main phases: *task definition*, *goal setting and planning*, *enactment of tactics and strategies*, and *adaptation* (Matcha et al., 2019; Winne & Hadwin, 1998). These four stages are described as:

1. *Task definition*: Learners develop a model of the task.
2. *Goal setting and planning*: Learners create goals relative to their model of the task and then select cognitive operations—study tactics and learning strategies—they forecast could achieve their goals.
3. *Enactment of tactics and strategies*: Learners apply tactics and strategies. As they do, tactics and strategies create provisional updates to knowledge and beliefs.
4. *Adaptation*: As operations create products and when information evaluating products is accessible, learners may monitor learning and adapt features of the three foregoing phases if progress deviates from standards specified in goals.

As seen in the description above, Winne and Hadwin (1998)'s model and Zimmerman and Campillo (2003)'s model share similar phases and processes that involve learning. Zimmerman and Campillo (2003)'s *forethought* phase corresponds to Winne and Hadwin (1998)'s *task definition* and

*goal setting and planning* stages. The *performance* phase involves *enactment of tactics and strategies*, and the *self-reflection* phase (Zimmerman & Campillo, 2003) corresponds to the *adaptation* phase (Winne & Hadwin, 1998). However, it is possible that some adjustments are made, depending on the learning context. For instance, *Adaptation* may happen during performing the task (i.e., the *performance* phase in Zimmerman & Campillo, 2003). Likewise, *self-reflection* might happen while performing the task.

Studies have applied these theoretical models to investigate student learning, and design and test interventions aimed at supporting students' self-regulated learning in various subject domains. In relation to my dissertation topic, some studies investigated students' self-regulated use of learning strategies when learners had the choice of using (or the order of using) those learning strategies. For instance, Foster et al. (2018) found that novice first-year college students, when they are given a choice as to whether to use example problems or practice problems, do not choose to select tasks following the example-based-learning principle (i.e., start with examples, and gradually transition to practice problems) in mathematics. van Harsel et al. (2022) ran a study with university students investigating students' use of worked examples, tutorial videos, and practice problems when the choice as to which of those types of learning materials to use for their mathematics learning was up to students. Results showed similar patterns to Foster et al. (2018)'s findings that students did not follow established learning principles in learning with the three types of materials. van Harsel et al. (2021) also conducted an experiment in which they tested the effect of a self-regulated learning intervention (i.e., a video teaching established principles for learning with examples and practice problems). They found that students who received the intervention showed choice behaviors that follow the established effective principles but it did not lead to better learning outcomes (van Harsel et al., 2021).

Studies have also used computer-supported learning environments to investigate students' choice behaviors and their relations with performance and learning (Greene et al., 2021). For example, a series of studies on "choice-based assessments" (Schwartz & Arena, 2013; e.g., "Posterlet" by Cutumisu, et al., 2019, 2015) have found that, when students were given a choice of seeking either positive or negative feedback on the design of posters that they made in a computer-supported interactive learning environment, the frequency of seeking negative feedback was significantly correlated with learning gain on poster design principles. Their findings show the importance of measuring and focusing on students' choice behaviors in research on learning technologies. Another example is Schwonke et al. (2013) in which they used "cue cards" that show hints that students can use to learn when and how to use system-provided hints in an ITS for geometry. An experiment with 60 students in Germany found that students with low prior knowledge benefited from the

metacognitive prompt. They also found that students, regardless of their prior knowledge level, solved problems more efficiently (less time spent on the task). However, the intervention did not affect how students used hints in the tutoring software (i.e., did not affect their choice behaviors) (Schwonke et al., 2013). Roll et al. (2011) developed the Help Tutor, an ITS that provides instruction on strategic help-seeking behaviors. Their classroom studies have found that the Help Tutor improved students' help-seeking behaviors and those strategic help-seeking behaviors transferred to a new learning environment (Roll et al., 2011). The most successful intervention aimed at supporting students' choice behaviors (not only *measuring* students' choices) to date is perhaps Chin et al (2019)'s experiment in which they tested the effect of instruction of design thinking strategies (Rauth et al., 2010) on students' choice behaviors in a transfer environment. They found that instructing design thinking helped lower-prior knowledge students choose to use important design-thinking principles (e.g., seeking feedback) in a transfer environment (Chin et al., 2019).

Despite the growing interest in measuring students' choice behaviors as described above (Chin et al., 2019; Schwartz & Arena, 2013), few studies address self-regulated use of visual representations. When use of visual representations is considered, most studies focus on *spontaneous* use of visual representations as an autonomous, "ideal" behavior that learners should engage with. For instance, Uesaka and Manalo (2010) found that teachers' verbal encouragement on diagram use and offering practice opportunities for using diagrams helped students spontaneously draw diagrams (on a paper) in math problem solving. Also, Wu et al. (2020) tested the effect of a computer-supported drawing prompt (i.e., suggesting the use of drawing to support understanding of the content) on students' spontaneous drawing (on paper). They found that students who received the prompt self-reported that they used drawings more than those who did not receive the prompt. However, the data came from students' self-report and the lack of fine-grained data on students' actual choice behaviors (when and how students drew) makes it difficult to understand the effect of the prompt on students' choice behaviors and how they might be associated with students' performance (Ben-Eliyahu & Bernacki, 2015; Bernacki, 2017; Zhou & Winne, 2012). Other studies also exist where students had an option of using or not using visual scaffold (Duffy & Azevedo, 2015) or choosing to use among multiple types of visual scaffolding (Mavrikis et al., 2018; Rummel et al., 2016). However, these studies do not focus on students' choices involving use of visual scaffolding (Duffy & Azevedo, 2015), and even when they do, they do not focus on how learners' choice behaviors and their learning and performance are associated (Mavrikis et al., 2018).

Furthermore, despite the importance of using temporal data in examining students' self-regulated choice behaviors (Greene et al., 2021), past studies on students' choice making in an interactive learning environment (and other contexts) do not provide any insights into how students' choice

behaviors might change over time. It is reasonable to think that learners' adoption of self-regulated choice behaviors is not consistent (Bernacki, 2017; Bernacki et al., 2015). For instance, learners may first use ineffective strategies to explore the learning environment and may gradually learn to use effective strategies (as they gain knowledge and skills during learning). Aggregated data (e.g., average number of using certain strategies) might obscure such meaningful changes in the choice behaviors (Zhou & Winne, 2012).

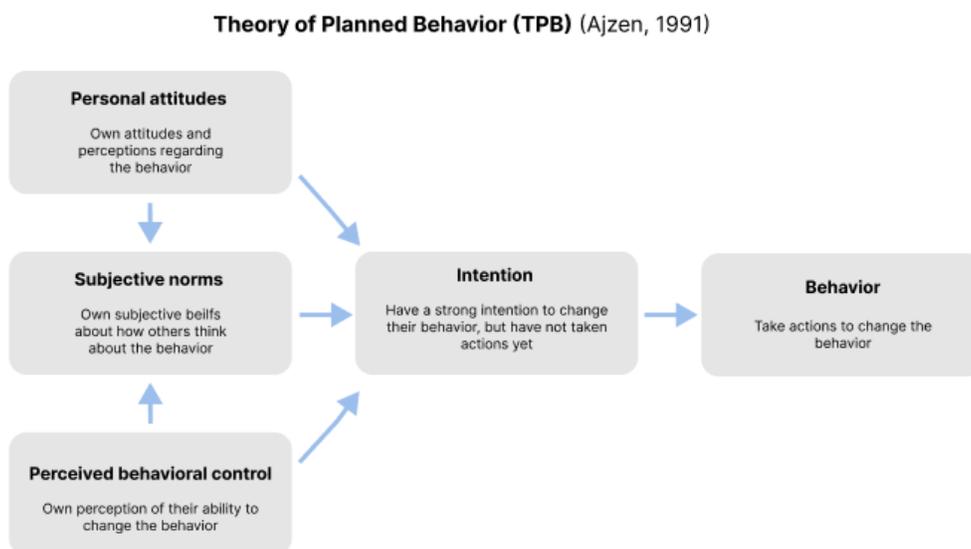
### **Behavior change perspective**

Another established area of research that is closely related to self-regulated learning is the field of behavior change. Behavior change theories, most frequently used in fields such as healthcare and criminology, aim to understand and support people's (stages of) behavior changes and their processes. Despite several shared characteristics with self-regulation theories in learning, behavior change theories differ in that they, as the name suggests, tend to focus more on behavioral outcomes than on metacognitive and motivational processes. Regarding this difference, Opezzo and Schwartz (2013) note that the goal of self-regulated learning interventions is to develop "habits of mind" whereas that of behavior change focuses on developing "habits of behaviors" (p. 485).

Behavior change models typically depict multiple stages that one would go through to change a less-desired behavior to an ideal behavior. For example, The Transtheoretical Model (TTM) of Behavior Change (Prochaska et al., 1992) describes five stages of behavior change (*pre-contemplation, contemplation, preparation, action, and maintenance*, see Figure 6). This linear model, however, assumes that a logical person would proceed through these stages of behavior change to attain an ideal behavior through coherent, reasonable decision-making processes, which is not often true (West, 2005). Another model used widely is the Theory of Planned Behavior (TPB, Ajzen, 1991). Figure 7 illustrates the model. This model, extended based on another similar theory called the Theory of Reasoned Action, focuses on one's perceptions regarding the behavior (*personal attitudes*), their own beliefs about how the behavior is perceived by others (e.g., "All of my friends smoke too, so smoking is not bad", *subjective norms*), and self-efficacy involving the behavior change (*perceived behavioral control*, Bandura, 2000). However, it simplifies the change from *intention to behavior*, failing to capture metacognitive processes that one would go through (see models of self-regulated learning introduced earlier).



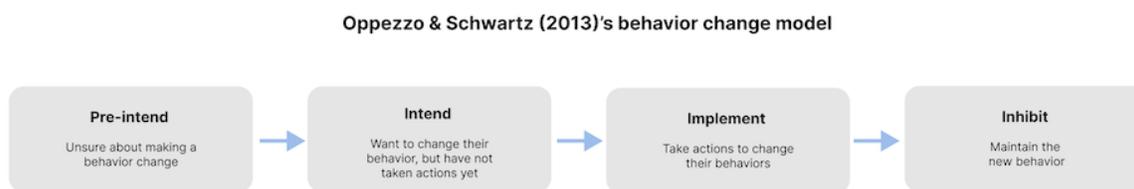
**Fig. 6.** The Transtheoretical Model of Behavior Change (TMM) by Prochaska et al. (1992). It consists of five stages and it is expected that people, when engaged in behavior change, follow these stages.



**Fig. 7.** Theory of Planned Behavior (TPB) by Ajzen (1991). The theory focuses on one's own beliefs and self-efficacy regarding the behavior change that contributes to their intention and behavior change.

Oppezzo and Schwartz (2013) abstracted these traditional models to create a four-stage model (*pre-intend*, *intend*, *implement*, and *inhabit* stages) and discussed its application to self-regulated learning in an interactive learning environment (Figure 8). The *pre-intend* phase is the state in which students are still not sure whether to invest their effort to adopt the new strategy (e.g., seeking negative feedback). When students understand the benefits of using the strategy, they will then move to the *intend* phase. In this phase, they have not yet implemented the strategy, but they have understood benefits of using it and they are interested in changing their behavior. In the *implement*

phase, students implement the strategy. The *inhabit* phase is where students sustain the behavior after implementing the strategy. To sustain the behavior, they often engage in metacognitive processes such as self-reflection and self-evaluation to determine their performance and outcomes of the adoption of the new behavior. In the *inhabit* stage, students no longer need to be motivated to use the strategy (Oppezzo & Schwartz, 2013).

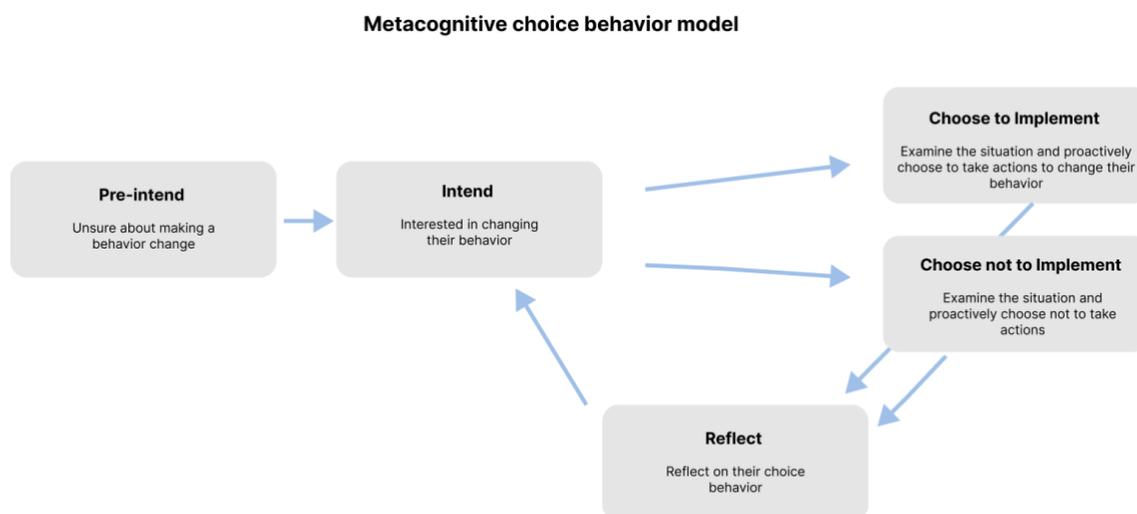


**Fig. 8.** A model proposed by Oppezzo & Schwartz (2013) based on several preceding behavior change models. It simplifies stages of change into four linear phases.

A distinctive characteristic of many behavior change models and interventions is that they target behavioral outcomes that are considered undoubtedly important by the society, and therefore they encourage people to maintain it after the behavior change has occurred. For instance, for a popular target behavior such as *stop smoking*, once the behavior change has occurred (i.e., when one stops smoking), the goal of many behavior change models is to maintain it to the level that the person does not require further motivation to keep the behavior. This assumption limits the applicability of many behavior change models to the fields such as the learning sciences where “ideal” target behavior outcomes can involve a complicated learning activity with deep metacognitive processes regarding the adoption of the behavior. For example, when the target outcome is the use of visual representations, there may be cases where the use of visual representations is not ideal (e.g., to avoid over-scaffolding students’ problem solving, Long & Alevan, 2017). There have not been attempts in behavior change literature that are aimed at supporting behaviors that involve strategic use of the target behavior (i.e., where a key question is “how and in what situation should students use the strategy?”).

My dissertation will address this complexity in researching students’ choice behaviors through integrating behavior change models and models of self-regulated learning. Specifically, I propose Metacognitive Choice Behavior Model (Figure 9), developed based on Oppezzo and Schwartz (2013) and Zimmerman and Campillo (2003)’s models introduced earlier. The Metacognitive Choice Behavior Model divides learners’ main behaviors into two categories (*choose to implement* and

*choose not to implement*) to address learners' multiple choice behaviors in the use of a target strategy and thought processes. It also incorporates a cyclic form of self-regulation (Zimmerman & Campillo, 2003) and therefore connects the *reflect* phase with the *intend* phase so that students' reflection feeds back to their intention of using the strategy for the next opportunity. An important characteristic of this model compared to other behavior change models is that the goal of the Metacognitive Choice Behavior Model is to promote metacognitive processes involving choice behaviors (e.g., deep thinking about when to use a target strategy) rather than promoting a certain choice behavior than others.



**Fig. 9.** Metacognitive Choice Behavior Model. In this model, choice behaviors are divided into two behavior types (*choose to implement* or *choose not to implement*) to appropriately capture the choice behaviors and the thought processes. Also, the *reflect* phase feeds back to the *intend* phase since learners' reflection of their choice behaviors might influence their intention regarding the use of the target strategy when the next opportunity arises.

### 3 Chapter 3: Learning with Visual Representations

#### 3.1 Study 1: Designing Visual Scaffolding with Pedagogical Affordance Analysis (PAA)<sup>1</sup>

In the first part of my dissertation work, I explored how we might design visual representations for supporting student learning of conceptual knowledge in early algebra, starting out with a user-centered research study with middle-school mathematics teachers in the US. This qualitative investigation aimed at identifying pedagogical affordances and constraints of tape diagrams. Through this work, I developed Pedagogical Affordance Analysis (PAA), a novel systematic, human-centered, action-oriented method for eliciting pedagogical affordances and constraints of an instructional tool by leveraging educators' pedagogical knowledge (Nagashima, Yang et al., 2020). PAA can be applied to both an existing tool (e.g., for the purpose of when analyzing how an existing tool is aligned with the in-tended goal of instruction) and early prototypes of a to-be-designed tool (e.g., for the purpose of improving the design or functionality of an early prototype) (Coburn & Penuel, 2016). It employs a human-centered approach in identifying affordances and constraints that are specific to an instructional goal defined up front.

In what follows, I describe PAA and its procedure. I then present a case study of its application to designing tape diagrams. To the best of our knowledge, PAA is the only available systematic method that elicits educators' pedagogical knowledge in understanding pedagogical affordances and constraints.

#### **Pedagogical Affordance Analysis (PAA): Definition and procedure**

I define pedagogical affordances as action possibilities of an instructional tool that could potentially help achieve a particular instructional goal (also see Masoudi et al., 2019, McGrenere & Ho, 2000). We define pedagogical constraints as action possibilities or incapability of an instructional tool that

---

<sup>1</sup> You can read more about this work in the following publications:

Nagashima, T., Yang, K., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2020). Pedagogical Affordance Analysis: Leveraging teachers' pedagogical knowledge for eliciting pedagogical affordances and constraints of instructional tools. In M. Gresalfi & I. S. Horn (Eds.), 14th International Conference of the Learning Sciences

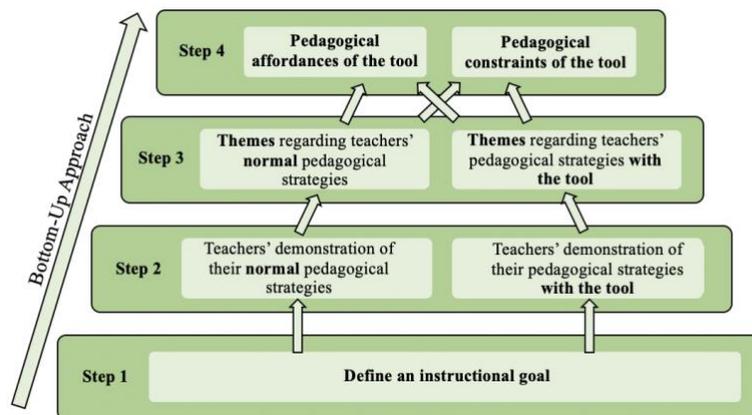
Nagashima, T., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2020). Enhancing conceptual knowledge in early algebra through scaffolding diagrammatic self-explanation. In M. Gresalfi & I. S. Horn (Eds.), 14th International Conference of the Learning Sciences (pp. 35-43). Nashville, TN: International Society of the Learning Sciences.

would put a limit on achieving the goal. Constraints, in other words, are limitations of the tool (or constraints on the utility of the tool) with respect to a stated pedagogical goal. By leveraging educators' unique knowledge, researchers and designers can find pedagogical affordances and constraints that would otherwise be missed if educators were not included in the processes of analysis and design. PAA is characterized by its three unique features:

- *Action-oriented*: In PAA, designers or researchers work with educators, or professional practitioners who possess real-world pedagogical knowledge, in contrast to the top-down approaches, which do not involve teaching experts. Educators are asked to demonstrate their pedagogical knowledge on one or more pedagogical tasks that are relevant to the targeted instructional goal(s). PAA aims to identify pedagogical affordances and constraints by analyzing the demonstrated pedagogical strategies rather than directly asking educators to identify affordances and constraints.
- *Goal-oriented*: PAA requires a specific instructional/learning goal, an integral part in designing and adopting instructional tools in general. PAA aims to elicit pedagogical affordances and constraints in relation to the defined goal. This allows designers to identify pedagogical affordances that are relevant and specific to the goal. It also helps identify pedagogical constraints, which are not typically explored but offer useful information regarding the design and analysis of the tool.
- *Comparative*: In PAA, participants are asked to demonstrate their usual pedagogical strategies without the tool and then potential approaches using the target tool on the same task. Unlike other available approaches, PAA systematically elicits pedagogical affordances and constraints by comparing and contrasting those two types of demonstrations, discovering affordances and constraints relative to participants' preferred pedagogical approaches. This approach also helps identify design implications that can be obtained through comparison (e.g., incorporating an element of a teachers' usual pedagogical strategy into the design of the target tool to overcome a pedagogical constraint).

PAA comprises four steps (Figure 10), some of which are informed by existing methods such as Cognitive Task Analysis (Clark & Estes, 1996) and prior attempts at measuring PCK (e.g., Krauss et al., 2008).

1. In Step 1, designers and educators meet to set an instructional/learning goal to be targeted by the tool of interest.
2. In Step 2, designers (potentially together with educators) decide and provide teachers one or more pedagogical tasks targeted at the given goal and ask them to demonstrate a pedagogical strategy that they would usually choose for each task.
3. Following the demonstration of their usual strategy, for each of the same set of tasks, teachers are asked to demonstrate a pedagogical strategy that they would choose if they were using the target tool. Questions regarding the strategies may be asked to the educator if necessary. The session is video- or audio-recorded for later analysis.
4. Once data collection is complete, in Step 3, designers analyze the demonstrated strategies with and without the tool using a grounded theory approach (Strauss & Corbin, 1994). They elicit themes regarding the pedagogical strategies, separately for the demonstrations with and without the target tool. Designers and educators may return to Step 2 and are encouraged to reconsider their sample of participants when necessary (i.e., theoretical sampling) (Strauss & Corbin, 1994).
5. Finally, in Step 4, designers synthesize the themes across those two demonstration types through comparison and contrast, identifying pedagogical affordances and constraints of the tool for the goal defined.



**Fig. 10.** Procedure of Pedagogical Affordance Analysis, which consists of four steps. PAA identifies affordances and constraints through examining educators' pedagogical demonstrations.

### **Applying Pedagogical Affordance Analysis**

I applied PAA to our research project in which I explored how we might re-design tape diagrams to support students' conceptual learning in early algebra.

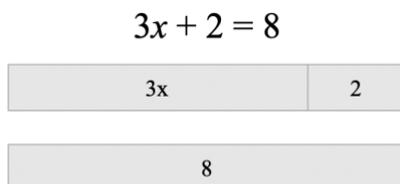
#### *Method*

I conducted PAA with eight middle school mathematics teachers in the United States who had taught algebra in their career. They participated either in-person ( $n = 2$ ) or remotely via teleconferencing software ( $n = 6$ ). Two of the teachers reported seeing tape diagrams in the past, and none reported ever using tape diagrams in their teaching.

We defined *enhancing conceptual knowledge in equation solving among low-achieving middle-school students with tape diagrams* as our target instructional goal for our application of PAA (Figure 10, Step 1). The goal in this case study was primarily defined by a researcher, namely, the author of this proposal, informed by the goal of the larger research project and the literature review conducted. However, teachers and I discussed the importance of learning conceptual knowledge in mathematics at the beginning of the PAA sessions. All eight teachers in the study agreed that teaching conceptual knowledge in algebra is both important and difficult.

In defining tasks used in PAA, as teachers' pedagogical includes knowledge about students' errors and how to help students correct their own misunderstandings (Krauss et al., 2008; Park & Oliver, 2008; Shulman, 1986), we decided to assign teachers tasks in which they would conceptually explain common errors of algebra problem solving to students. We first asked teachers to generate a few examples of common errors and suboptimal strategies they see frequently in their students' work. When they were not able to generate such errors, we showed them examples of common student errors in equation solving reported in Booth et al. (2014). For each of the errors, we first asked teachers to demonstrate, while thinking aloud, their usual pedagogical approaches (i.e., instructional approaches that teachers would usually choose in their teaching) to helping students correctly understand the concept (Figure 10, Step 2). Following that, we introduced tape diagrams, explaining that they show relationships among different quantities in an equation. We presented a simple tape diagrams together with a sample algebraic equation, in which tapes corresponded to the two sides of the equation (see Figure 11 for an example). The alignment of the tapes was not necessarily fixed in a certain position and the size of the tapes was not necessarily proportional to the actual value being represented by the tapes. We asked teachers to demonstrate, while thinking aloud, the strategies they would choose if they were to use tape diagrams in their explanations (Figure 10, Step 3). Two learning sciences graduate students analyzed approximately eight hours of video recordings following a grounded theory approach. They conducted the analysis separately for the

demonstrations with and without tape diagrams. They went from open coding to axial coding, and then conducted selective coding to elicit themes. (Figure 3, Step 4).



**Fig. 11.** Example tape diagrams.

### *Results*

By analyzing how teachers explained students' common errors and suboptimal strategies, we found five themes regarding their usual pedagogical strategies (usual strategies: US) and eight themes regarding their strategies using tape diagrams (tape diagram strategies: TDS) (Table 5). We describe these themes in detail below.

**Table 5.** Themes regarding teachers' usual pedagogical strategies and strategies with tape diagrams.

<b>Themes regarding usual pedagogical strategies</b>	<b>Themes regarding strategies with tape diagrams</b>
US1: Teachers use familiar real-world examples and plain numbers so that students can relate to their own prior knowledge	TDS1: Teachers use the lengths of tapes as a visually-intuitive representation of mathematical equivalence
US2: Teachers choose pedagogical approaches and tools that can be used for a variety of problems and operations	TDS2: Teachers use tape diagrams to visually show students how equations and equation transformations can be represented
US3: Teachers want students to make a transition from concrete to abstract thinkers	TDS3: Teachers use the size of tapes to help students understand equation transformations
	TDS4: Teachers use tape diagrams to help students avoid errors

---

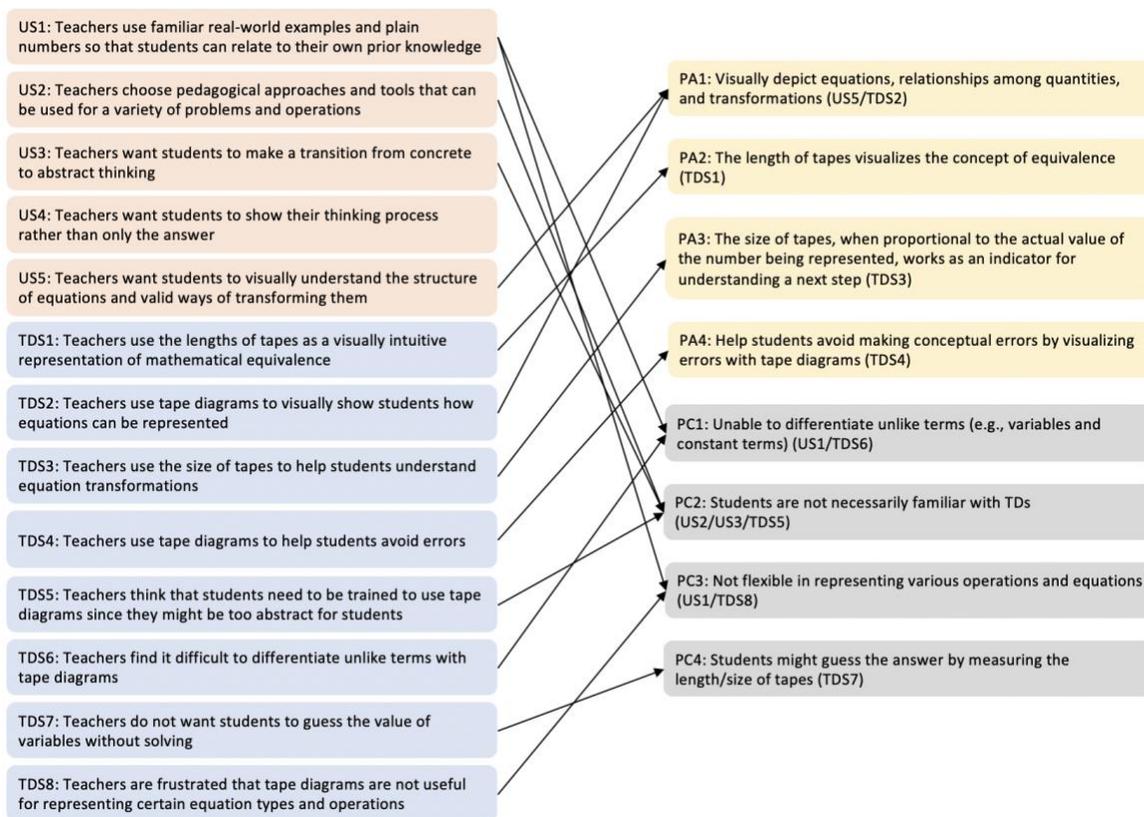
US4: Teachers want students to show their thinking process rather than only the answer	TDS5: Teachers feel that students need to be trained to use tape diagrams since they might be too abstract for students
US5: Teachers want students to visually understand the structure of equations and valid ways of transforming them	TDS6: Teachers find it difficult to differentiate unlike terms with tape diagrams
	TDS7: Teachers do not want students to guess the value of variables without solving
	TDS8: Teachers are frustrated that tape diagrams were not useful in representing certain equation types and operations

---

I then synthesized the themes to identify pedagogical affordances and constraints of tape diagrams (Table 6, Figure 12). Specifically, I compared or contrasted each TDS theme with each US theme, and then elicited an affordance or constraint of tape diagrams. For instance, TDS2 and US5 are both about communicating concepts visually. This produced an affordance, which we call PA1: *Visually depict equations, relationships among quantities, and transformations*. Similarly, contrasting TDS8 with US2 elicited a tape diagrams' pedagogical constraint of *not being flexible in representing various operations and equations* (PC3). When no relevant (similar or opposite) US themes aligned with a specific TDS is found, TDS themes themselves were classified either as affordances or constraints of tape diagrams, depending on whether the theme is about helping or hindering the goal of learning conceptual knowledge in algebra (examples of these are PA2, PA3, PA4, and PC4). Figure 12 illustrates which TDSs and USs informed pedagogical affordances and constraints we identified.

**Table 6.** Pedagogical affordances and constraints of tape diagrams for teaching conceptual knowledge (relevant themes for each, in parentheses).

<b>Pedagogical affordances of tape diagrams</b>	<b>Pedagogical constraints of tape diagrams</b>
PA1: Visually depict equations, relationships among quantities, and transformations (US5/TDS2)	PC1: Unable to differentiate unlike terms (e.g., variables and constant terms) (US2/TDS6)
PA2: The length of tapes visualizes the concept of equivalence (TDS1)	PC2: Students are not necessarily familiar with tape diagrams (US1/US3/TDS5)
PA3: The size of tapes, when proportional to the actual value of the number being represented, works as an indicator for understanding a next step (TDS3)	PC3: Not flexible in representing various operations and equations (US2/TDS8)
PA4: Help students avoid making conceptual errors by visualizing errors with tape diagrams (TDS4)	PC4: Students might guess the answer by measuring the length/size of tapes (TDS7)



**Fig. 12.** Relationships between Usual Strategies (US), Tape Diagram Strategies (TDS), Pedagogical Affordances (PA) and Pedagogical Constraints (PC)

Note that the elicited pedagogical affordances and constraints are specific to the instructional goal of *teaching conceptual knowledge with tape diagrams in early algebra*. In other words, it is very possible that some or all of these pedagogical affordances and constraints would not have been elicited if the goal had been defined differently. Also, as affordances are analyzed in relation to the individual(s) interacting with the object, those with different background knowledge, skills, and experiences might perceive different affordances and constraints (McGrenere & Ho, 2000). Therefore, the fact that I found these pedagogical affordances and constraints with teachers does not guarantee that middle-school students, the target learners, will perceive all of these pedagogical affordances and constraints (Mawasi et al., 2021; Wisittanawat & Gresalfi, 2021). Therefore, I then

explored how we might use the elicited pedagogical affordances and constraints to redesign tape diagrams as a way to offer scaffolding support for middle-school students (Clements & McMillen, 1996).

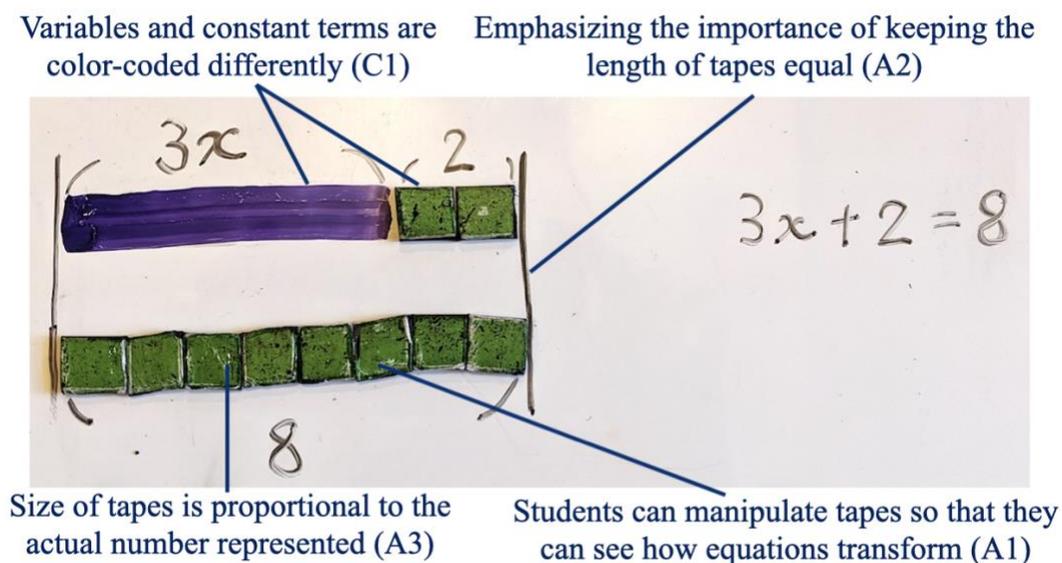
In the case study, my purpose of using PAA was not only to identify pedagogical affordances and constraints of tape diagrams but to leverage the identified affordances and constraints to redesign tape diagrams to help students to build up conceptual knowledge in algebra. Even small design features (e.g., visual signaling cues) used to highlight important instructional information can have an impact on student perception, understanding, and learning, particularly among struggling students (e.g., Barbieri et al., 2019; de Koning et al., 2010; Yung & Paas, 2015). In the following section, we describe how I redesigned the tape diagrams.

### **3.2 Study 2: Supporting Conceptual Understanding through Confirmatory Diagrammatic Self-Explanation**

#### **Design process**

After conducting the PAA, I worked with one of the participating teachers to redesign tape diagrams based on the pedagogical affordances and constraints we found. The teacher was chosen because of their willingness for helping further. The teacher and I met approximately 90 minutes to discuss the elicited affordances and constraints. During the discussion, we also brainstormed additional features that could effectively scaffold students' understanding of concepts.

The discussion largely focused on designing visual cues that could emphasize the affordances. Given that many teachers demonstrated their strategies by manipulating tape diagrams in the PAA, we first created a physically-manipulable prototype of tape diagrams with additional features based on the pedagogical affordances and constraints identified. Specifically, using magnets and a whiteboard, we developed a prototype that emphasizes PA1, PA2, and PA3 (Figure 13). The design also attempted to overcome PC1 by color-coding variables and constants, informed by one of the usual strategies (US5). The teacher and I generated these design features based on their prior teaching and design experience as well as by revisiting some of the teachers' demonstrations during the PAA. In prior studies using tape diagrams, these features had not been explicitly designed (e.g., sizes of tapes are inconsistent) or never been attended to at all (e.g., color-coding variables and constants). We also decided, after discussion, that it is better to avoid modeling certain equation types (e.g., equations with negatives), informed by PC3.



**Fig. 13.** Physically-manipulable tape diagrams, designed to emphasize pedagogical affordances and overcome/avoid pedagogical constraints. Green magnetic blocks are used to visualize constant terms.

### User study with middle-school students

#### *Method*

To investigate whether the affordances and constraints of tape diagrams are materialized when middle-school students interacted with them, I conducted one-on-one, in-person user-study sessions with seven students in the United States, whose grade level ranged from Grades 5-7. During the sessions, participants were presented with four worked examples of equations in symbolic notation containing common errors and asked to explain the errors in the examples. I used the four common errors reported by teachers the most. After students attempted to explain errors, I then introduced the participants to the redesigned tape diagrams (Figure 13). I gave a short description of what tape diagrams are and then asked them to identify the errors in the worked examples by allowing them to manipulate the tape diagrams. All sessions were video-recorded. Two researchers analyzed a total of approximately 4.5 hours of video recordings of participants' explanations of the incorrect worked

examples, and coded if and how participants' explanations of errors reflected the pedagogical affordances and constraints of tape diagrams that we found.

### *Results*

First, we found that, when asked to explain errors with tape diagrams, students who had not been able to identify errors without tape diagrams (i.e., when given only the equation in symbolic notation) tended to refer to the features of tape diagrams in their explanations, whereas those who had identified errors correctly without tape diagrams relied on their own schema of understanding concepts when explaining with tape diagrams. For example, a student who had not successfully identified any error without tape diagrams referred to the size of tape diagrams (PA3) when explaining the error of not keeping the sides of an equation equal (e.g.,  $3x + 2 = 8$  becoming  $3x + 2 - 2 = 8 - 8$ ) with tape diagrams: “you take away 2 and 8 and then, you can see that 2 isn't as the same size as 8. So, you are not taking away an even amount.” These same students also frequently mentioned when using tape diagrams that variables and constants were “different”, a design feature (namely, color coding of constant and variable terms) that, as explained, we added to overcome PC1. On the other hand, students who identified errors correctly with the symbolic notation only explained the error of not keeping the sides of an equation equal with tape diagrams by saying, “they are not equal” or “unbalanced”, which does not seem to be particularly referring to tape diagrams' unique affordances. One of them even mentioned, “heavier,” when explaining the unbalanced situation (referring to the visual representation familiar to him: a balance scale). However, the analysis also revealed that our prototype over-scaffolded students in figuring the value of the variable. Specifically, the use of unit blocks (green-colored magnets in Figure 6) allowed several students to find the answer without solving an equation, by counting the number of blocks corresponding to the size of the variable. This behavior was seen especially among students who had not been able to identify errors without tape diagrams. This instance indicates that a feature meant to support PA3 inadvertently activated PC4.

### *Discussion*

The user study suggested the potential effectiveness of emphasizing specific affordances of tape diagrams to support conceptual understanding of equation solving. The evaluation sessions suggest that the re-designed tape diagrams seemed particularly useful for students who had been unable to explain errors without the visual. It also informed us that overcoming PC1 through color-coding variables and constants differently would potentially be effective. However, the use of unit blocks as a way to visually show the meaning of the size of tapes did not show any particular benefit; rather,

it activated a pedagogical constraint by making it easy to guess the value of the variable (PC4). Because the teacher I worked with and I both agreed that the behavior of guessing should be avoided to encourage students to think of equation-solving steps conceptually and algebraically, as we describe below, we decided not to use unit blocks in the following re-design iteration.

### **Design process**

In the next design phase, to make a scalable tool that can be used by many students at a time, I decided to design a digital learning tool that leverages tape diagrams. Since prior studies on tape diagrams found that merely presenting tape diagrams of an equation did not benefit lower-grade and low-achieving students (Booth & Koedinger, 2012; Chu et al., 2017), I also decided to design an activity that integrates the redesigned tape diagrams as a core component of students' learning processes in addition to the re-design of tape diagrams themselves. The following paragraphs describe the design process of the tape diagrams themselves and the learning activity.

The re-design of the tape diagrams involved one change while keeping other features consistent, based on the findings from the think-aloud evaluation in the first evaluation study. Specifically, I kept the vertical lines to emphasize the equality, color-coded tapes to visually differentiate variables and constants, and the size of tapes, which was set proportional to the actual number represented. To avoid activating the pedagogical constraint of facilitating the guessing of the solution of an equation by measuring the length of tapes they perceive (PC4), during this second re-design interaction, I decided to use tapes/bars instead of unit blocks (see Figures 14-17).

To further help students realize pedagogical affordances of tape diagrams for conceptual understanding of algebra, I designed a learning activity that integrates tape diagrams as a core component of students' learning processes. I extended an existing Intelligent Tutoring System (ITS) for algebra (Long & Alevan, 2014) by embedding a novel form of "diagrammatic self-explanation" (Ainsworth & Iacovides, 2005), by which I mean a scaffold that helps students use a diagrammatic representation to confirm and correct their conceptual understanding in step-by-step problem solving with tape diagrams. Specifically, in the new form of diagrammatic self-explanation, which I call "confirmatory diagrammatic self-explanation," students are asked to *explain* their equation transformations by choosing an appropriate tape diagram representation from among three options given, in which two were conceptually-incorrect diagrams. The activity, designed based on the pedagogical affordances and constraints found in the PAA and through a short prototyping session with a teacher, aims to further emphasize the pedagogical affordances and avoid mis-using constraints. For example, by visualizing step-by-step equation transformations, we aimed to emphasize the affordance that tape diagrams visually depict equations, relationships among

quantities, and transformations (PA1). Also, the ITS shows incorrect tape diagram options that were modeled after students' common errors found in literature and in our user study sessions (PA4: *Tape diagrams help students avoid making conceptual errors by visualizing errors with tape diagrams*). Additionally, the activity does not include equations with negative numbers or complex equations such as those with parentheses (PC3).

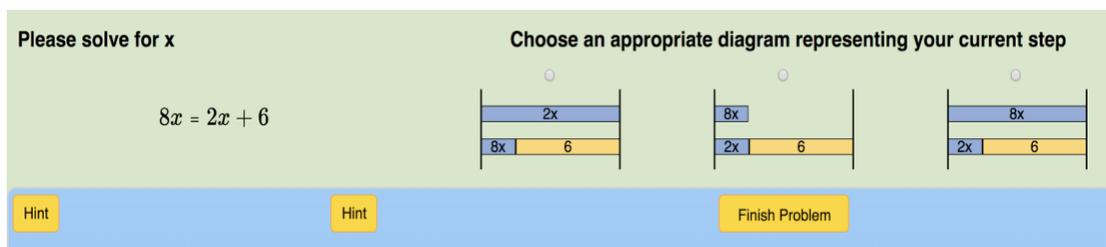
Our diagrammatic self-explanation was also informed by established instructional principles reported in the field of the learning sciences. We incorporated two instructional principles, namely, self-explanation (Bisra et al., 2018; Chi et al., 1989; Rittle-Johnson & Loehr, 2017) and contrasting cases (Schwartz et al., 2011), in the learning activity.

Self-explanation is a learning strategy in which learners attempt to make sense of what they study by generating explanations to themselves, which helps integrate new information with their own prior knowledge (e.g., Chi et al., 1989; Rittle-Johnson et al., 2017). Self-explanation has been found to enhance conceptual understanding across many areas in mathematics (e.g., Aleven & Koedinger, 2002; Rau et al., 2015; also see meta-reviews by Bisra et al., 2018; Rittle-Johnson et al., 2017). Pairing explanations with visual representations, which often contain information that is difficult to represent in text (Larkin & Simon, 1987), could potentially have an even greater effect, as both visual representations and explanations help learners process information and connect related information together (Rau, 2017). Even more effective in fostering conceptual knowledge might be prompting students to explain their thinking through visual representations, as opposed to verbal self-explanation of visual representation (Ainsworth & Loizou, 2003). However, as the traditional form of diagrammatic self-explanation (i.e., explaining through generating drawings or through sketching) may increase cognitive load due to the unstructured nature of the prompts (Ainsworth & Scheiter, 2021; Fiorella & Zhang, 2018; Wu & Rau, 2018), I wanted to create a better scaffolding prompt that can support struggling students.

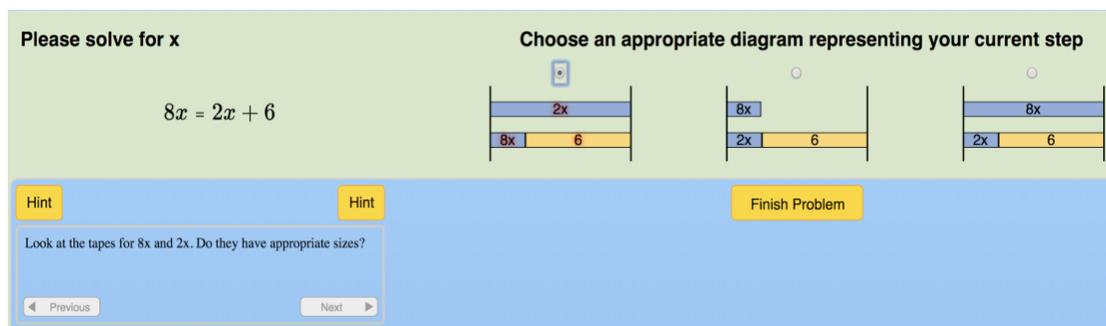
The use of contrasting cases is another established instructional strategy; Multiple examples that differ in an important instructional feature are shown next to each other to help learners notice the features. Prior research suggests that contrasting cases will best benefit learners when learners are engaged in a sense-making activity, such as explaining the cases (Schwartz et al., 2011; Sidney et al., 2015).

My confirmatory diagrammatic self-explanation activity was designed to foster sense-making of the diagrams and making connections between symbolic and diagrammatic representations by using both contrasting cases and self-explanation. First, students select an appropriate diagrammatic representation for the given equation from among three contrasting cases presented by the system (Figure 14). When they make a mistake (i.e., select the wrong diagram), the ITS gives a feedback

message (Figure 15). They then solve the equation in its symbolic form on the left-hand side of the screen (Figures 16, 17). They are asked to *explain* each of the equation transformations they perform (e.g., subtract 2 from both sides of  $3x + 2 = 8$ ) by selecting the appropriate diagrammatic representation, again from among three options given (contrasting cases, Figure 17). Diagram choices are generated automatically based on the equation transformation input that the student provides. These diagram options were designed so that each incorrect diagram differs from the correct diagram with respect to a single conceptual aspect (Schwartz et al., 2011). For example, the three diagrams in Figure 17 show the action of dividing the equation by 2, correctly (left) and incorrectly (middle and right). The diagram on the left and the diagram in the middle differ in terms of the lengths of tapes on each side in the resulting diagrams (i.e., whether the tape on the top and bottom have the same length). The diagram on the left and that on the right have the same length but differ in the action performed: division (left) or subtraction (right).



**Fig. 14.** Student selects a diagrammatic representation that corresponds to the given equation on the left.

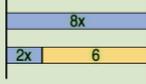


**Fig. 15.** Feedback message is given when an incorrect attempt is made. The message here says, “Look at the tapes for 8x and 2x. Do they have appropriate sizes?”

Please solve for x

$$8x = 2x + 6$$

Choose an appropriate diagram representing your current step



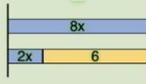
$\frac{8x}{2} = \frac{2x}{2} + \frac{6}{2}$

Fig. 16. Student solves the first step of the given equation.

Please solve for x

$$8x = 2x + 6$$

Choose an appropriate diagram representing your current step



$\frac{8x}{2} = \frac{2x + 6}{2}$

$4x = x + 3$

Hint Hint Finish Problem

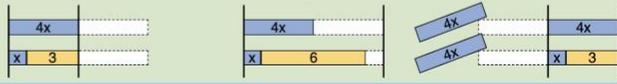


Fig. 17. Student explains the problem-solving step by selecting an appropriate diagrammatic representation.

## Study 2.2: A Classroom Experiment for Testing the Effectiveness of Confirmatory Diagrammatic Self-Explanation

### Method

I then tested whether the ITS, whose design was extended based on the elicited pedagogical affordances and constraints, would help achieve the goal of supporting low-prior knowledge students' understanding of conceptual knowledge in equation solving (e.g., lower grade, lower prior

knowledge students). I conducted a  $2 \times 2$  (Diagram/No-Diagram: diagrammatic self-explanation or no diagrammatic self-explanation, Grade: 5th or 6th grade) between-subjects pretest-intervention-posttest experiment in middle-school classrooms.

### *Participants*

A total of 45 students participated in the study (19 5th and 26 6th graders). Four students in the 6th grade were in advanced classes (two of them were in Pre-algebra and the other two were in Algebra I). All others were in their grade-level math class.

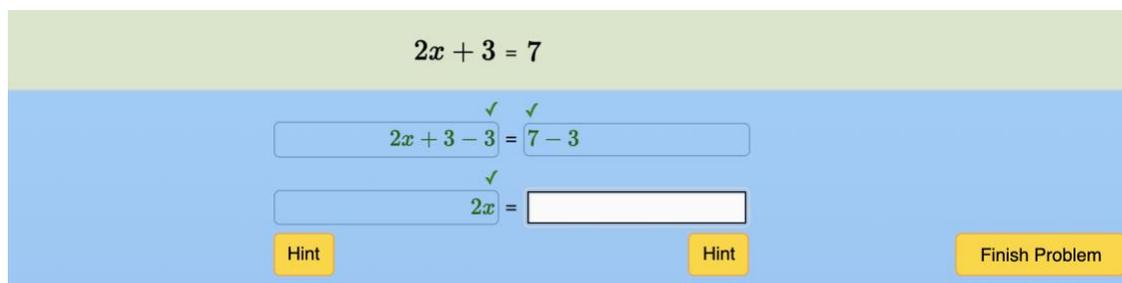
### *Materials*

Together with my collaborators at the University of Wisconsin-Madison, I developed pretest and posttest assessments to assess students' conceptual and procedural knowledge of basic algebra. The test contained several items drawn from prior studies (e.g., Chu et al., 2017; Fyfe et al., 2018; Rittle-Johnson et al., 2011) as well as some new items. The conceptual knowledge items consisted of seven multiple-choice questions and four open-ended questions, which assessed a wide range of conceptual knowledge constructs based on the categorization made in Crooks and Alibali (2014), including equality, variables and like/unlike terms, inverse operations, isolating variables, and the concept of keeping both sides of an equation equal. We also included three procedural knowledge items, including two problem-solving items (e.g., "solve for  $x$ :  $3x + 2 = 8$ ") and one multiple-choice item to test if tape diagrams also enhance students' procedural skills. In addition, to assess students' understanding of tape diagrams, we included two tape diagram items (one multiple-choice, one open-ended). We developed two isomorphic versions of the test that varied only with respect to the specific numbers used in the items; participants received one form as pretest and the other as posttest (with versions counterbalanced across subjects).

### *Procedure*

The study took place during regular mathematics class in the students' classrooms. Students within each grade were randomly assigned to either the Diagram condition or the No-Diagram condition. The advanced students mentioned above were separately randomly assigned so that each condition would have two Pre-algebra students and one Algebra I student. In the Diagram condition, students practiced with the ITS version with the confirmatory diagrammatic self-explanation. In the No-Diagram condition, students used a version of the tutor showing the equation-solving part only (Figure 18). Thus, the only difference between the Diagram and No-Diagram conditions was whether or not students self-explained their solution steps in the form of diagrams. Per teacher report, the

students had never used tape diagrams in class. Both ITSs had 20 problems across five different problem types (Table 7).

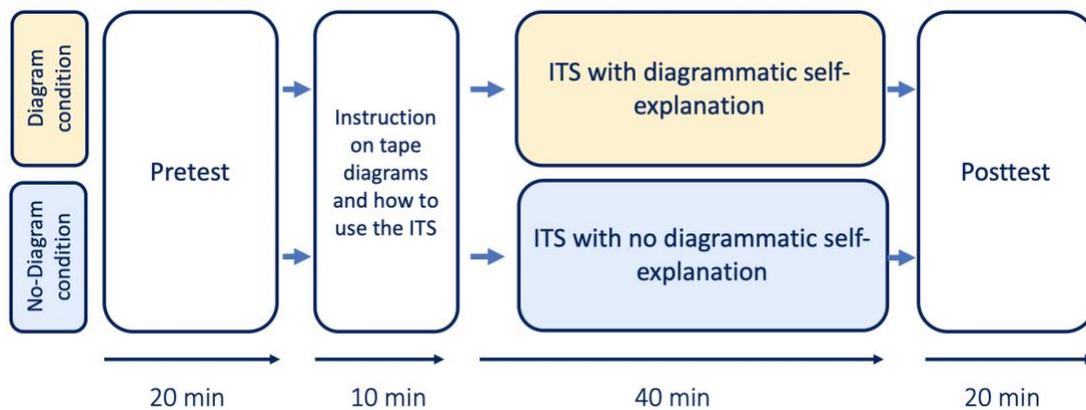


**Fig. 18.** A version of the ITS with no diagrammatic self-explanation

**Table 7.** Types of equations the tutor included and the number of problems in the tutor.

Level	Equation type	Example	Number of problems in the tutor
Level 1	$x + a = b$	$x + 3 = 6$	6
Level 2	$ax + b = c$ , $ax = b$	$2x + 3 = 7$ , $6x = 12$	6
Level 3	$ax = bx + c$	$8x = 2x + 6$	4
Level 4	$ax + b + cx + d$	$4x + 1 = 3x + 10$	4

All students first worked on the paper-based pretest for 20 minutes. Then the teacher passed out two instructional handouts and read them out loud. The handouts explained how to use the tutor and described tape diagrams, and read them out loud. Students then practiced equation solving via their assigned tutor version using school-provided computers. After 40 minutes of working with the tutor, students took the paper-based posttest for 20 minutes. Students were given access to both tutor versions after the study. Figure 19 shows the study procedure.



**Fig. 19.** Study procedure

## Results

Open-ended items were coded for whether student answers were correct or incorrect by two researchers (*Cohen's kappa* = .83). Data from three 5th-graders and one 6th-grader were excluded because they did not complete the study; therefore, data from 41 students, namely, 16 5th-graders (8 Diagram, 8 No-Diagram) and 25 6th-graders (13 Diagram, 12 No-Diagram) were included in the analysis. Table 6 presents raw pretest and posttest performance on conceptual knowledge (CK), procedural knowledge (PK), and understanding of tape diagrams (TD) items. The maximum scores for CK, PK, TD were 11, 3, and 2, respectively (Table 8).

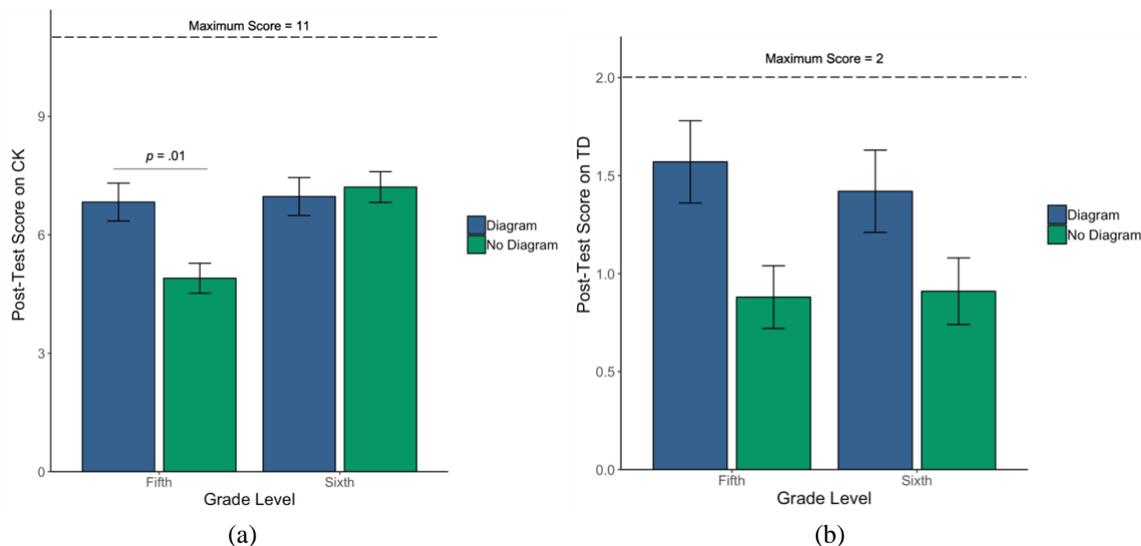
Overall pretest performance did not differ either between the conditions,  $F(1, 37) = .27, p = .60$ , or between grades,  $F(1, 37) = .94, p = .33$ ; pretest scores on each of the item subsets (CK, PK, and TD) also did not differ between conditions or grades. The average number of tutor problems attempted did not differ between conditions (Diagram:  $M = 16.14, SD = 3.35$ , No-Diagram:  $M = 17.75, SD = 3.61$ ),  $F(1, 37) = .51, p = .48$ , or grades (5th:  $M = 16.13, SD = 3.77$ , 6th:  $M = 17.44, SD = 3.34$ ),  $F(1, 37) = .44, p = .51$ .

To analyze how the conditions may have affected students' posttest performance, I conducted three ANCOVAs, each with the same independent variables (diagram condition and grade) and covariate (pretest, to control for prior knowledge), but with a different dependent variable (CK, PK, or TD). First, we investigated the effect of the condition (diagrammatic self-explanation or not) on the conceptual knowledge items (CK) on the posttest. For scores on CK items, I found a main effect of diagrammatic self-explanation,  $F(1, 36) = 8.01, p < .01, partial \eta^2 = .06$ , and a significant

interaction between diagram condition and grade,  $F(1, 36) = 6.18, p = .02, \text{partial } \eta^2 = .15$ . A post-hoc analysis revealed that 5th-graders benefited from diagrammatic self-explanation ( $F[1, 13] = 8.31, p = .01, \text{partial } \eta^2 = .39$ ), whereas no difference between the Diagram and No-Diagram conditions was found for 6th-graders (Figure 20, panel a). No significant main effects nor interactions were found for students' performance on PK items. For TD items, there was a main effect in favor of diagrammatic self-explanation,  $F(1, 36) = 5.58, p = .02, \text{partial } \eta^2 = .22$ , but no main effect of grade nor any interaction (Figure 20, panel b).

**Table 8.** Means and standard deviations (in parentheses) for CK, PK, and TD on the pretest and posttest.

	<b>CK</b> (maximum score: 11)		<b>PK</b> (maximum score: 3)		<b>TD</b> (maximum score: 2)	
	pretest	posttest	pretest	posttest	pretest	posttest
5 <sup>th</sup> Diagram	4.75 (1.28)	6.38 (1.85)	2.13 (.99)	2.00 (.93)	0.88 (.35)	1.50 (.54)
5 <sup>th</sup> No-Diagram	5.38 (2.07)	4.88 (1.89)	2.38 (.74)	2.50 (1.07)	0.75 (.46)	0.88 (.64)
6 <sup>th</sup> Diagram	6.46 (2.15)	7.23 (2.09)	1.62 (.96)	2.08 (.96)	0.92 (.28)	1.46 (.52)
6 <sup>th</sup> No-Diagram	6.50 (2.71)	7.25 (2.42)	1.33 (.98)	1.67 (1.23)	0.83 (.39)	0.92 (.80)



**Fig. 20.** Students' posttest scores on conceptual knowledge items (a) and tape diagram items (b) by grade and condition

To further investigate why only 5th-graders benefited from having diagrams, I examined the strategies that students used to solve the procedural items on the pretest and posttest. I examined strategies because students' strategy use might affect how they perform and learn with tape diagrams (Chu et al., 2017). Also, the design of tape diagrams attempted to promote the use of formal problem-solving strategies rather than informal strategies such as guessing by replacing unit blocks with bars to represent numbers. Two researchers coded the strategies by adapting the coding scheme used by Chu et al. (2017) and Koedinger et al. (2008), which includes both formal (algebraic) and informal (non-algebraic) ways of solving equations (Table 9; *Cohen's kappa* = .82). For this strategy coding, I was primarily interested in the Algebra strategy (see Table 9) because the goal of the equation-solving instruction in the school where the study took place was to help students learn the formal algebraic strategy. For the two problem-solving items on the pretest, there was a significant difference in use of the Algebra strategy between the two grades, but not between the Diagram/No-Diagram conditions. Whereas 80% of 6th-graders used the Algebra strategy at least once, only 6% of 5th-graders used it ( $\chi^2[1, n = 41] = 21.24, p < .01$ ). The difference shrank dramatically at posttest, but remained significant: 88% of 6th-graders used Algebra strategy, whereas 56% of 5th-graders used it ( $\chi^2 [1, n = 41] = 5.33, p = .02$ ). We used McNemar's test to compare the frequency of Algebra strategy use at pretest and posttest. Fifth-grade students increased their use of the Algebra strategy

significantly from pretest to posttest ( $p = .01$ ), but 6th-grade students did not ( $p = .72$ ). These findings suggest that equation-solving practice with the ITS helped the 5th-graders transition from informal to formal strategy use. This shift is masked in the raw pretest to posttest accuracy data on equation-solving items, but it is visible in the strategy choices that students made.

**Table 9.** Strategies used to solve equations, adapted from Chu et al. (2017) and Koedinger et al. (2008).

Strategy name	Description	Example answer for $3x + 2 = 8$
Algebra	Student uses algebraic manipulations to find an answer	$3x = 6$ $x = 6/3 = 2$
Unwind	Student works backward using inverse operations to find an answer	$8 - 2 = 6$ $6/3 = 2$
Guess & Check	Student tests potential solutions by substituting different values	$3*2 + 2 = 8$ $6 + 2 = 8$
Other	Student uses other non-algebraic strategies	$3 + 2 = 5$ $8/5 = 1.6$
Answer Only	Student provides an answer without showing any written work	$x = 2$
No Attempt	Student leaves problem blank or explicitly indicates that she/he does not know how to solve the problem	“I don’t know”

### Discussion

In Study 2, conducted in school classrooms, we experimentally investigated whether our diagrammatic self-explanation, designed through Pedagogical Affordance Analysis, can enhance the learning of conceptual knowledge in early algebra. The results showed that confirmatory diagrammatic self-explanation supported learning of conceptual knowledge among lower-grade students. Also, the post-test performance on tape diagram items showed that diagrammatic self-

explanation helped students, regardless of their grade or prior knowledge, to correctly understand the tape diagram representation. These findings differ from those reported in prior literature on using tape diagrams in algebra, which showed increased performance with diagrams and correct understanding of diagrammatic representations only among higher-grade or higher-ability students (e.g., Booth & Koedinger, 2012; Chu et al., 2017). Regarding students' procedural knowledge, we did not find that diagrammatic self-explanation helped students solve problems more accurately.

Further investigation suggested that our diagrammatic self-explanation with the redesigned tape diagrams is useful when students are transitioning from informal to formal algebraic strategies in equation solving. Overall, pretest performance did not differ between 5th- and 6th-graders, but a detailed coding of students' problem-solving strategies revealed that 5th-graders initially relied on informal strategies to solve equations, whereas 6th-graders relied mostly on algebraic strategies. A follow-up interview with a teacher from the school also confirmed that, prior to the study, the 5th-graders had only been exposed to arithmetic equation solving whereas 6th-graders had consistently been required to use the algebraic strategy to solve equations. In other words, 5th grade students may have a weak understanding of algebraic notation and operations. We also observed the same pattern in Study 1, in which students who were not able to correctly identify errors in worked examples used features of tape diagrams when explaining the errors.

### 3.3 Study 3: Supporting Learning and Performance through Anticipatory Diagrammatic Self-Explanation <sup>2</sup>

#### Design and motivation

Studies 1 and 2 showed that the use of tape diagrams through a scaffolded form of diagrammatic self-explanation helped students acquire conceptual knowledge in early algebra. In the next study phase, I extended the *confirmatory* diagrammatic self-explanation tutor to design a novel, scaffolded way of using visual representations in problem solving: *anticipatory diagrammatic self-explanation* (Nagashima, Bartel, Tseng et al., 2021; Nagashima, Bartel, Yadav et al., 2021; Renkl., 1997).

---

<sup>2</sup> You can read more on Study 3 in Nagashima, T., Bartel, A. N., Yadav, G., Tseng, S., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevin, V. (2021). Using anticipatory diagrammatic self-explanation to support learning and performance in early algebra. In E. de Vries, J. Ahn, & Y. Hod (Eds.), *15th International Conference of the Learning Sciences* (pp. 474–481). International Society of the Learning Sciences

In anticipatory diagrammatic self-explanation, students are asked to *explain* future problem-solving steps by selecting a visual representation that shows a correct and strategic step (Figures 21–23). This contrasts with confirmatory diagrammatic self-explanation, in which students are asked to *explain* the step that they have already solved. Anticipatory diagrammatic self-explanation is informed by the idea of “anticipative reasoning”, proposed by Renkl (1997) as a highly successful self-explanation pattern that they observed in their study.

Anticipatory diagrammatic self-explanation would potentially support both student learning and performance during the learning activity; while it keeps all the core elements in the confirmatory diagrammatic self-explanation (e.g., step-by-step diagrammatic steps), it may support a new cognitive process in its interaction where students may deeply engage with diagrammatic steps to think, “what would be correct and strategic to do next?” In algebra problem solving, such anticipatory self-explanation, in contrast to confirmatory self-explanation (tested in Study 2), may support inference generation about strategic problem-solving steps. If students consider the mathematical symbols as the target representation to learn, engaging in step-level anticipatory self-explanation could help students understand strategic next steps, which would improve both understanding of strategic solution steps and problem-solving performance.

Which diagram shows the current step?

Finish Problem Hint Hint

**Fig. 21.** The ITS starts by asking a learner to select a correct diagram for the given equation. The ITS gives correctness feedback on the learner’s choice of diagram.

Which diagram shows the current step?

Finish Problem Hint Hint

**Fig. 22.** Next, the ITS asks the learner to *explain* (by selecting a diagram) what would be a correct and strategic step to take next. The ITS gives feedback on the choice of diagram.

The screenshot shows a software interface for solving the equation  $2x + 3 = 7$ . At the top right, it says "Please solve for x" and displays the equation  $2x + 3 = 7$ . Below this, there are two diagrams. The first diagram shows a horizontal bar divided into two segments: a blue segment labeled  $2x$  and a yellow segment labeled  $3$ . Below the bar, a bracket indicates the total length is  $7$ . The second diagram shows the same bar, but the yellow segment  $3$  is being moved to the right, and a new yellow segment labeled  $4$  is being added to the left. Below this bar, a bracket indicates the total length is  $4$ . The user has selected the second diagram. At the bottom, there is a "Finish Problem" button and two "Hint" buttons. To the right of the diagrams, there is an input field with an equals sign, suggesting the user should enter the step in symbols.

**Fig. 23.** After selecting a correct and strategic step, the learner enters the step in symbols.

## Method

In Study 3, I investigated the effectiveness of anticipatory diagrammatic self-explanation on student learning and performance. The research question of the study was: *will anticipatory diagrammatic self-explanation, embedded in an ITS, support students' learning and performance?* I hypothesized that the anticipatory diagrammatic self-explanation would promote students' conceptual understanding, enhance procedural skills, and help students learn formal algebraic strategies (H1). I also hypothesized that the anticipatory diagrammatic self-explanation would enhance performance during problem solving in the ITS in that students with the support will perform better on learning process measures while solving symbolic problem-solving steps (e.g., fewer hint requests and fewer incorrect attempts per step), and will solve a similar number of problems as students who do not receive the scaffolded self-explanation support (H2).

## Participants

I conducted a classroom experiment at two private schools in the United States. Participants included 55 6th graders and 54 7th graders across nine class sections taught by four teachers. The experiment was conducted in October 2020, when both schools adopted a hybrid teaching mode in which the majority of students ( $n = 102$ ) attended study sessions in-person and the rest attended remotely ( $n = 7$ ). Teachers were present in the in-person classrooms.

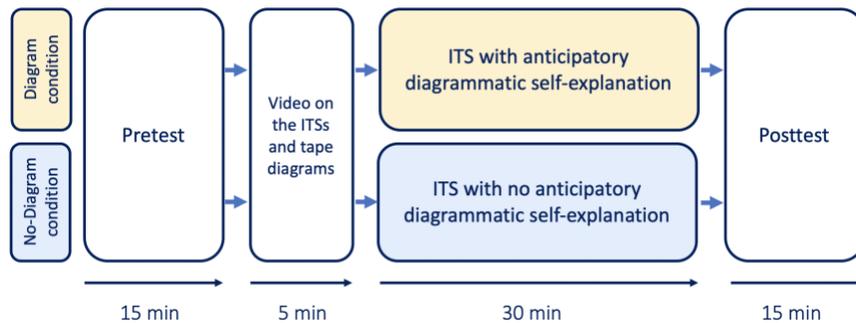
## Test instruments

I developed web-based pretest and posttest assessments (using Qualtrics) to assess students' conceptual and procedural knowledge of basic algebra. The tests contained several items drawn from Study 2 as well as new items we created based on discussions with collaborators at the University of Wisconsin-Madison. The conceptual knowledge items consisted of eight multiple-choice questions and one open-ended question. Also included were four problem-solving items (e.g., “solve for  $x$ :  $3x + 2 = 8$ ”), with two items that were similar to those included in the ITS and two transfer items involving negative numbers. I developed two isomorphic versions of the test that varied only with respect to the specific numbers used in the items; participants received one form as pretest and the other as posttest (with versions counterbalanced across subjects).

### *Procedure*

For each class, the study covered two regular mathematics class periods. The classes were virtually connected to the experimenters and remote learners through a video conferencing system. Students were randomly assigned to either the Diagram condition or the No-Diagram condition. In the Diagram condition, students used the ITS with anticipatory diagrammatic self-explanation. In the No-Diagram condition, students used the ITS with no self-explanation support (the same ITS version I used in Study 2).

On the first day, all students first worked on the web-based pretest for 15 minutes. Then a teacher or the experimenter showed a 5-minute video to all students, which described how to use the ITS and what tape diagrams represent to all students. Next, students practiced equation solving using their randomly-assigned ITS version for approximately 15 minutes. On the second day, students started the class by solving equation problems in the assigned ITS for approximately 15 minutes. After working with the ITS, students took the web-based posttest for 15 minutes. Figure 24 illustrates the study procedure.



**Fig. 24.** Study procedure

## Results

### *Effects on learning outcomes*

One 6th grader was absent for the second day and excluded from the analysis; therefore, I analyzed data from the remaining 108 students, namely, 54 6th-graders (28 Diagram, 26 No-Diagram) and 54 7th-graders (27 Diagram, 27 No-Diagram). Open-ended items were coded for whether student answers were correct or incorrect by two researchers (*Cohen's kappa* = .91).

I first tested hypothesis H1 (*benefits of anticipatory diagrammatic self-explanation with respect to learning outcomes*). I analyzed the data using hierarchical linear modeling (HLM) because the study was conducted in nine classes taught by four teachers at two schools. According to both AIC and BIC, a two-level model showed the best fit, in which students (level 1) were nested in classes (level 2). The inclusion of teachers (level 3) and schools (level 4) did not improve the model fit. I ran two HLMs with posttest scores on CK and PK as dependent variables, type of ITS assigned as the independent variable, and pretest scores (either CK or PK given the dependent variable) as a covariate. For both CK and PK, there was no significant effect of the Diagram/No-Diagram condition (CK:  $t(99.3) = -1.030$ ,  $p = .31$ , PK:  $t(99.4) = -0.292$ ,  $p = .77$ ). I also ran two additional HLMs, regressing pretest-posttest raw gains for CK and PK (dependent variables) on type of ITS. There was a significant gain from pretest to posttest for CK ( $t(108) = 2.778$ ,  $p < .01$ ) but not for PK ( $t(106) = 1.153$ ,  $p = .26$ ), and no significant effect of ITS type. This suggests that students in both ITS conditions improved in conceptual knowledge but not in procedural knowledge.

I then analyzed the strategies that students used to solve the problem-solving items on the pretest and posttest, following the analysis in Study 2 (*Cohen's kappa* = .73, Koedinger et al., 2008). On the pretest, 11 students in the Diagram condition and 17 students in the No-Diagram condition used the Algebraic strategy on one or more problem-solving items. More students did so on the posttest; 26 students in the Diagram condition and 23 students in the No-Diagram condition used the Algebraic strategy. We used McNemar's test to compare the frequency of use of the Algebra strategy at pretest and posttest for each condition. The increase in frequency was significant ( $p < .01$ ) for students in the Diagram condition but was not significant ( $p = .11$ ) for students in the No-Diagram condition. This pattern also held when we limited the analysis to problems involving negative numbers (transfer problems); there was a pretest-posttest increase of only 1 student in the No-Diagram condition, but 12 students in the Diagram condition ( $p < .01$ ). These findings suggest that, although students who learned with anticipatory diagrammatic self-explanation did not have greater gains on tests of conceptual and procedural knowledge, they were more likely to learn the formal algebraic strategy and to apply it to problems with no diagram support, even for problem types that they did not practice in the ITS (H1, partially supported).

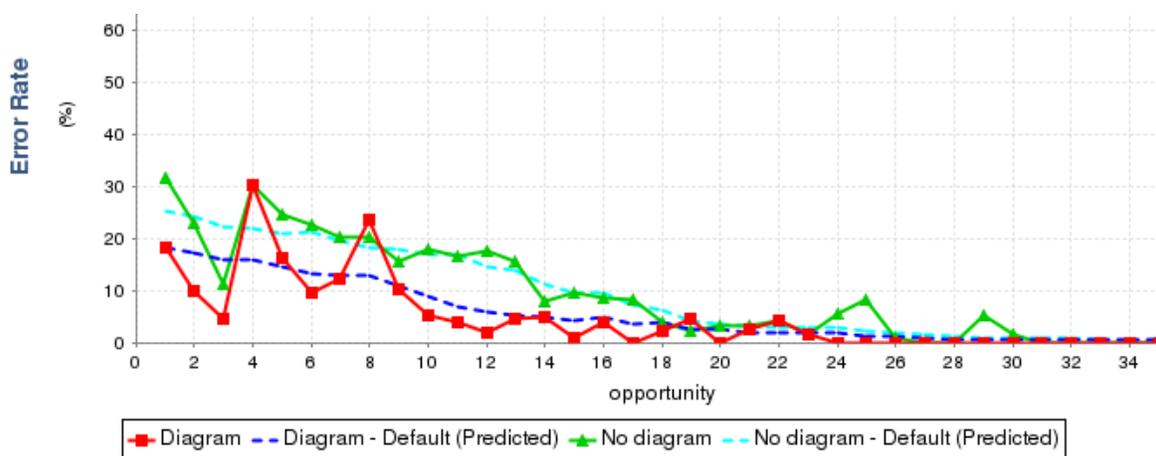
*Effects on learning processes*

Next, I tested hypothesis H2 (*benefits of anticipatory diagrammatic self-explanation with respect to learning processes*), using log data from the ITS. Specifically, I examined “learning curves”, which plot students’ performance within the ITS over time (Rivers et al., 2016). Figure 25 depicts learning curves for the two conditions. The y-axis shows the error rate on steps in tutor problems, averaged across students and skills, and the x-axis shows the sequence of opportunities for practicing each skill. Learning curve analysis assumes that learning occurs when a curve starts with a relatively high initial error rate and gradually goes down as students practice the target skills. The curves are fit to student performance data using the Additive Factors Model (AFM), a specialized form of logistic regression (Rivers et al., 2016). In this study, students practiced a variety of equation-solving skills (e.g., subtracting variable terms). We expected that students who learned with diagrammatic self-explanation support would perform better in the ITS than their peers who did not receive the support (H2). On the symbolic problem-solving steps in the ITS, students in the Diagram condition had a lower error rate than students in the No-Diagram condition. Figure 25 shows learning curves averaged across all symbolic equation-solving skills students in both conditions practiced. Students in the Diagram condition made fewer errors than those in the No-Diagram condition, especially on the earlier opportunities. Both groups improved as they solved more problems (i.e., both curves show a gradual decline), suggesting that, after much practice, the No-Diagram condition eventually lowered their error rate to the same level as the Diagram condition.

In parallel to the trend observed in the learning curves, we found that, when restricting the analysis to symbolic steps only (i.e., excluding diagrammatic self-explanation steps), students who received the self-explanation support trended toward using fewer hints ( $t(89.52) = -1.812, p = .07$ ) and spent significantly less time on each symbolic problem-solving step ( $t(99.51) = -2.238, p = .03$ ) than students who did not receive the self-explanation support (Table 10). The average number of problems solved in the ITS during the (fixed amount of) available time did not differ significantly across conditions, ( $t(99.30) = -0.528, p = .60$ ) (Table 10).

**Table 10.** Average number of problems solved, number of incorrect attempts, number of hint requests, and average time spent on symbolic steps in the ITS (standard deviation).

	Average number of problems solved	Average number of hints requested per step	Average time spent per step
Diagram	15.40 (9.02)	0.68 (0.96)	15.99 (9.91)
No-Diagram	16.17 (11.16)	1.02 (1.38)	20.27 (13.95)



**Fig. 25.** Learning curves for the Diagram condition (red) and the No-Diagram condition (green) averaged across the skills students practiced during the symbolic problem-solving steps. Dark and light blue lines show predicted curves based on the AFM (dark blue: Diagram condition, light blue: No-Diagram condition).

In summary, the students in both conditions practiced a similar number of problems in the ITS in a similar amount of time overall, and the anticipatory diagrammatic self-explanation helped students spend less time and ask for fewer hints on symbolic steps (H2, partially supported). In addition, the learning curves indicate that students in both conditions learned equation-solving skills eventually, but the students in the Diagram condition had a smoother experience with fewer errors.

## Discussion

We found that anticipatory diagrammatic self-explanation embedded in an Intelligent Tutoring System (ITS) helped students learn to apply a formal, algebraic problem-solving strategy to problems

outside the ITS and to transfer problems involving negative numbers (H1). Anticipatory diagrammatic self-explanation also supported student performance within the ITS, measured by lower learning curves, less frequent use of hints, and less time spent on each symbolic equation-solving step (H2). Anticipatory self-explanation did not lead to differences in posttest scores, contrary to H1, but it helped students learn more efficiently; students learned the formal algebraic strategy while solving a similar number of problems with less time and fewer errors and hint requests, and they achieved similar gains on conceptual knowledge (H2).

I attribute these findings to the design and learning principles used in supporting anticipatory diagrammatic self-explanation. Specifically, I think that the process of selecting the next correct-and-strategic problem-solving step, depicted diagrammatically, helped students perform better and faster on the corresponding step with symbols. On steps with symbols, students had a diagrammatic representation of the step available to them on the screen. They could refer to this representation as they sought to express the step using mathematical symbols. Engaging in this cognitive process may have helped students understand step-level formal strategies in a visual form (e.g., visually seeing that constant terms are taken out from both sides of an equation). Comparing and contrasting the different tape diagrams may have supported students in selecting steps that were both correct and strategic, and it may have helped them avoid using informal strategies, such as guessing. It may be, as well, that the better performance resulting from the anticipatory diagrams gave students a bit more confidence to take on the challenge of moving towards formal algebra.

### **3.4 Study 4: Testing the Effectiveness of Anticipatory Diagrammatic Self-Explanation on Students with Different Prior Knowledge Levels<sup>3</sup>**

#### **Motivation**

Using the anticipatory diagrammatic self-explanation tutor, I conducted another study 1) to investigate whether the benefits of anticipatory diagrammatic self-explanation we found in Study 3 would be influenced by students' prior knowledge and 2) to further examine how diagrams scaffold students during anticipatory diagrammatic self-explanation.

---

<sup>3</sup> You can read more on Study 4 in Nagashima, T., Bartel, A. N., Tseng, S., Vest, N.A., Silla, E. M., Alibali, M. W., & Aleven, V. (2021). Scaffolded self-explanation with visual representations promotes efficient learning in early algebra. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society* (pp. 1858-1864). Cognitive Science Society.

## Method

### *Participants*

I conducted a classroom experiment at a single private school in the United States. Participants included 30 5th graders, 17 6th graders, 23 7th graders, and 21 8th graders (total  $n = 91$ )<sup>4</sup>. These students were taught by two teachers across seven class sections. We conducted the experiment in October 2020 when the school was operating under a hybrid teaching mode due to the COVID-19 pandemic. In this hybrid mode, the majority of students ( $n = 89$ ) attended the study in-person while two students attended remotely. Both participating teachers noted that some students had seen tape diagrams in learning materials, but the instruction had never focused specifically on tape diagrams.

### *Procedure*

The study took place during two regular mathematics class periods, in which most of the students and the teacher were present “live” in the actual classroom, and in which experimenters and remote learners joined through a video conferencing system. Students in each class were randomly assigned to either the Diagram condition or the No-Diagram condition. The study procedure was the same as that of Study 3 (Figure 17).

## Results

### *Effects on learning outcomes*

Of the 91 participants, we excluded five students who did not complete the posttest and two other students who were at ceiling at the pretest. Therefore, the following analyses focus on remaining 84 students (26 5th graders, 16 6th graders, 23 7th graders, and 21 8th graders), of whom 41 were in the Diagram condition and 43 were in the No-Diagram condition.

Two researchers separately evaluated all answers for the open-ended items on the pretest and posttest (840 student answers) and achieved high inter-rater reliability (*Cohen's kappa* = .81). Table 11 shows pretest and posttest scores on the conceptual knowledge (CK, max: 9) and procedural knowledge (PK, max: 4) items. One-way repeated measures ANOVAs showed a significant pretest-posttest gain across conditions on the procedural knowledge items ( $F(1, 83) = 12.88, p < .01$ ) and positive but non-significant pretest-posttest gain for the conceptual knowledge items ( $F(1, 83) = 2.86, p = .09$ ). I then conducted two separate linear regressions, one with conceptual knowledge posttest scores and one with procedural knowledge posttest scores as dependent variables. In both models, condition (Diagram or No-Diagram), prior knowledge pretest score (combined CK and PK

---

<sup>4</sup> Data of 6<sup>th</sup> and 7<sup>th</sup> graders from this sample was also included in the analyses for Study 3

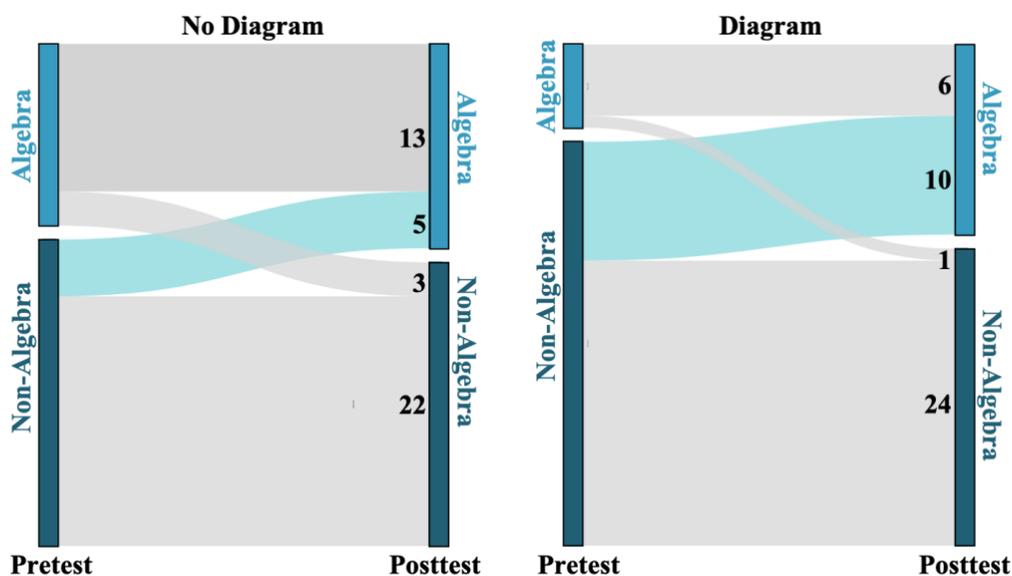
scores), and the interaction between the two served as predictors. Additionally, the number of problems solved in the ITS and grade level were included as covariates. Grade level was treated as a continuous variable, with 5th, 6th, 7th, and 8th grade coded as -1.5, -.5, .5, 1.5, respectively. In both models, there was no significant main effect of condition (CK:  $\beta = -0.20$ ,  $t(78) = -0.28$ ,  $p = .78$ , PK:  $\beta = -0.28$ ,  $t(78) = -0.60$ ,  $p = .55$ ) and no significant interaction of condition and pretest scores (CK:  $\beta = .09$ ,  $t(78) = .76$ ,  $p = .45$ , PK:  $\beta = .03$ ,  $t(79) = .39$ ,  $p = .69$ ).

**Table 11.** Mean scores on the pretest and posttest (standard deviations).

Condition	Pretest		Posttest	
	CK	PK	CK	PK
Diagram	3.63 (1.93)	1.15 (1.35)	4.10 (2.37)	1.56 (1.52)
No-Diagram	4.30 (2.29)	1.63 (1.57)	4.47 (2.53)	2.19 (1.82)

Also, to examine whether engaging with anticipatory diagrammatic self-explanation influenced students' use of formal problem-solving strategies, for the procedural pretest and posttest items, two researchers coded for whether students used the formal algebraic strategy in solving equations (Koedinger et al., 2008). If an answer used algebraic manipulations to reach the solution, it was coded as "Algebra" strategy. Otherwise, we coded it as "Non-Algebra" strategy. Two researchers coded all 672 student answers (*Cohen's kappa* = .64).

There was no difference between the conditions in the number of students who used the Algebra strategy at least once on the procedural items neither on the pretest (Diagram: 12 out of 41, No-Diagram: 20 out of 43,  $\chi^2[1, n = 84] = 2.65$ ,  $p = .10$ ) nor on the posttest (Diagram: 16 out of 41, No-Diagram: 22 out of 43,  $\chi^2[1, n = 84] = 1.25$ ,  $p = .26$ ). However, for the two transfer items, I found a significant increase in the number of students using the Algebra strategy from pretest to posttest for the Diagram condition (from 7 to 16,  $p = .01$ ), but not for the No-Diagram condition (from 16 to 18,  $p = .72$ : Figure 26).



**Fig. 26.** The change, by condition (left: No-Diagram condition; right: Diagram condition), from pretest to posttest, in the number of students who used the Algebra strategy

#### *Effects on learning processes*

To investigate students' performance in the ITS, I analyzed log data collected by the ITS. Specifically, I explored the total number of problems solved, the average number of incorrect attempts at each problem-solving step, the average number of hints requested at each step, and the time spent on each step (Table 12). I only compared the process measures on the symbolic steps, excluding the transactions for the diagrammatic steps, to make fair comparisons between the conditions.

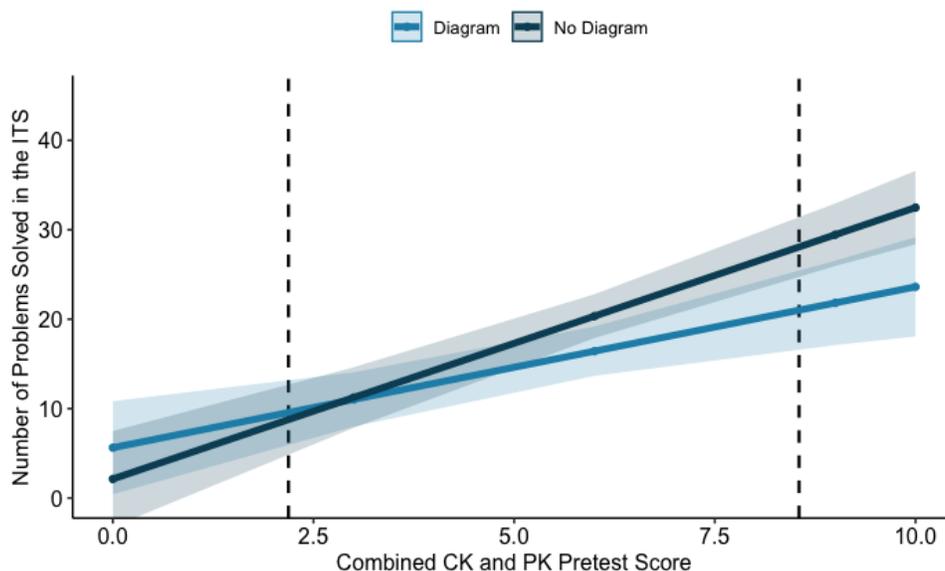
**Table 12.** The means and standard deviations (in parentheses) of the process measures.

Condition	Number of problems solved	Average number of incorrect attempts per step	Average number of hints requested per step	Average time spent per step
Diagram	14.22 (8.47)	1.07 (2.03)	0.31 (0.52)	18.42 (13.02)
No-Diagram	20.14 (13.94)	1.09 (1.57)	0.56 (0.72)	21.30 (17.15)

To examine whether learners in the Diagram condition showed *efficient* learning, we ran four separate linear regressions with each of the process measures as a dependent variable. In all four models, condition, pretest score, and their interaction were included as independent variables. Additionally, grade level was included as a covariate. Also, we added the number of problems solved as a covariate to three of the four models (the ones in which it was not the dependent variable) because the number of problems solved was strongly/moderately correlated with each of the three other dependent variables.

First, we found a main effect of pretest scores on the number of problems solved,  $\beta = 3.04$ ,  $t(79) = 7.49$ ,  $p < .01$ , indicating that as prior knowledge increases, students solved more problems in the ITS. This increase was steeper for students in the No-Diagram condition than the Diagram condition,  $\beta = -1.24$ ,  $t(79) = -2.12$ ,  $p = .04$  (Figure 27). We then tested simple main effects of condition at one standard deviation below the mean for combined pretest scores and one standard deviation above the mean for combined pretest scores (see dotted vertical lines in Figure 20). Results showed that among those who scored above average on the pretest, learners in the No-Diagram condition solved significantly more problems than those in the Diagram condition.  $\beta = 3.53$ ,  $t(79) = 2.69$ ,  $p < .01$ . However, there was no difference in the number of problems solved between conditions for learners who scored below average on the pretest,  $\beta = -0.40$ ,  $t(79) = -0.31$ ,  $p = .76$ .

Regarding hint use and average time spent per step, we found a significant main effect of condition (hint use:  $\beta = -0.71$ ,  $t(78) = -3.08$ ,  $p < .01$ ; time per step:  $\beta = -12.18$ ,  $t(78) = -2.89$ ,  $p < .01$ ) but no significant interactions between condition and pretest score. There were no significant main nor interaction effects on the average number of incorrect attempts made per step. These results indicate that anticipatory diagrammatic self-explanation helped learners spend less time and request fewer hints on symbolic steps than learners with no self-explanation support, but it did not help them make fewer errors on symbolic steps.



**Fig. 27.** An interaction between condition and pretest score on the number of problems solved. The slope is steeper for the No-Diagram condition than the Diagram condition. The two dots lines indicate our two tests of simple main effects; one standard deviation below and one standard above the mean for combined pretest score.

To uncover *how* the anticipatory diagrammatic self-explanation scaffolded student performance, I examined relations between performance on the diagrammatic steps and the symbolic steps using ITS log data from the participants in the Diagram condition ( $n = 41$ ). I tested if any of the performance measures for diagrammatic steps predicted learners' performance on symbolic steps. I ran three additional linear regressions with the same set of predictors of primary interest: pretest scores, the average number of incorrect attempts for each diagrammatic step, and the average time spent for each diagrammatic step. I did not include the average number of hints requested since only one student used hints for diagrammatic steps. I included grade level and the number of problems solved as covariates in order to keep the models consistent with other models presented earlier. The dependent variables for the three models were the average number of incorrect attempts for each symbolic step, the average time spent for each symbolic step, and the average number of hints requested for each symbolic step. When controlling for these other variables, the average number of incorrect attempts on diagram steps significantly predicted more incorrect attempts on symbolic steps

( $\beta = 6.17$ ,  $t(35) = 2.50$ ,  $p = .02$ ) and more time spent on symbolic steps ( $\beta = 25.64$ ,  $t(35) = 2.34$ ,  $p = .03$ ). There was also a significant association between more incorrect attempts on diagrammatic steps and lower hint use on symbolic steps ( $\beta = -1.46$ ,  $t(35) = -2.32$ ,  $p = .03$ ).

## Discussion

Findings from this study indicate that, regardless of their prior knowledge, anticipatory diagrammatic self-explanation helped learners solve symbolic steps faster and ask for fewer hints within the ITS, and supported them in the transition from informal strategies to the formal algebra strategy use for transfer problems with negative numbers on the procedural items in the pretest and posttest. Also, despite the additional diagrammatic steps, which almost doubled the number of steps for each problem, there was no difference, for students with lower prior knowledge, in the number of problems solved between those who received anticipatory diagrammatic self-explanation and those who did not. By contrast, for learners with higher prior knowledge, diagrammatic steps led to fewer problems solved, suggesting that the diagrams introduced additional workload. Still, the results suggest that learners can use inferential activity with tape diagrams to guide their symbolic problem solving. Learners with lower prior knowledge may have used the scaffolding to help with selecting strategic problem-solving steps. For those with higher prior knowledge, although the new representation may have largely captured what they already knew how to do and may not have scaffolded them to solve more problems, it still helped them process symbolic steps faster with fewer hints.

How did anticipatory diagrammatic self-explanation support learning and performance? The analysis revealed that making more incorrect attempts during anticipatory diagrammatic self-explanation was associated with more time spent and more incorrect attempts made on the symbolic steps, even after controlling for prior knowledge. However, making incorrect diagram selections was also associated with fewer hint requests on the symbolic steps. These results suggest that, although students who make errors on the diagrammatic steps tend to make more errors and spend more time on symbolic steps, anticipatory diagrammatic self-explanation also serves as a guiding step that learners could use when entering the next symbolic step. Making incorrect diagrammatic self-explanations and receiving feedback on their incorrect attempts may allow learners to reflect on their selection deeply, rather than processing the multiple-choice diagrammatic step shallowly, leading to fewer hint requests made on the symbolic steps. However, we also acknowledge that the observed relation between diagrammatic steps and symbolic steps might be a manifestation of a behavior known as “gaming the system” (Baker et al., 2008). That is, the multiple-choice diagrams with

feedback may have invited quick guessing and therefore students may not have fully engaged with diagrams and the ITS.

I acknowledge several limitations of the study. A sample size of 84 participants is not large. The additional analysis for the participants in the Diagram condition was performed with an even smaller subset of the data. Therefore, the findings from this study might not correctly reveal causal and predictive relationships. Also, the study was conducted in a private school in the United States. Future studies are necessary to investigate the impact of the intervention with students at other types of schools and in other locations.

### **3.5 Study 5: Reducing Scaffolding in Anticipatory Diagrammatic Self-Explanation through Interleaving Practice**

#### **Motivation**

Studies 3 and 4 have shown consistent findings that anticipatory diagrammatic self-explanation helps students' learning of the formal problem-solving strategy while supporting their problem-solving performance in the tutor. These findings made me think of an open, important question of *how much visual scaffolding is appropriate to support student learning and performance*. While providing visual support may enhance problem-solving performance when the scaffold is available, it might only engage shallow processing of the target knowledge as learners could easily rely on the provided visual rather than connecting the visual with the target content meaningfully (Zhang & Fiorella, 2021). Of particular interest is the fading of visual scaffolding (i.e., “when should learners solve problems with the visual scaffolding and when to solve problems with no visual support?”) (Belland et al., 2017). Research has shown that fading scaffolds, such as interleaved practice, can support robust learning (Rohrer et al., 2015). To date, however, few studies have been conducted on the topic of the fading of visual scaffolding (Rau et al., 2010, 2013). There is a lack of thorough understanding of the effect of sustaining and fading visual support and how learners interact with visual scaffolding. In-depth investigations into the cognitive processes involved in problem solving with visual scaffolding would help understand the effect more.

In the context of early algebra, sustaining or fading visual scaffolding presents an important question both for the scientific and practical communities. To date, research on visual representations for algebra has only focused on either providing visual support all the time or not providing it at all when investigating the effect of learning with visual representations, including my own studies (Booth & Koedinger, 2010). This *all-or-nothing* contrast does not reflect how classroom teaching is

conducted; many mathematics textbooks mix problems with visuals and without visuals in teaching algebra (Fukuda et al., 2021). Examining how sustaining and fading affects learning and problem-solving performance will provide important scientific knowledge regarding whether providing visual support all the time might or might not over-scaffold learning. For instance, always providing tape diagrams may effectively foster conceptual and strategic understanding of problem-solving procedures because students can connect the visual information depicted in a tape diagram with symbolic representations in an equation. On the other hand, always using such scaffolding in solving equation problems also contains a risk of students' over-reliance on tape diagrams when solving symbolic equations (Booth & Koedinger, 2012). That is, students' learning might get focused on rather superficial diagram-to-symbols translation knowledge that might not help to acquire deeper knowledge of using visual representations to strategically solve symbolic equations. In Study 5, I was interested in whether sustaining visual scaffolding to support problem solving, previously shown to be beneficial in my studies, would *over-scaffold* student learning by comparing two instructional strategies: interleaving between problems with diagrams and problems without diagrams and providing diagrams for all problem-solving opportunities. Specifically, I examined the following research questions<sup>5</sup>:

- *RQ1: Does providing visual scaffolding at every problem-solving opportunity during algebra problem solving over-scaffold student learning, compared to interleaving the visual scaffolding?*
- *RQ2: Does providing visual scaffolding at every problem-solving opportunity during algebra problem solving support efficient problem-solving performance in the ITS, compared to interleaving the visual scaffolding?*
- *RQ3: How does the visual support influence students' performance in the ITS?*

RQ1 addresses a fundamental question of whether sustaining the visual scaffolding may or may not over-scaffold learners in gaining conceptual and procedural knowledge. Literature on interleaved practice argues that interleaving different problem types may support learning because it requires extra cognitive effort from learners (e.g., choose different problem-solving strategies) (Rohrer et al., 2015). Such extra cognitive effort might help students focus on connecting visual information with the information in the symbolic representation. In the context of anticipatory diagrammatic

---

<sup>5</sup> To read more about this study, please see: Nagashima, T., Ling, E., Zheng, B., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2022). How does sustaining and interleaving visual scaffolding help learners? A classroom study with an Intelligent Tutoring System. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*. Cognitive Science Society (pp. 1751-1758).

explanation, interleaving visual scaffolding might help foster deeper thinking on the part of students about problem-solving procedures because students would have to think about future steps on their own without visual support for problems without visual scaffolding. This extra effort could result in enhanced conceptual and procedural knowledge. On the other hand, when students only receive problems in which the visual scaffolding is provided, they might only engage with the shallow processing of the content (i.e., copying what is shown in the tape diagrams to the symbolic problem solving without deeply engaging with its conceptual and procedural meanings). Therefore, I hypothesized:

- *Hypothesis 1. Students who received interleaved visual scaffolding will gain more conceptual and procedural knowledge from pretest to posttest, compared to students who received the visual scaffolding all the time (i.e., sustaining the visual support over-scaffolds learners)*

RQ2 asks how sustaining versus interleaving visual scaffolding might affect students' problem-solving performance during a learning activity (in the ITS). This is measured by performance measures, including time spent and accuracy in problem solving (Long & Alevan, 2014). My hypothesis for RQ2 was:

- *Hypothesis 2. Students who received the sustained visual scaffolding will solve problems in the ITS more efficiently compared to students who received the visual scaffolding in an interleaved way.*

I examined RQ3 to uncover, through several different analytical approaches, how students interact with the visual scaffolding during problem solving, which is an underexplored area of research on the use of visual representations in algebra. I specifically compared students' performance across conditions 1) on problems in which all students, regardless of the visual support frequency, received with the visual support and 2) on problems in which students with the interleaved practice received no visual support whereas students who were given the visual scaffolding all the time did. By comparing and contrasting students' performance on these two types of problem events, we can have a better understanding of the scaffolding effects in the ITS. Specifically, I investigated where any observed differences between the conditions (if any) would come from. Therefore:

- *Hypothesis 3.1. Students in both conditions will show similar within-ITS performance on problems with the visual support, regardless of the assigned condition, because students in both conditions receive the same scaffolding.*

- *Hypothesis 3.2. Students in the Interleaved condition will perform worse on problem-solving items in the ITS than those in the All-Diagram condition on problems where only students in the All-Diagram condition receive the scaffolding.*

Lastly, using a Knowledge Component modeling approach (a standard technique used in the field of educational data mining) (Long et al., 2018; Nguyen et al., 2019), I investigated 3) to what extent students use overlapping vs. separate knowledge on symbolic steps with diagrams and without diagrams. By labeling Knowledge Components, or fine-grained problem-solving skills in an intelligent tutor (Koedinger et al., 2012) differently for solving problems with the visual scaffolding and solving problems without the visual scaffolding, we will examine if students' actual performance can be modeled better with such a separation of knowledge. I hypothesized:

- *Hypothesis 3.3. A Knowledge Component model that includes the effect of visual scaffolding shows a better fit, assuming that performance with the visual scaffolding is better than that without the visual scaffolding.*

## **Method**

### *Participants*

I conducted a classroom experiment at a public middle school in the United States. The school is the only middle school in the school district in which over 65% of its students come from low-income families. At the participating school, 44.7% of students were considered at the “below basic” level, based on the state’s standardized assessment results in 2019. Participants in the study included 77 7th-grade students. These students were taught in five class sections by one teacher. I conducted the experiment in May 2021 when the school was operating under a hybrid teaching mode due to the COVID-19 pandemic, in which 30 students participated remotely from their own home environment and the remaining 47 joined from their classroom with their teacher. The participating teacher noted that students’ prior exposure to tape diagrams was minimal. Among the 77 students, 16 students were in the Individualized Education Program (i.e., IEP).

### *Materials*

Web-based pretest and posttest, modified based on those used in Study 3 and 4, were used in the study. Test items included six conceptual knowledge items (CK) and seven procedural knowledge items (PK). The procedural knowledge items consisted of four items with no tape diagrams (PK-NoDiagram) and three problems that show a corresponding tape diagram (PK-Diagram). The three problems with tape diagrams and three of the four problems without tape diagrams shared the same

equation types, and the fourth problem without tape diagrams was a transfer problem that used negative numbers. Two isomorphic versions were created and assigned to students in a counter-balanced way.

The ITS with anticipatory diagrammatic self-explanation was used in the study. As illustrated in Figure 28, students in the All-Diagram condition used the ITS that provided anticipatory diagrammatic self-explanation support for all problems whereas those in the Interleaved condition used it only for *odd-numbered* problems (i.e., first problem, third problem, fifth problem, and so on). For *even-numbered* problems, students in the Interleaved condition received a version with no diagrammatic steps available. These two ITS versions differ only in whether the ITS provides diagrammatic steps or not; in other words, the only difference between the two conditions was that, for even-numbered problems, students in the All-Diagram condition solved problems with diagrams, but those in the Interleaved condition solved without diagrams. Students, regardless of their assigned condition, received the same list of problems in a fixed order which increases problem difficulty as students move on.

Problem #	Interleaved	All-Diagram	Equation Type
Problem 1			$x + a = b$
Problem 2			$x + a = b$
Problem 3			$a + x = b$
Problem 4			$a + x = b$
⋮			⋮
Problem 9			$ax + b = c$
Problem 10			$ax + b = c$

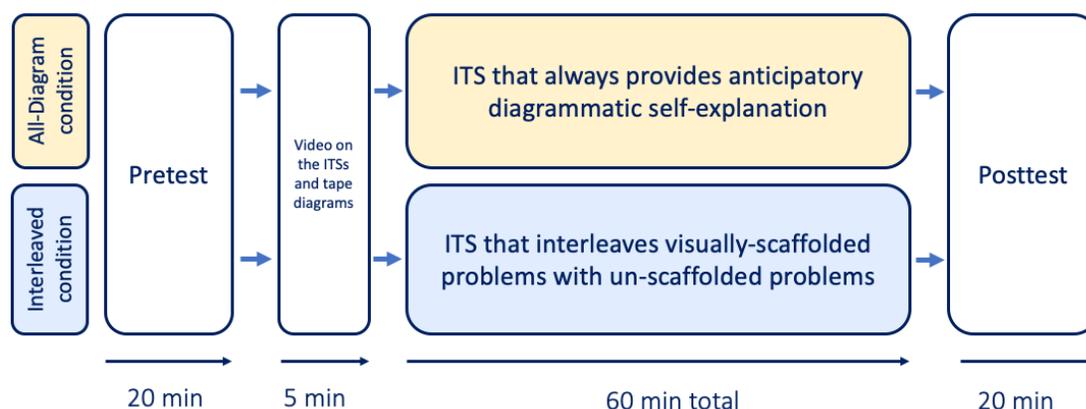
**Fig. 28.** How we sustained and interleaved visual scaffolding in the experiment. In the Interleaved condition, students received the visual scaffolding only on odd-numbered problems.

### Procedure

The study took place during five regular mathematics class periods, in which approximately half of the students and the teacher were present live in the actual classroom, and experimenters and remote learners joined through a video conferencing system. Students in each class were randomly assigned to either the All-Diagram condition or the Interleaved condition. Sixteen students in IEP were pre-identified and separately randomly assigned to the groups among them. Also, based on teacher-reported information regarding students' regular participation mode (i.e., remote or in-person class participation), we randomly assigned students to conditions separately among students joining remotely and those joining from the classroom.

In the first class session, students started by working on the pretest for 20 minutes. Then the experimenter showed students in both conditions a five-minute video describing how to use the ITS

and what tape diagrams represent. Starting in the second class period, students spent 15-20 minutes using the assigned ITS version to practice algebra problem solving. (from the 2nd to 4th periods, the total ITS learning time in both conditions was approximately 60 minutes). On the final day, students took the web-based posttest for 20 minutes. Students were given access to both ITS versions about a week after the study. Figure 29 illustrates the study procedure.



**Fig. 29.** Study procedure

## Results

Of the 77 participants who completed the pretest, we excluded 14 students who did not complete either the tutor learning activity, the posttest, or both. I think that this high attrition was due to the fact that the school was conducting the hybrid instruction (Nagashima, Yadav et al., 2021). The following analyses focus on the remaining 63 students, of which 32 were in the Interleaved condition and 31 were in the All-Diagram condition.

### *Effects on learning outcomes*

Table 13 shows students' pretest and posttest scores. To test Hypothesis 1 (i.e., *students who receive interleaved visual scaffolding will gain more conceptual and procedural knowledge, compared to students who receive the visual scaffolding all the time*), I conducted three separate linear regressions,

with conceptual knowledge posttest score (CK), posttest score on procedural items with tape diagrams (PK-Diagram), and posttest score on procedural items without tape diagrams (PK-NoDiagram) as dependent variables, respectively. In all of the models, condition (All-Diagram or Interleaved, coded as 0 or 1) and prior knowledge (i.e., pretest scores) served as predictors. We found no significant main effect of condition for conceptual learning ( $\beta = -0.42$ ,  $t(62) = -1.22$ ,  $p = .23$ ), procedural learning on items with diagrams ( $\beta = -0.17$ ,  $t(62) = -0.77$ ,  $p = .44$ ), and procedural learning on items without diagrams ( $\beta = -0.29$ ,  $t(62) = -1.12$ ,  $p = .27$ ). This finding shows that Hypothesis 1 was not supported; contrary to our expectation, students who solved problems using the tutor that gives visual scaffolding for all problem-solving opportunities did not over-scaffold learning.

**Table 13.** Pretest and posttest score (standard deviation) by condition.

Condition	Pretest			Posttest		
	CK (max: 4)	PK- Diagram (max: 3)	PK- NoDiagram (max: 4)	CK (max: 4)	PK- Diagram (max: 3)	PK- NoDiagram (max: 4)
All-Diagram	2.82 (1.53)	1.27 (1.15)	2.18 (1.42)	3.30 (1.55)	1.45 (1.25)	2.21 (1.45)
Interleaved	2.78 (1.39)	1.03 (1.06)	1.84 (1.22)	2.75 (1.50)	1.12 (1.13)	1.72 (1.45)

#### *Effects on learning processes*

To address Hypothesis 2 (i.e., *students who receive the visual scaffolding all the time will solve problems in the ITS more efficiently compared to students who receive the visual scaffolding in an interleaved way*), I ran four separate linear regressions with four problem-solving performance measures (the number of problems solved, average number of hints per symbolic step, average number of incorrect steps per symbolic step, and average time spent per symbolic step) as dependent variables. I only compared students' performance on symbolic steps and excluded interactions with the diagram steps from the log data. In all four models, condition and pretest score were included as independent variables. Table 14 shows descriptive data on the performance measures.

**Table 14.** Average performance measures per symbolic step (standard deviation).

Condition	Average number of hint requests	Average number of incorrect attempts	Average time spent
All-Diagram	0.16 (0.23)	0.45 (0.63)	6.85 (2.70)
Interleaved	0.53 (0.64)	0.86 (0.92)	12.0 (6.92)

First, I did not find a main effect of condition on the number of problems solved ( $\beta = -3.44$ ,  $t(60) = -1.05$ ,  $p = .30$ ). In other words, students solved a similar number of problems in the tutor, regardless of condition. In the models for the average number of hints per symbolic step, the average number of incorrect attempts per symbolic step, and time spent on symbolic steps, I found a significant main effect of condition on the average number of hint requests per step ( $\beta = 0.28$ ,  $t(60) = 2.83$ ,  $p < .01$ ) and average time spent per step ( $\beta = 3.32$ ,  $t(60) = 2.21$ ,  $p = .03$ ). I did not find a main effect of condition on the number of incorrect attempts made per step,  $\beta = 0.31$ ,  $t(60) = 1.67$ ,  $p = .10$ . Altogether, students in the All-Diagram condition solved problems with fewer hint requests and less time on symbolic steps, suggesting that students in the All-Diagram condition solved symbolic steps more efficiently than those in the Interleaved condition. These results show that students who received the sustained visual scaffolding performed better on symbolic problem-solving measures such as hint use and time spent on symbolic steps. These findings partially support Hypothesis 2.

#### *Diagram's scaffolding effects*

To address Hypotheses 3.1 – 3.3 (i.e., *students who receive the visual scaffolding all the time will solve problems in the ITS more efficiently compared to students who receive the visual scaffolding in an interleaved way*), I ran four separate linear regressions with four problem-solving performance measures (the number of problems solved, average number of hints per symbolic step, average number of incorrect steps per symbolic step, and average time spent per symbolic step) as dependent variables. I only compared students' performance on symbolic steps and excluded interactions with the diagram steps from the log data. In all four models, condition and pretest score were included as independent variables. Table 15 shows descriptive data on the performance measures.

**Table 15.** Performance measures on odd-numbered problems and even-numbered problems by condition (standard deviation).

Condition	Odd-numbered problems			Even-numbered problems		
	Hint requests	Incorrect attempts	Time spent	Hint requests	Incorrect attempts	Time spent
All-Diagram	0.23 (0.31)	0.70 (1.02)	15.1 (9.15)	0.08 (0.19)	0.17 (0.29)	7.51 (5.14)
Interleaved	0.62 (0.75)	0.99 (1.01)	21.4 (16.0)	0.37 (0.55)	0.49 (0.50)	14.6 (10.4)

I first compared students' performance on odd-numbered problems to examine if students, regardless of the condition, performed similarly on scaffolded problems. I ran three separate linear regressions with the number of hints used per symbolic step, the number of incorrect attempts per symbolic step, and time spent per symbolic step as dependent variables, and condition and pretest scores as predictors. I found that students in the All-Diagram condition used significantly fewer hints ( $\beta = 0.29$ ,  $t(60) = 2.48$ ,  $p = .02$ ) and trended towards spending less time ( $\beta = 5.87$ ,  $t(60) = 1.89$ ,  $p = .06$ ). No significant difference was found on the number of incorrect attempts per symbolic step,  $\beta = 0.19$ ,  $t(60) = 0.78$ ,  $p = .44$ . Therefore, Hypothesis 3.1 was not supported; students in the All-Diagram condition performed better on problems in which students in both conditions received the same scaffolding.

Then, I compared students' performance on even-numbered problems to understand whether students in the Interleaved condition performed poorly on problems with no visual scaffolding (on even-numbered problems, only students in the All Diagrams condition had visual scaffolding; students in the Interleaved condition worked with the symbolic representation only; see Figure 21). The same set of regression models revealed that students in the All-Diagram condition requested significantly fewer hints ( $\beta = 0.24$ ,  $t(60) = 2.67$ ,  $p = .01$ ), made significantly fewer incorrect attempts ( $\beta = 0.30$ ,  $t(60) = 2.85$ ,  $p = .01$ ), and spent significantly less time ( $\beta = 6.64$ ,  $t(60) = 3.31$ ,  $p = .01$ ) on symbolic steps. This finding indicates that students in the All-Diagram condition did better on problems in which only those students in the All-Diagram condition received the scaffolding (Hypothesis 3.2, supported).

#### *Knowledge component modeling*

Finally, for H3.3, we conducted Knowledge Component modeling to investigate potential mechanisms that may have influenced the observed differences between the conditions. A

Knowledge Component (KC) is defined as “an acquired unit of cognitive function or structure that can be inferred from performance on a set of related tasks” (Koedinger et al., 2012). Studies on ITSs have used Knowledge Component modeling (i.e., modeling student’s knowledge state and growth based on student’s performance on a set of KCs) to design and improve instruction in the software (Huang et al., 2021). KC models use a specialized form of logistic regression known as Additive Factors Models (Rivers et al., 2016). Improving KC models is critical for better understanding student learning and performance, and for better designing instructional support in intelligent software. Model fit can be evaluated by three metrics, namely, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and 3-fold cross validation metrics reported as root mean squared error (RMSE), where lower values suggest better model fit (Nguyen et al., 2019). We applied the KC modeling approach to our dataset to investigate whether the scaffolding effect of having diagrams can be manifested in the model fit. To conduct the KC modeling analysis, we used LearnSphere’s DataShop (<https://pslcdatashop.web.cmu.edu/>).

In our ITS log data, the original KC model had general algebra problem-solving KCs (see Table 14). To see if the visual scaffolding effect would be manifested in KCs, we created an additional set of KCs that treats the various skills involved in “solving equations with diagrams” as separate skills, depending on whether the problems had visual scaffolding or not (Table 16). We compared the original KC model with an updated model that considers “solving equations in symbols without diagrams” and “solving equations with diagrams,” only for the Interleaved condition (because the All-Diagram condition had diagrams for all the problem-solving opportunities). Applying the KC modeling approach allows us to see whether students in the Interleaved condition practiced separate sets of skills in the tutor. We found that the updated model improved the model fit on AIC and all the RMSE values (but not for BIC, see Table 4), which suggests that treating problem-solving skills with and without diagrams as distinct better represents the actual student behavior (Stamper et al., 2013). This suggests that students were solving equations using different skills between problems with anticipatory diagrammatic self-explanation and problems without visual support (Hypothesis 3.3, supported).

**Table 16.** Lists of Knowledge Components for the original and updated models. The updated model treats KCs differently for problems with diagram and without diagrams and therefore has additional skills.

KC type	Original model	Updated model	Description
KCs for the symbolic problem-solving steps	cancel-const	cancel-const	Canceling out constant terms ( $ax + \mathbf{b} - \mathbf{b} = c - b$ )
		cancel-const-d	Canceling out constant terms when diagrams are present
	cancel-var	cancel-var	Canceling out variable terms ( $ax - cx = \mathbf{cx} - \mathbf{cx} + b$ )
		cancel-var-d	Canceling out variable terms when diagrams are present
	combine-like-const	combine-like-const	Combining like terms (constant terms, $ax + b - b = c - \mathbf{b}$ )
		combine-like-const-d	Combining like terms (constant terms) when diagrams are present
	combine-like-var	combine-like-var	Combining like terms (variable terms, $\mathbf{ax} - \mathbf{cx} = cx - cx + b$ )
		combine-like-var-d	Combining like terms (variable terms) when diagrams are present
	divide	divide	Dividing an equation by variable coefficient ( $ax$ )
		divide-d	Dividing an equation by variable coefficient when diagrams are present
	division-complex	division-complex	Dividing an “ $ax + b = cx$ ” equation by “ $a$ ”
		division-complex-d	Dividing an “ $ax + b = cx$ ” equation by “ $a$ ” when diagrams are present
	division-simple	division-simple	Dividing an “ $ax = b$ ” equation by “ $a$ ”
		division-simple-d	Dividing an “ $ax = b$ ” equation by “ $a$ ” when diagrams are present
subtraction-const	subtraction-const	Subtracting a constant term	
	subtraction-const-d	Subtracting a constant term when diagrams are present	
subtraction-var	subtraction-var	Subtracting a variable term	
	subtraction-var-d	Subtracting a variable term when diagrams are present	
KCs for the diagram steps	selectd-given-eq	selectd-given-eq	Selecting a diagram for the given equation
	selectd-subtract-const	selectd-subtract-const	Selecting a diagram for subtracting a constant term
	selectd-subtract-var	selectd-subtract-var	Selecting a diagram for subtracting a variable term
	selectd-divide	selectd-divide	Selecting a diagram for divisions

**Table 17.** Model metrics values for the KC models.

<b>KC model</b>	<b>AIC</b>	<b>BIC</b>	<b>RMSE (student blocked)</b>	<b>RMSE (item blocked)</b>	<b>RMSE (unblocked)</b>
Original	4,431.81	4,894.95	0.3682	0.3088	0.3075
Updated	4,347.60	4,933.32	0.3675	0.3048	0.3047

### **Discussion**

Comparing two conditions that differed on the frequency of receiving visual scaffolding during algebra problem solving, this study found that engaging anticipatory diagrammatic self-explanations for all problem-solving opportunities was a helpful level of scaffolding and did not “over-scaffold” learning. Contrary to my expectation that interleaving visual scaffolding might foster better learning, reducing the frequency had a host of detrimental effects. In the following, I will provide detailed discussions on the results.

First, sustaining visual scaffolding for every problem-solving opportunity in the ITS did not over-scaffold learning. Students who received the sustained scaffolding for every problem-solving opportunity did not perform differently from students who with interleaved visual scaffolding on any of the posttest item categories. In fact, students in the interleaved condition, on average, scored lower on the posttest than the pretest on two of the three test item categories whereas students in the All-Diagram condition improved from the pretest to the posttest, when observed descriptively. This descriptive difference might be due to the fact that students in the Interleaved condition did not have a very smooth learning experience, as evidenced by the performance differences in the tutor. Interleaved visual scaffolding, which made problem solving harder and slower, might have only added “un-desirable difficulties” which did not lead to enhanced learning (Bjork & Bjork, 2011; Koedinger & Alevan, 2007).

Also, consistent with my prior studies (Study 3 & 4), engaging in more anticipatory diagrammatic self-explanation resulted in better problem-solving performance in the ITS. In-depth analyses using the ITS log data revealed that students in the All-Diagram performed better not only on even-numbered problems, in which only the All-Diagram condition received the scaffolding, but also on odd-numbered problems, in which students in both conditions solved exactly the same problem. It is possible that the performance differences were due to the increased practice of using the visual support to solve equation problems in the All-Diagram condition; students in the All-Diagram condition might have become fluent in using the visual scaffolding and therefore solved more

efficiently even on odd-numbered problems. These results indicate that the overall performance differences did not come only from the problems with no visual scaffolding, but rather came from the entire learning experience, including their interaction with the scaffolded problems. This finding implies that sustaining the visual scaffolding throughout the entire learning process benefits learners by helping them have a smooth learning experience. It is true that students in the All-Diagram condition received twice the visual support than the students in the Interleaved condition, and it would be reasonable to think that the increased exposure might have helped the students gain competence in using diagrams to solve symbolic problems.

Why did sustaining visual scaffolding benefit students? The Knowledge Component modeling analysis provides evidence that students in the interleaved condition exercised different types of skills (i.e., Knowledge Components) for problems with visual scaffolding and those without visual scaffolding. Students in the sustained condition, on the other hand, were consistently practicing the skills of “solving problems with diagrams.” It may be that students who received the scaffolding for every problem-solving opportunity benefited because their learning experience was focused and consistent.

However, the findings from the Knowledge Component modeling also indicate that students in the interleaved condition were engaged in learning that students in the sustained condition did not practice (i.e., solving equations without visual scaffolding). Given that students eventually need to be able to solve equation problems without visual scaffolding (e.g., more advanced equation problems), it could be that students’ practice with interleaved visual scaffolding may lead to better learning outcomes in later phases of equation solving that involve more complicated problem types. The current study did not capture this potential benefit because these later stages were not reached. Future research could explore this possibility.

## 4 Chapter 4: Supporting Self-Regulated Use of Visual Representations

### 4.1 Motivation

Chapter 3 illustrated how I explored the design of a visual representation for early algebra (namely, tape diagrams) and the design and empirical evaluations of instruction using the visual representation through multiple instructional strategies (i.e., diagrammatic self-explanation). The focus of assessment, or the target instructional goal, in these completed studies had been students' learning of domain-specific knowledge and skills in early algebra, which include conceptual understanding, procedural skills, and strategies in solving algebra problems (Chu et al., 2017; Crooks & Alibali, 2014; Koedinger et al., 2008). In the later work, I made a shift in the targeted learning goal. Instead of only focusing on students' domain-specific knowledge/skills, I also targeted students' self-regulatory behaviors involving the use of the visual scaffold (diSessa, 2004). That is, I explored how we might facilitate learners' self-regulatory use of diagrams *and* whether and how it would affect their domain-level learning. This shift in the research question also means a shift in how to assess learning. As illustrated in Table 18, rather than restricting the assessment environment in a way that students will not be allowed to use available resources, I used a learning/assessment environment where students are allowed to use available help (i.e., diagrams) in solving problems. Such an environment is more aligned with the authentic learning environment in which learners are not restricted in terms of the use of resources around them to navigate everyday problem solving (Schwartz & Arena, 2013). Drawing on the idea of "choice-based assessment" (Cutumisu et al., 2015, 2019; Schwartz & Arena, 2013), I investigated students' choices with regard to the use of visual representations in a learning/assessment environment and how their choices relate to their domain-level learning. Choice-based assessments have been implemented in a number of prior studies in the field of the learning sciences (e.g., Cutumisu et al., 2015) and learner choice has been argued as representing a valid self-regulated learning construct (Bransford & Schwartz, 1999; Roll et al., 2011).

**Table 18.** Comparison between Studies 1-5 and Studies 6-8.

	<b>Studies 1-5</b>	<b>Studies 6-8</b>
<b>Target learning goal</b>	Domain-specific knowledge and skills	Self-regulated learning and domain-specific knowledge and skills
<b>Why is it important?</b>	Middle-school students learn algebra knowledge and skills to prepare for future learning in advanced STEM courses	Choosing to use visual representations strategically is a critical self-regulated learning skill that learners need to acquire
<b>Assessment approach</b>	Domain knowledge/skills test where learners are not supposed to use visual representations	Choice-based assessment where learners can decide whether (or not) to use visual representations

## 4.2 Study 6: Investigating Patterns of Students' Diagram Use in a Choice-based Diagram Tutor

I first designed Choice-based Diagram Tutor, which gives learners a control over whether and when to engage with the anticipatory diagrammatic self-explanation in the tutor, instead of asking learners to process anticipatory diagrammatic self-explanation at every problem-solving opportunity (Figures 30-33).



**Fig. 30.** In the Choice-based Diagram Tutor, students start with an interface in which only the symbolic equation-solving window is available.

Which diagram shows the current step?

$2x + 3 = 7$

Finish Problem Use Diagrams? Hint

**Fig. 31.** When students press the “Use Diagrams?” button, the anticipatory diagrammatic self-explanation step appears, and students are asked to work on the diagram step. Clicking on the “Hint” button will show multi-step text-based hints for the diagram step. Correctness feedback is provided based on students’ selection.

Please solve for x

$2x + 3 = 7$

Finish Problem Use Diagrams? Hint

**Fig. 32.** Once the diagram step is done, students will work on the corresponding symbolic step. The tutor provides targeted feedback in text format. Hints are also available, and hints do not direct students to use diagrams.

Please solve for x

$2x + 3 = 7$

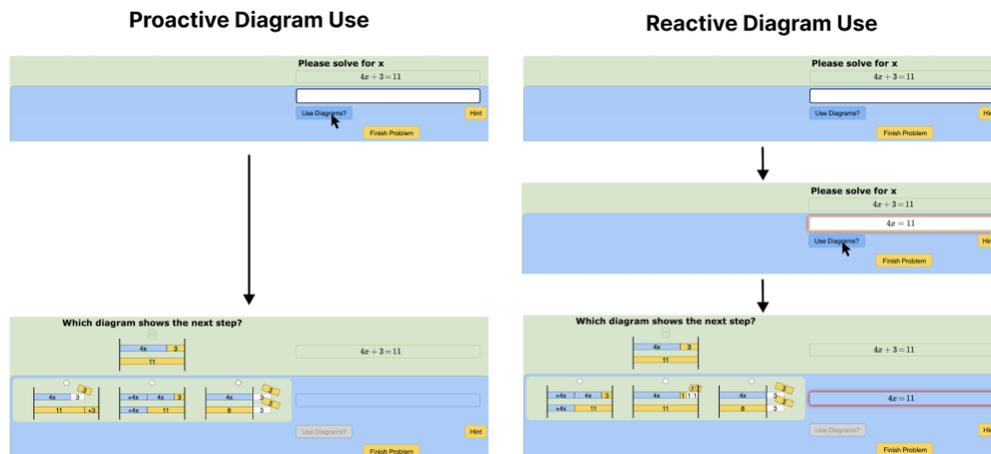
Finish Problem Use Diagrams? Hint

**Fig. 33.** For any equation-transformation steps, students can also choose not to use diagrams. At any point during symbolic problem solving, students can press the “Use Diagrams?” button to get the visual scaffolding. The tutor does not prompt students to direct students’ attention to a certain choice behavior.

The Choice-based Diagram Tutor helps us understand students’ choice behaviors when the use of visual scaffolding is optional. Strategically choosing to use or not to use the visual representation in our equation-solving tutor constitutes an appropriate task for measuring learners’ self-regulated choice in learning with visual representations (diSessa, 2004). Chin et al. (2019) discuss that a critical aspect in designing a choice-based assessment environment is to select a task for which learners have a natural tendency *not to use* the strategy of interest. Using tape diagrams when practicing equation solving is considered as having this characteristic at least from two viewpoints. First, engaging with anticipatory diagrammatic self-explanation requires additional steps. Because diagram steps are given for every *transformation* step (i.e., steps that make the state of the equation closer to isolating the variable), the number of steps would almost double, compared to the problem with no diagrammatic self-explanation support. Although my prior studies have consistently demonstrated that such additional steps do not affect the total number of problems students solve in a fixed time period compared to students who do not have diagrammatic steps (Nagashima, Bartel, Tseng et al., 2021; Nagashima, Bartel, Yadav et al., 2021), it is a reasonable expectation that middle-school students would have a tendency to avoid having to do additional steps. Furthermore, tape diagrams introduce a new representation that many middle-school students are not familiar with. Even though tape diagrams are increasingly incorporated in classroom instruction (Murata, 2008), they are still, at least in the U.S., not yet prevalent (Nagashima, Yang et al., 2020). Therefore, from a learner’s point of view, it might not make sense (or might not be attractive) to choose to engage with such a new representation when they are given the choice.

In addition to the decision of whether or not to use diagrams, which I henceforth call *spontaneous diagram use* (Uesaka & Manalo, 2012), I also investigated students’ self-regulated use of diagrams in the Choice-based Diagram Tutor. Drawing on models and theories of self-regulated learning, in the context of the Choice-based Diagram Tutor, I categorized two types of diagram use (Figure 34):

- *Proactive diagram use*: using diagrams before attempting to solve the corresponding symbolic step
- *Reactive diagram use*: using diagrams after making one or more incorrect attempts on the symbolic step



**Fig. 34.** Proactive diagram use represents the use of diagrams *before* making any attempts on the symbolic step, whereas reactive diagram use represents the use of diagrams *after* making one or more incorrect attempts.

Any use of diagrams in the tutor falls into one of these categories. These choice categories are informed by the literature on help-seeking in computer-based tutoring systems (Aleven & Koedinger, 2001; Wood & Wood, 1999) and by Zimmerman and Campillo's (2003) cyclical model of self-regulation. As described in Chapter 2, Zimmerman and Campillo's (2003) model captures phases of self-regulated behaviors in which learners evaluate the difficulty of the target task ("forethought" phase), self-monitor learning strategies ("performance" phase), and evaluate and reflect on the use of the strategy ("self-reflection" phase). Under this model, in the Choice-based Diagram Tutor, students may choose to use diagrams by assessing the difficulty of the given equation, self-monitor how they perform by using diagrams, and adjust their use of the strategy through self-reflection. Therefore, it is reasonable to interpret proactive diagram use as self-regulated, well-planned use of diagrams (below I examine this conjecture in light of data from students' work with the Choice-based Diagram Tutor), whereas reactive diagram use would represent unplanned diagram use.

### **Pilot classroom study**

I conducted a classroom pilot study in 2021 to investigate how learners interact with the choice-based diagram tutor<sup>6</sup>. The purposes of the study were twofold. First, I collected *baseline* data regarding how much students would choose to use tape diagrams in the Choice-based Diagram Tutor when no additional prompts to support students' choice were used and how that would affect their performance and domain-level learning. In other words, I was interested in knowing students' "natural tendency" regarding the use of the visual scaffold when they are given the choice (Chin et al., 2019) and what choice can be considered as self-regulated. Second, I investigated if students' use of diagram is associated with their perceptions regarding the utility of using diagrams, and their learning outcomes and performance. The following research questions were explored:

*RQ1: When given the choice, how frequently will students choose to engage with anticipatory diagrammatic self-explanation?*

*RQ2: Will the self-regulated diagram use be related to their perceptions regarding using diagrams during problem solving and other behavioral measures (e.g., prior knowledge)?*

*RQ3: Will students learn and perform better if they autonomously choose to engage with anticipatory diagrammatic self-explanation?*

*RQ4: Will students who proactively use diagrams more often perform and learn better?*

### *Method*

Thirty 6<sup>th</sup> grade students from two classes at a private middle school in the U.S. participated in the study. All students participated from their in-person classroom with their own teachers. On the first day, participants were asked to work on a 10-minute, web-based pretest. On the second day, one of the teachers gave a brief lecture on tape diagrams to all students and then students used the choice-based tutor to practice problem solving for about 25 minutes. To complete the study, a posttest was conducted. The posttest included several survey questions, which asked students their perceptions regarding using tape diagrams during algebra problem solving. Per teachers' report, students had never seen tape diagrams at school.

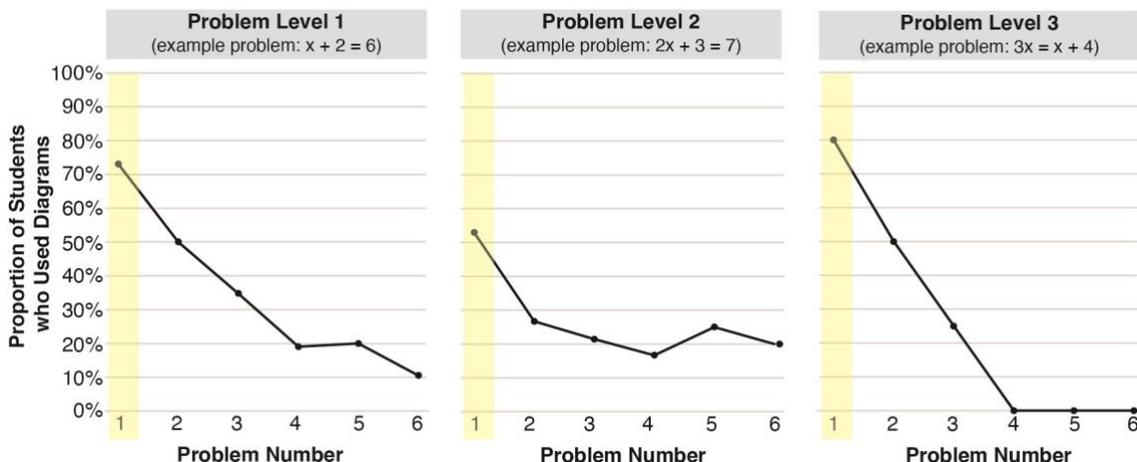
---

<sup>6</sup> To read more about this work, please see: Nagashima, T., Tseng, S., Ling, E., Bartel, A. N., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevan, V. (2022). Students' self-regulated use of diagrams in a choice-based Intelligent Tutoring System. In *Proceedings of the Annual Meeting for the International Society of the Learning Sciences*, Hiroshima, Japan.

The pretest and posttest included seven conceptual knowledge items and four procedural knowledge items. Among these 11 items, two conceptual items and two procedural items used tape diagrams in their problems. Two versions of the test were created and assigned to students in a counter-balanced way across pretest and posttest. The survey questions embedded in the posttest asked, on a scale of 1-100, 1) if students liked tape diagrams in the tutor, 2) if they felt confident about using tape diagrams for problem solving, 3) if they felt that tape diagrams helped learn algebra skills, 4) if they felt that tape diagrams helped improve the accuracy of problem solving in the tutor, and 5) their general perception regarding how good they are at solving equation problems without using diagrams. The survey items were developed partly based on prior research (e.g., Uesaka et al., 2007) but also designed specifically for the topic of my research to explore various factors that might affect students' diagram use. All students were given the choice-based diagram tutor, which included 22 problems of four different problem types (called "levels" in the tutor): Level 1)  $x + a = b$ , Level 2)  $ax + b = c$ , Level 3)  $ax = bx + c$ , and Level 4)  $ax + b = cx + d$ , assigned in this order. According to the teachers, students had seen Level 1 problems but had never solved equations in a formal way (by subtracting "a" from both sides of " $x + a = b$ ")

### *Results*

Four students did not complete all study sessions and were therefore excluded from the analysis. Remaining data from 26 students were included in the analyses. To answer RQ1 (*When given the choice, how frequently will students choose to engage with anticipatory diagrammatic self-explanation?*), we calculated the frequency of using diagrams (i.e., clicks made on the "Use Diagrams?" button). On average, students solved 9.12 problems ( $SD = 5.14$ ) and chose to use diagrams 2.73 times ( $SD = 2.29$ ). That is, on average, tape diagrams were used 0.3 times per problem. When looking at patterns of diagram use, I found that students tended to request diagrams the most for the first problem in each level, but their use decreased as they solved more problems of the same type in each level (Figure 35).



**Fig. 35.** Proportion of students who spontaneously used diagrams at least once for each problem.

Next, to explore RQ2 (*Will the self-regulated diagram use be related to their perceptions regarding using diagrams during problem solving and other behavioral measures?*), I first ran a linear regression with the rate of (spontaneous) diagram use in the tutor as a dependent variable and the scores for the five survey questions as independent variables. Pretest score was also added as an independent variable. The model showed that none of the six independent variables predicted the frequency of using diagrams in the tutor. Contrary to my expectation, this result suggests that students' prior knowledge and their (post-hoc) perceptions regarding tape diagrams and math ability are not associated with the choice behavior. Table 19 shows students' pretest and posttest scores.

**Table 19.** Students' pretest and posttest scores (standard deviations).

	<b>All items</b>	<b>Conceptual knowledge (max: 7)</b>	<b>Procedural knowledge (max: 4)</b>	<b>Tape diagram knowledge (max: 4)</b>
Pretest	6.15 (2.01)	3.42 (1.36)	2.73 (1.00)	2.15 (1.05)
Posttest	6.96 (2.42)	4.04 (1.63)	2.92 (1.05)	2.69 (1.29)

To investigate RQ3 (*Will students learn and perform better if they autonomously choose to engage with anticipatory diagrammatic self-explanation?*) and RQ4 (*Will students learn and perform better if they proactively use diagrams?*), we constructed linear regression models with the following dependent variables: posttest score, posttest score for items without tape diagrams, average time spent per symbolic step, average number of hints used per symbolic step, and average number of incorrect attempts for the symbolic steps. We included each of the spontaneous, proactive, and reactive diagram use frequencies in each model as independent variables. Overall pretest score was included to control for students' initial knowledge level. Spontaneous diagram use predicted higher overall posttest scores; however, this relation did not reach statistical significance ( $\beta = 1.47, p = .07$ ). Spontaneous diagram use did not predict any other dependent variables. Proactive diagram use, however, predicted higher overall posttest scores ( $\beta = 2.01, p < .01$ ) and higher posttest scores on items without tape diagrams ( $\beta = 1.03, p = .04$ ). Reactive diagram use, on the other hand, was associated with lower overall posttest scores ( $\beta = -4.75, p = .03$ ) and lower posttest scores on items without tape diagrams ( $\beta = -3.39, p = .01$ ). Reactive diagram use also predicted greater hint use ( $\beta = 5.21, p < .01$ ).

### *Discussion*

Although students' use of diagrams was generally low, students' proactive diagram use predicted higher learning outcomes. This association was also observed on test items that did not include tape diagrams, suggesting near transfer of learning. However, as we cannot establish any causal relationships from the results, it is possible that certain characteristics that we did not control for led both to more proactive diagram use and better learning. For example, proactive diagram use might be indicative of superior monitoring ability or a propensity to monitor one's comprehension more frequently, which might lead to other behaviors that might contribute to better learning.

Study 6 contributes understanding of self-regulated learning with visual representations. Specifically, it provides evidence that effective use of visual representations involves more than simply using the visual representations spontaneously; rather, proactive use, which I consider as a form of self-regulated diagram use that involves assessment of task difficulty and planning of whether or not to use diagrams, leads to better learning. In a related, pilot study with eight middle-school students who thought aloud as they used the tutor, I found that students who tended to proactively use diagrams thought deeply about the given problem before attempting to solve it and correctly understood diagrams. On the other hand, those with frequent reactive diagram use did not seem to understand the diagrams, and therefore seemed to process diagrams in a shallow way (e.g.,

selecting diagrams that look right). I acknowledge though that different interpretations of these patterns of diagram use may be possible.

### **4.3 Study 7: Generating Ideas for Metacognitive Interventions with Children**

Study 6 left open an important question: *how might we help students become self-regulated learners who use diagrams proactively and learn well?* To approach this question, I conducted co-design sessions with middle-school students to generate ideas and designed metacognitive interventions.

#### **Method**

Research assistants and I recruited a total of eight students in the U.S. The students included one 4th grader, one 5th grader, one 6th grader, four 7th graders, and one 8th grader, and were recruited through our previous contacts. Each student remotely met with a researcher for an hour-long session in which they first 1) used the Choice-based Diagram Tutor to solve 5-10 problems (to familiarize themselves with the tutor) and then 2) brainstormed ideas for promoting self-regulated diagram use. Specifically, for the latter part, we gave students three following prompting questions to facilitate the idea-generation activity. Each student and the researcher used and shared a Figma board (<https://www.figma.com/>) to draw and write down ideas as they generated. Sessions were video recorded for a later analysis.

Prompting questions:

- “What would be some features that you wish the system to have to help you see the benefits of using diagrams?”
- “What would be some features that will help you feel motivated to think about using diagrams or not?”
- “What would be some features that you wish the system to have to help you think carefully about whether or not to use diagrams for solving equation problems?”

#### **Analysis**

The eight co-design sessions resulted in a total of approximately eight hours of video recordings. To consolidate and categorize generated ideas, three researchers performed Affinity Diagramming (Lucero, 2015). The research team met twice to engage in all of the core phases of Affinity Diagramming (familiarization with data, initial coding, mid-level categorization, high-level categorization).

## Results

Initial coding produced 117 ideas from students, which were grouped together according to similarities. This process produced 13 mid-level themes. Then, the 13 ideas that share similar themes were grouped together to make five high-level ideas. In the following, I describe these five ideas.

*1: "Tell me that diagrams are there to help, they are not there for no reason"*

Students wanted to have examples, videos, or tutorials that help them understand why they should consider using diagrams. Because students were not familiar with tape diagrams, they were, by default, reluctant to even consider using diagrams without such information. Specifically, students wanted an explanation of 1) tape diagrams and 2) how they could be useful.

*2: "I want to be prompted to consider using diagrams when they can be helpful"*

Students frequently mentioned, when they thought-aloud while using the tutor, that they would use diagrams when they were not sure how to solve the presented problem. Students believed that such a moment would be perfect timing for them to think about using diagrams, and therefore recommended that being prompted to consider using diagrams at such moments would be helpful.

*3: "Show me how diagrams are helping (or not helping) me"*

Students expressed that they would like to see if there are any improvements they make when they used diagrams. What differences do diagrams make? They care about whether using diagrams influenced their own problem solving, not only general, established benefits of it. Students mentioned that they would be motivated to think about using diagrams if they can clearly see advantages and disadvantages of using diagrams to aid their *own* problem solving.

*4: "A diagram badge can help me think about using diagrams"*

Referring to some educational games (e.g., Duolingo: <https://www.duolingo.com/>), students wanted to have more motivational features, such as a badge that would be given based on their own progress. Many students generated the idea of creating a badge for using diagrams; they think that recognizing their progress and use of diagrams in the form of a badge would help them keep motivated to learn and think about whether to use diagrams.

*5: "Why is this diagram wrong? I need an explanation"*

Finally, several students were confused about the selection of diagrams for diagrammatic steps in the tutor. While students visually were able to tell how the diagram choices look different, to further help

them understand the benefits of using diagrams, students wanted to see an explanation of why correct/incorrect diagram options are correct/incorrect.

### **Designing interventions to support students' self-regulated use of diagrams**

To design an intervention for supporting students' self-regulated diagram use, I used the Metacognitive Choice Behavior Model that I proposed in Chapter 2 (Figure 9, also see Figure 39 below) as a design guide. Specifically, I wanted to create interventions that would affect the behavioral stages separately (at different points during the use of the tutor) rather than designing a single intervention that could address all the phases. As Opezzo and Schwartz (2013) argue, interventions may be needed at different stages to promote specific behaviors effectively. In my design, I focused on the following phases:

#### *Pre-intend to intend transition*

For one to move from the *pre-intend* to the *intend* stage, they need to understand the benefits of using the new strategy. Once they successfully understand the benefits, they will be more interested in thinking about whether or not to use the strategy. I thought that, in the context of the tutor, this stage would be an appropriate time to address the first idea (“*Tell me that diagrams are there to help, they are not there for no reason*”) because the first idea represents a need for understanding benefits of using diagrams.

#### *Intend to choose to implement/choose not to implement transition*

For one to move from the *intend* to the *choose to implement* and the *choose not to implement* phases, they need to “see the opportunity” to think whether or not to use the strategy (Opezzo & Schwartz, 2013). That is, students would need to recognize the conditions under which diagrams might be helpful. If students do not see such an opportunity, they would not be able to even consider about using the strategy at the appropriate timing. Therefore, the second idea from the co-design sessions (“*I want to be prompted to use diagrams when they are helpful*”) could be addressed during these phases.

#### *Reflect phase*

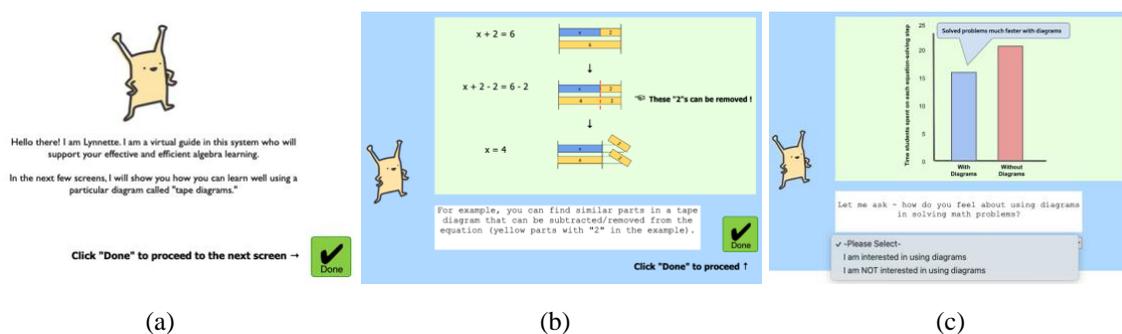
Based on the model, it is expected that self-regulated learners reflect on their own performance and progress after deciding whether to use the strategy and actually using (or not using) it. Reflection is a key to promoting further strategic use of the strategy because students can self-judge their own performance and learn from it (e.g., by increasing their use of the strategy if they see its benefit for

their own work or reducing their use of strategy if they do not see its benefits). In the *reflect* phase, the ideas 3 (“*Show me how diagrams are helping me*”) and 4 (“*A diagram badge can help me think about using diagrams more*”) could be appropriately addressed as it indicates a need for reflecting on the choice of using or not using diagrams.

Drawing on this framework and students’ ideas, research assistants and I designed the following interventions to support students’ self-regulated use of diagrams:

### A tutorial teaching the benefits of using diagrams

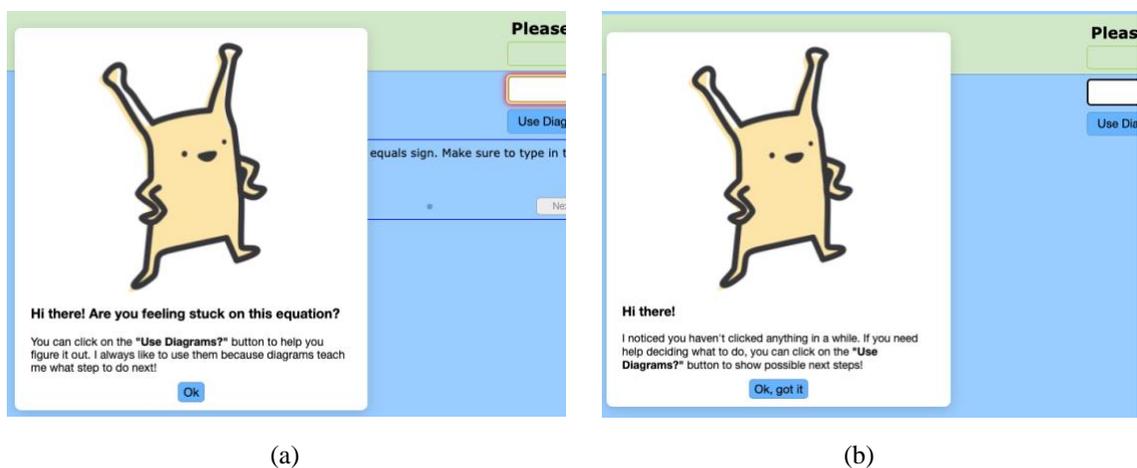
To support students’ transition from the *pre-intend* to *intend* phases, we developed an introductory tutorial (shown before students start solving problems) that shows what tape diagrams are and how they could be useful in solving equations (Figure 36). The tutorial was composed of three screens. In the first screen (a), students are introduced to a fictional character (named “Lynnette”). The second screen (b) shows students that tape diagrams can help them visually see the structure of equations. The third screen (c) teaches students that research has shown that using diagrams help students solve problems much faster than students who solved the same problems without using any diagrams. It is designed as a brief introduction with minimum interactions. At the end of the tutorial, it asks students whether they are interested in using diagrams or not and collects students’ response in the tutor log.



**Fig. 36.** A tutorial consisting of (a) an introduction screen, (b) a screen illustrating how tape diagrams can be useful, and (c) a screen teaching what previous research has shown about using tape diagrams. At the end of the tutorial, (c) students are asked for their intention regarding using diagrams.

### Adaptive recommendations providing opportunities for considering diagram use

To help students see opportunities where the use of the strategy might be particularly helpful (i.e., from the *intend* to the *choose to/not to implement* phases), we designed and developed an adaptive recommendation feature that prompts students to think about whether or not to use diagrams. Note that it does not strongly recommend the use of diagrams; it rather encourages students to *think about* using diagrams. The recommendation was designed as a pop-up screen that shows up (a) when students make three consecutive incorrect attempts on a symbolic step or (b) when students stay idle for 90 seconds on any step. As Figure 37 shows, different messages are shown for these cases. Students can click on the “Ok” and “Ok, got it” button to close the window.



**Fig. 37.** Adaptive recommendation popup windows shown when (a) students make three consecutive mistakes on symbolic steps and when (b) students are idle for more than 90 seconds.

### A student-facing dashboard that helps students reflect on their use of diagrams

To help students reflect on their choice behaviors, we designed and developed a student-facing dashboard that presents the student's own problem-solving performance (correctness rate) on symbolic steps when they used diagrams and when they did not use diagrams, represented as two separate bars in a bar graph. It also shows badges for using diagrams and mastering pre-defined Knowledge Components assigned in the tutor. The dashboard is given at the end of every problem level/set and shows the information only from the current problem level. To proceed to the next level, students need to indicate their feeling of using diagrams using the 5-likert scale "smiley" survey and respond to an open-ended textbox asking them for the reasoning behind their choice (Figure 38).

Lynnette Red Level 5

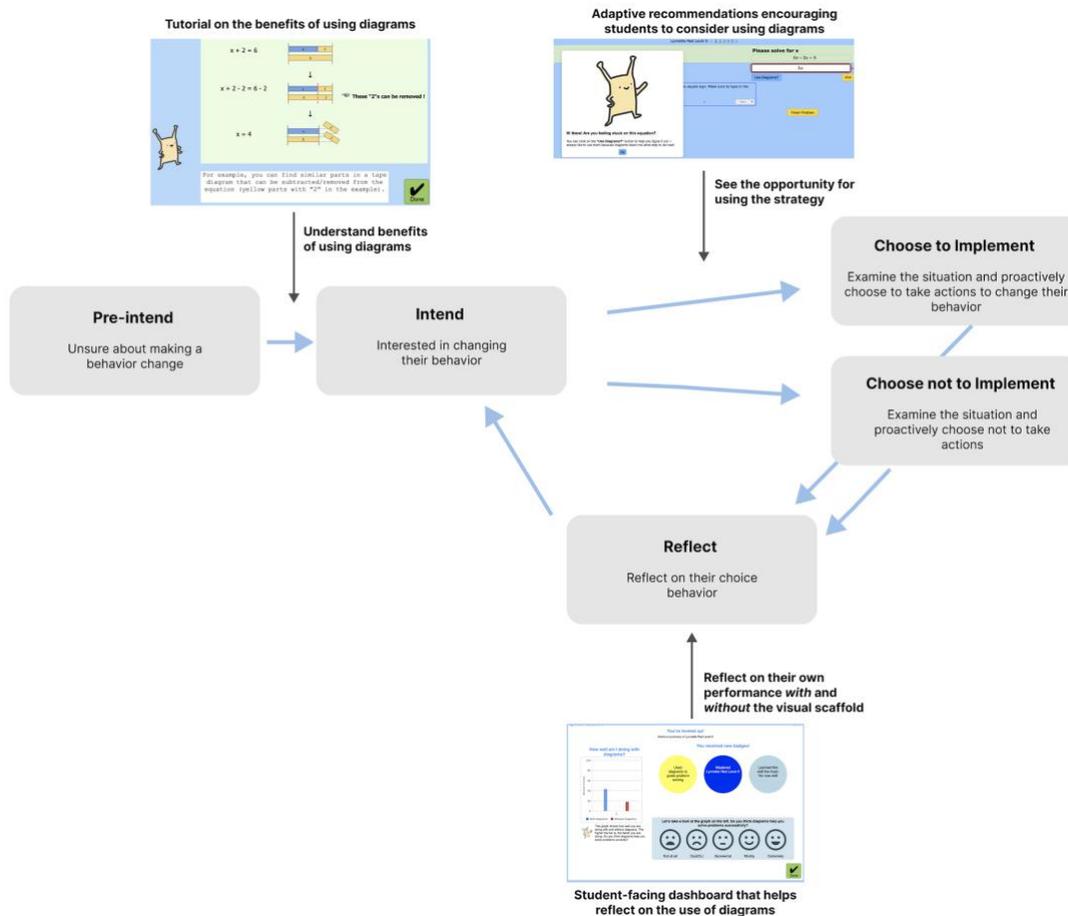
tomo



**Fig. 38.** A dashboard screen that is presented at the end of each problem level. Students are shown a bar graph showing their own performance with and without diagrams (left) and badges (right). At the bottom right, students are asked to select one smiley icon to indicate their current feeling about using diagrams. Once they make a selection, the bottom-right smiley area is replaced with an open textbox asking students for the reasoning for the selection.

These three ideas were then mapped onto the Metacognitive Choice Behavior Model. Following Oppezzo and Schwartz (2013)'s argument that different treatments may be needed to promote behaviors in each stage of behavior change, we embedded the three core intervention elements into the Choice-based Diagram Tutor to develop a *metacognitive intervention package*. This design of the intervention would make it hard to test the effect of individual intervention elements (also see the experiment design in Section 4.5) but would allow us to maximize the potential treatment effect on students' learning outcomes.

As Figure 39 shows, I expected that the tutorial would help students understand the benefits of using diagrams, therefore support their transition from the *pre-intend* phase to the *intend* phase (i.e., students would get interested in thinking about using the strategy). The adaptive recommendations are designed to support students' transition from the *intend* to *choose to implement* and *choose not to implement* phases. The recommendations, which are given based on students' interactions with the tutor, would help students see opportunities to consider whether or not to use the strategy (Oppezzo & Schwartz, 2013). Lastly, the dashboard would help students *reflect* on their choice behaviors by comparing and contrasting their performance when they use diagrams and when they do not use diagrams. Self-reflection with the dashboard may also support students' decision making for the next opportunity (e.g., "I am doing better with diagrams, so I will keep using diagrams").



**Fig. 39.** Each of the three intervention elements (tutorial, recommendations, and dashboard) works at different stages in supporting students' self-regulated choice-making behaviors.

Further, my colleagues and I also designed and implemented feedback messages for diagram selections, one of the ideas that students generated during co-design sessions. For example, as shown in Figure 40, when students choose an incorrect diagram, the tutor provides a feedback message helping students to understand why it is an incorrect representation. Since it was not clear how this feature would fit in the Metacognitive Choice Behavior Model (not as clear as other intervention elements), as we describe later, we rather implemented this feature as a general improvement of the

Choice-based Diagram Tutor (i.e., not part of the metacognitive intervention package, which we tested in Study 8 below).

The screenshot shows a math problem interface. At the top, it asks "Which diagram shows the next step?". The initial equation is  $4x + 3 = 11$ . Below this, three diagrams are presented as options:

- Diagram 1: A bar model with a top bar divided into a blue section labeled  $4x$  and a yellow section labeled  $3$ , and a bottom bar labeled  $11$ . A radio button next to it is selected.
- Diagram 2: A bar model with a top bar divided into a blue section labeled  $4x$  and a yellow section labeled  $3$ , with a  $+3$  label to the right. The bottom bar is labeled  $11$  with a  $+3$  label to the right.
- Diagram 3: A bar model with a top bar divided into a blue section labeled  $4x$  and a yellow section labeled  $3$ , with a  $+3$  label to the right. The bottom bar is labeled  $8$  with a  $+3$  label to the right.

Below the diagrams, there is a feedback message box that reads: "That does the same thing to both sides of an equation, but adding 3 does not help you solve the problem." There are "Previous" and "Next" navigation buttons. A "Use Diagrams?" button and a "Hint" button are also visible.

**Fig. 40.** Feedback messages are shown when students make a (correct and incorrect) diagram selection. Messages tell students why the selected diagram is correct or incorrect.

#### 4.4 Study 8: Supporting Students' Self-Regulated Use of Diagrams and Math Learning

In May 2022, I conducted an in-vivo classroom study (Koedinger et al., 2009) to experimentally investigate the effectiveness of the designed metacognitive intervention package on students' self-regulated use of diagrams, their performance (when using the Choice-based Diagram Tutor), their learning (from pretest and posttest), and learning transfer in the classroom context.

##### Research questions

This experiment tests the effect of the metacognitive intervention package against a control condition in which the package was not used ("Metacognitive" condition vs. "Non-Metacognitive" condition, respectively). The following questions were examined:

*Will the metacognitive intervention package designed with children help learners:*

1. *(RQ1) use diagrammatic self-explanation strategically during learning, when it is provided as an optional resource,*
2. *(RQ2) achieve greater performance and learning gains in doing so,*
3. *(RQ3) do better on future learning tasks in a transfer environment with diagrams, and*
4. *(RQ4) perceive the usefulness of using diagrams compared to learners without the package?*

*Also, (RQ5) how is diagram use associated with learning outcomes?*

### **Method: Participants**

A total of 179 students from two schools in the U.S. participated in the study, taught by two teachers. The sample included 38 5<sup>th</sup> graders, 37 6<sup>th</sup> graders, 86 7<sup>th</sup> graders, and 18 8<sup>th</sup> graders. Both teachers noted that students' prior exposure to tape diagrams was minimal. Students in each class were randomly assigned to either the Metacognitive condition or Non-Metacognitive condition. In both schools, all students participated in the study from their in-person classroom.

### **Method: Measures**

I used four types of assessment measures:

#### *Domain knowledge pretest and posttest*

Pretest and posttest that measure students' conceptual understanding and procedural skills in early algebra were developed based on items from Study 3, 4, 5, and 6 (Nagashima, Bartel et al., 2020; Nagashima, Bartel, Yadav., 2021) as well as based on prior literature (Booth et al., 2013). Each test had 16 multiple-choice conceptual knowledge items and five open-ended procedural items. Two isomorphic versions of the test were developed and assigned in a counter-balanced way across pretest and posttest.

#### *Transfer test*

To measure if students' propensity to proactively use diagrams will transfer in another environment, I designed a transfer assessment where students were first given a story problem of early algebra (which students never practice in the tutor) and asked about their intention of whether to use diagrams or not to solve the problem (Figure 41). The assessment was designed so that, depending on whether students *proactively* decided to use diagrams, a corresponding next screen was assigned (i.e., those who chose to use visual help were given a screen with visual help whereas those who chose not to use visual help were given a screen without any visual help). Students, regardless of their choice,

were then asked to 1) construct a symbolic equation from the story problem and 2) solve the equation (Figure 42). However, I realized that text had a misleading error (the sum of three consecutive *even* numbers can never be 45, an odd number).

Please read the question below:

If the sum of the smallest and largest of three consecutive even numbers is 45, what is the value of the second largest number in the series?

You can solve this problem by making a simple equation that you practiced in the software.  
**HINT: if we let the smallest, unknown number be "x", we can write down the second largest number as "x + 2" and the largest number as "x + 4"!**

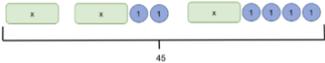
Now, you can choose to solve this problem without any help or you can choose to use a new kind of diagram. Would you want to use the diagram?

I will use the diagram to solve the problem

I will solve the problem without the diagram

Please say a bit more about your decision above about using diagrams. Why or why not?

**Fig. 41.** First screen of the transfer item asking students for their intention of whether to use diagrams.

<p>Now, let's try to solve the problem using the diagram below:</p> <p>If the sum of the smallest and largest of three consecutive even numbers is 45, what is the value of the second largest number in the series? You can solve this problem using a simple equation that you practiced in the software.</p> <p><b>HINT: if we let the smallest, unknown number be "x", we can write down the second largest number as "x + 2" and the largest number as "x + 4"!</b></p>  <p>What equation can you make?</p> <input type="text"/> <p>What is the answer (the second largest number)? Please show your work.</p> <input type="text"/>	<p>Now, let's try to solve the problem:</p> <p>If the sum of the smallest and largest of three consecutive even numbers is 45, what is the value of the second largest number in the series? You can solve this problem using a simple equation that you practiced in the software.</p> <p><b>HINT: if we let the smallest, unknown number be "x", we can write down the second largest number as "x + 2" and the largest number as "x + 4"!</b></p> <p>What equation can you make?</p> <input type="text"/> <p>What is the answer (the second largest number)? Please show your work.</p> <input type="text"/>
---	---

**Fig. 42.** Once students indicate the intention on whether to use diagrams, depending on their selection, one of these screens would be assigned. There is no way for them to go back to the previous screen and change their choice. The picture on the left shows a version with diagrams, the one on the right shows one without diagrams.

*Survey on the perception of diagram use*

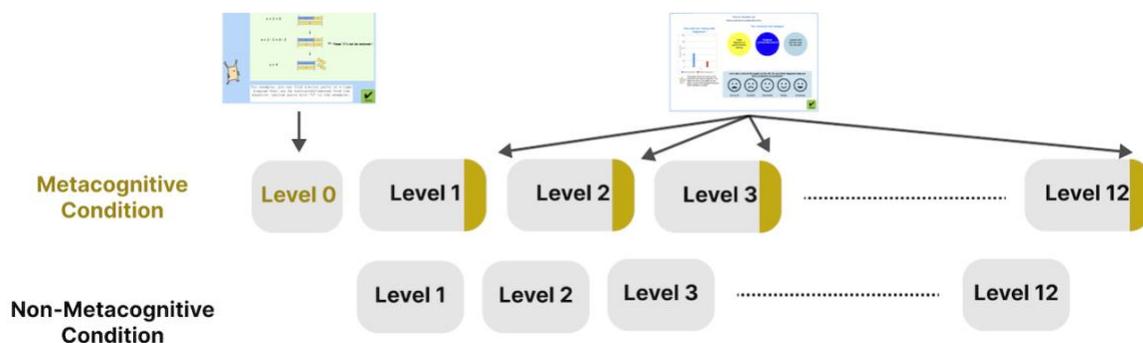
Drawing on Uesaka et al. (2007)'s questionnaire that measured students' perceptions regarding use of diagrams, I created a brief survey with questions about students' perceptions of diagram use. The survey had five 5-likert scale statements (1 = Strongly Disagree, 5 = Strongly Agree, see Appendix). It was administered both on pretest and posttest, administered immediately after the conceptual and procedural knowledge items mentioned above (on the same platform, Qualtrics).

*Choice behaviors and performance measures in the tutor*

We also collected log data in the Choice-based Diagram Tutor to observe behaviors regarding students' diagram use, as well as other performance measures such as error rates and hint use rates (Long & Alevan, 2014).

**Method: Materials**

I used two versions of the Choice-based Diagram Tutor in the study. One version, Metacognitive Choice-based Diagram Tutor, was assigned to students in the Metacognitive condition. In the Metacognitive Choice-based Diagram Tutor, the three metacognitive intervention elements (the tutorial, recommendation pop-ups, and dashboard) were embedded (Figure 43). Specifically, the tutorial was implemented at the very beginning of the tutor learning experience (as Level 0 problem set) and the recommendations were available throughout Levels 1-12. The dashboard was embedded at the end of each problem level (from Level 1 to Level 12) and students were required to interact with the dashboard to proceed to the next problem. The tutor version also had feedback messages for diagram selections embedded. The other version is the same Choice-based Diagram Tutor version we used in Study 6, except for the newly-added feedback messages for diagram selections. In other words, the only difference between the two tutor versions was the three intervention elements (tutorial, dashboard, and the adaptive recommendations). As described earlier, feedback messages for diagram selections were included in both conditions since they were implemented as a general improvement of students' tutor learning experience. Due to the integration of the metacognitive interventions package, students in the Metacognitive condition had more tasks to do than those in the Non-Metacognitive condition (Figure 43).



**Fig. 43.** Intervention assignment in the Metacognitive condition, in comparison to the Non-Metacognitive condition. In the Metacognitive condition, the tutorial was added as the first problem set (Level 0). The dashboard was added at the end of each problem set. The adaptive recommendation feature was available throughout all the problem sets, except Level 0 (not shown in the figure).

In both conditions, I used the same problems, assigned in the same order. It was designed so that students would see increasingly difficult problems as they proceed in the tutor. Table 20 shows the problem types implemented in both versions of the tutor.

**Table 20.** Problem levels and types added in both tutor versions, from Level 1 to Level 12

Problem level	Problem type	# of problems added
1	$x + a = b$	4
2	$ax + b = c$	4
3	$ax + b = c$	4
4	$ax = bx + c$	4
5	$ax = bx + c$	4
6	$ax + b = cx + d$	4
7	$ax + b = cx + d$	4
8	$ax + b = c$ (bonus content)	4
9	$ax = bx + c$ (bonus content)	4
10	$ax + b = cx + d$ (bonus content)	4
11	$ax = bx + c$ (bonus content)	4
12	$ax + b = cx + d$ (bonus content)	4

### Method: Procedure

The experiment took place during the schools' regular math class period, for the duration of five consecutive days. Figure 44 illustrates the experimental set up and procedure of the study. On the first day, following a brief introduction of the study by experimenters, students completed the web-based pretest. At the beginning of the second day, all students watched a video describing the tutor (i.e., the standard tutor functions available to the students in both conditions) and were told that some students may see additional features in the system. Starting on the second day till the fourth day, students used their assigned tutor to solve equation problems. On the fifth day, students completed the posttest and the transfer test. During the experiment, teachers (in the classroom) and experimenters (remotely) were present to support students when they had questions. After the study, all students were allowed to use both versions of the tutor (data logging stopped at the end of the fourth day, when students finished using the tutor in the study). Each study session lasted about 20 – 35 minutes. The average total time allocated for tutor learning was approximately 60 minutes.

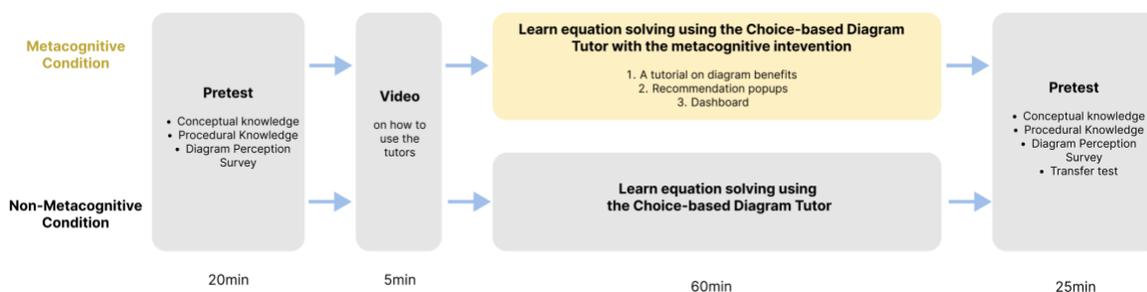


Fig. 44. Experimental design, types of assessments used, and the procedure of Study 8

### Results

Of the 179 students who participated in the study (who started the pretest, Metacognitive:  $n = 87$ , Non-Metacognitive:  $n = 92$ ), 168 students completed all parts of the study (pretest, tutor, and posttest). This dropout rate (6.1%) was within the expected range of 3.2% - 33.3%, calculated based on my previous in-person and remote classroom studies. I also excluded those who scored 100% on the pretest (Metacognitive:  $n = 1$ , Non-Metacognitive:  $n = 0$ ), and those who did not complete more than 50% of the test items on pretest and/or posttest, which we decided on before testing treatment effects (Metacognitive:  $n = 3$ , Non-Metacognitive:  $n = 4$ ). Using the 50% cutoff on pretest and posttest makes it possible to get a more accurate estimate of student learning than including all

students regardless of their completion status (Chan et al., 2022). Furthermore, I excluded all students from two advanced classes (i.e., honors classes in their schools) based on a teacher's suggestion (Metacognitive:  $n = 11$ , Non-Metacognitive:  $n = 9$ ), these students were included in the sample because teachers wanted the students to experience the study but expected that the students would be too advanced for the study content). I included 140 students who remained in the sample after applying the above-mentioned exclusion criteria. Of these students, 69 students were in the Metacognitive condition and 71 students were in the Non-Metacognitive condition. No statistically significant difference was found between the conditions on the dropout/exclusion rate,  $\chi^2(1, N = 179) = .05, p = .92$ .

Since the study aims to establish initial results of the treatment effects in a preferred environment (i.e., aiming to show whether the intervention *works* in the schools volunteered to participate in the study) rather than aiming to claim generalizable effects in a broader population, in the following analyses, I treat school as “fixed blocks” when examining treatment effects (Hedges et al., forthcoming).

*RQ1: Will the metacognitive intervention help learners use diagrams proactively during learning, when they are provided as an optional resource?*

From the tutor log data, I first collected data on the frequency of students' diagram requests (i.e., spontaneous diagram use). This data was obtained by counting the number of times students clicked on the “Use Diagrams?” button in the tutor. I further calculated students' proactive diagram use (any spontaneous diagram use would be categorized as either proactive or reactive diagram use, Figure 34). Table 21 shows descriptive statistics of these types of diagram use, divided by conditions. Additionally, I calculated students' *proactive diagram use proportion* by dividing the amount of proactive diagram use by the amount of spontaneous diagram use.

I ran three separate independent two-sample t-tests to investigate the effect of the intervention on students' use of diagrams. In each model, one of the three diagram use measures was added as a dependent variable. In all models, condition (coded as Metacognitive = 1, Non-Metacognitive = 0) was added as an independent variable. These t-tests showed that students in the Non-Metacognitive condition used diagrams more spontaneously ( $t(135.3) = 2.03, p = .04$ ). No statistically significant difference was found on other diagram use measures.

**Table 21.** Students' mean spontaneous diagram use, proactive diagram use, and proactive diagram use proportion per step. Standard deviations in parentheses.

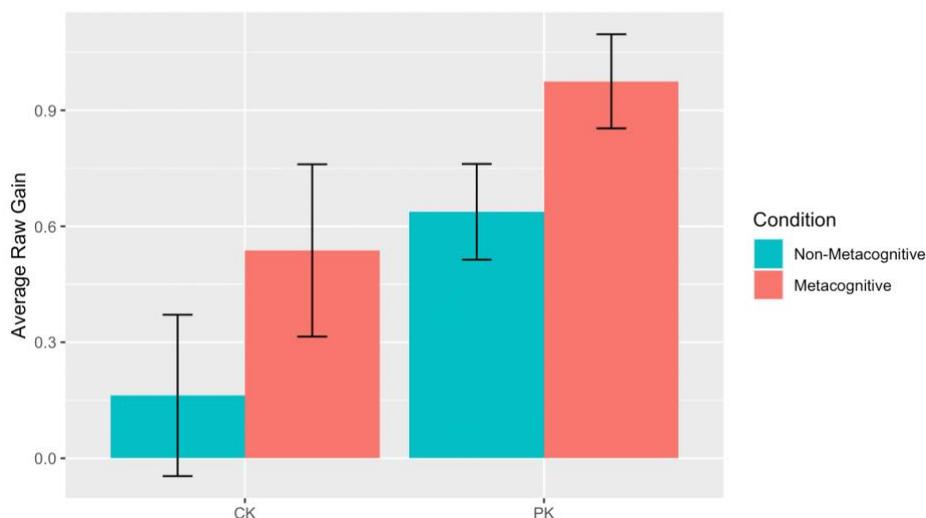
	<b>Spontaneous diagram use</b>	<b>Proactive diagram use</b>	<b>Proactive diagram use proportion</b>
Metacognitive	.12 (.10)	.08 (.06)	59.1%
Non-Metacognitive	.16 (.12)	.10 (.06)	62.6%

*RQ2: Will the metacognitive intervention help learners achieve greater performance and learning?* Two researchers separately coded student responses to the procedural items for their correctness with a high inter-rater reliability (*Cohen's kappa* = .94). Incomplete items were coded as incorrect. Table 22 shows students' average pretest and posttest scores on conceptual and procedural knowledge items (standard deviations in parentheses).

**Table 22.** Students' mean pretest and posttest scores and (standard deviations).

	<b>Conceptual knowledge</b> (max = 16)		<b>Procedural knowledge</b> (max = 5)	
	Pretest	Posttest	Pretest	Posttest
Metacognitive	9.78 (2.72)	10.48 (2.49)	1.74 (1.74)	2.88 (2.03)
Non-Metacognitive	9.28 (2.22)	9.38 (2.15)	1.73 (1.74)	2.35 (1.88)

To examine the treatment effect on student learning, I ran two separate linear regressions with condition as an independent variable (Metacognitive = 1, Non-Metacognitive = 0), and conceptual knowledge and procedural knowledge were added as a dependent variable, respectively. In both models, students' pretest score (on conceptual knowledge and procedural knowledge, respectively) was added as a covariate to control for students' incoming knowledge before the study. Students in the Metacognitive condition learned greater conceptual knowledge ( $\beta = .95$ ,  $t(137) = 2.52$ ,  $p = .01$ ) and procedural knowledge ( $\beta = .53$ ,  $t(137) = 2.08$ ,  $p = .04$ ) than those in the Non-Metacognitive condition (Figure 45).



**Fig. 45.** Students' learning gains from pretest to posttest on conceptual knowledge (CK) and procedural knowledge (PK) in each condition.

I also looked at students' test scores only on problems that involved tape diagrams (Table 23) and ran linear regressions with the same predictors. Models showed that students with the metacognitive intervention achieved greater gains on procedural items with tape diagrams compared to those without the intervention ( $\beta = .24$ ,  $t(137) = 2.07$ ,  $p = .04$ ) but did not show greater significant gains on conceptual items with tape diagrams ( $\beta = -.03$ ,  $t(137) = -.38$ ,  $p = .70$ ).

**Table 23.** Students' mean pretest and posttest scores on items with tape diagrams. Standard deviations in the parentheses.

	Conceptual knowledge (max = 3)		Procedural knowledge (max = 2)	
	Pretest	Posttest	Pretest	Posttest
Metacognitive	1.42 (.58)	1.26 (.53)	.59 (.77)	1.10 (.89)
Non-Metacognitive	1.45 (.65)	1.31 (.50)	.55 (.73)	.86 (.78)

Next, to examine the treatment effects on students' problem-solving performance in the tutor, I used the following performance measures from the tutor log data: frequency of hint use (i.e., the number of hint requests per symbolic step) and error rate (i.e., the number of incorrect attempts per symbolic step) (Table 24). We also counted the number of problems students solved in the tutor in each condition (Table 24). These measures are typical measures examined in ITS literature (e.g., Long & Aleven, 2014). We constructed three separate linear regressions with condition as an independent variable, and the frequency of hint use per step (hint request rate), incorrect attempts per step (error rate), and the number of problems solved were added as a dependent variable, respectively. In all models, students' overall pretest score was added to control for their prior knowledge. The models revealed no effect of the intervention on the hint request rate ( $\beta = -.001$ ,  $t(137) = -.60$ ,  $p = .54$ ), the error rate ( $\beta = .09$ ,  $t(137) = 1.30$ ,  $p = .20$ ), and the number of problem solved ( $\beta = .67$ ,  $t(137) = .32$ ,  $p = .75$ ).

**Table 24.** Students' average problem-solving performance in the tutor across the conditions. Standard deviations in parentheses.

	<b>Average number of hint request per symbolic step</b>	<b>Average number of incorrect attempts per symbolic step</b>	<b>Number of problems solved</b>
Metacognitive	.003 (.005)	.43 (.55)	32.64 (13.52)
Non-Metacognitive	.003 (.005)	.36 (.20)	31.25 (13.10)

*RQ3: Will the metacognitive intervention help learners perform better on a future learning task in a transfer environment with diagrams?*

Students' responses to each of the transfer tasks (initial proactive choice on whether to use diagrams or not, constructing an equation, and solving the equation) were independent of each other and coded as a binary variable. Two researchers coded students' responses with a high inter-rater reliability (*Cohen's kappa* = .89). Table 25 shows the number of students who successfully solved these transfer tasks and those who did not. We constructed three logistic regressions with students' score (successful = 1 or unsuccessful = 0) for each of these transfer tasks separately. In each model, condition was added as an independent variable. The models showed no significant effect of the treatment on any of the transfer task performance. However, due to the misleading problem text

mentioned earlier, this result does not give any conclusive evidence about the effect of the intervention on a transfer task.

**Table 25.** Number of students in each condition who successfully achieved each of the transfer tasks (and proportions).

	<b>Intention to use diagrams</b>	<b>Construct an equation</b>	<b>Solve the problem</b>
Metacognitive	48 (69.66%)	5 (7.25%)	5 (7.25%)
Non-Metacognitive	48 (67.61%)	4 (5.63%)	2 (2.82%)

*RQ4: Will the metacognitive intervention help learners perceive the usefulness of using diagrams?*

The average of each student's responses to the five survey questions was calculated (range: 1-5). Table 26 shows students' average ratings for the survey on pretest and posttest. We ran a linear regression with condition and the average survey rating on the pretest as predictors and the average score on the posttest as the dependent variable. Students' average rating on the pretest significantly predicted their rating on the posttest ( $\beta = .51$ ,  $t(130) = 5.67$ ,  $p < .01$ ) but the condition did not predict the rating.

**Table 26.** Students' average rating on the usefulness of using diagrams on pretest and posttest (standard deviation in parentheses)

	<b>Pretest</b>	<b>Posttest</b>
Metacognitive	2.99 (.89)	3.32 (1.11)
Non-Metacognitive	3.07 (.90)	3.39 (1.00)

*RQ5: How is diagram use associated with learning outcomes?*

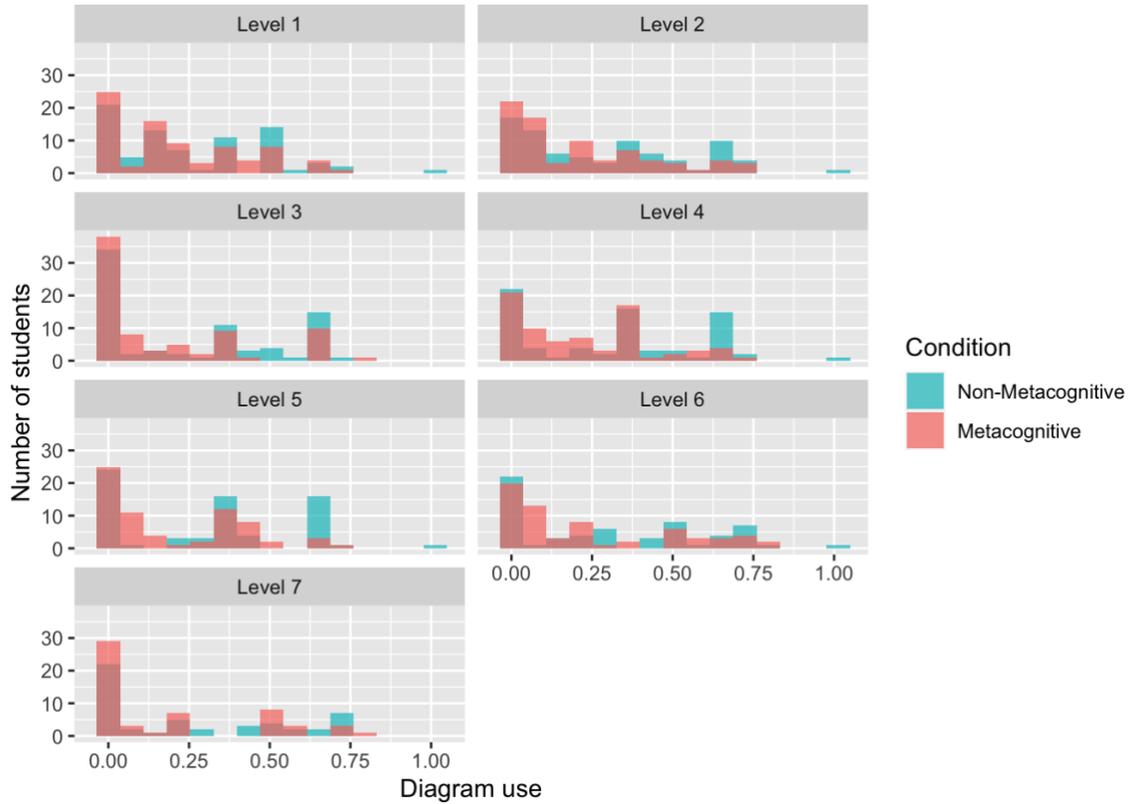
To further get insights into the relations between students' diagram use and their learning, we conducted several correlational analyses. Correlational analyses are typically used in examining associations between learner choices and outcomes (Cutumisu et al., 2019). We separately conducted correlational analyses for the Metacognitive and the Non-Metacognitive conditions. First, in the Non-Metacognitive condition, we found that students' diagram use, proactive diagram use, and proactive diagram use proportion were not significantly associated with students' pretest knowledge nor their learning gains from pretest to posttest. In other words, students in the Non-Metacognitive condition

used diagrams regardless of their prior knowledge and their use of diagrams did not have any relations with their learning. On the other hand, in the Metacognitive condition, students with lower prior knowledge (as measured by the pretest) tended to use diagrams more frequently both in a spontaneous ( $r = -.47, p < .01$ ) and a proactive manner ( $r = -.41, p < .01$ ). Students with lower prior knowledge also tended to have higher proactive use proportion (i.e., when they used diagrams, they tended to use them proactively,  $r = .43, p < .01$ ). Also, students who tended to use diagrams proactively (than reactively) gained more (conceptual and procedural) knowledge from pretest to posttest ( $r = .28, p = .02$ ). None of these relations were observed for students in the Non-Metacognitive condition.

#### *When did students use diagrams?*

Next, I conducted a further investigation in tutor log data. Specifically, I investigated when students used diagrams in the tutor and how that might differ across conditions. As the investigation so far only looked at students' diagram use that was averaged across all problems and problem levels, findings so far do not reveal when students used diagrams, which may be a critical factor in students' learning processes. Please note, however, that the following analyses do not involve any testing of the statistical significance of any differences, and they are rather descriptive and observational differences.

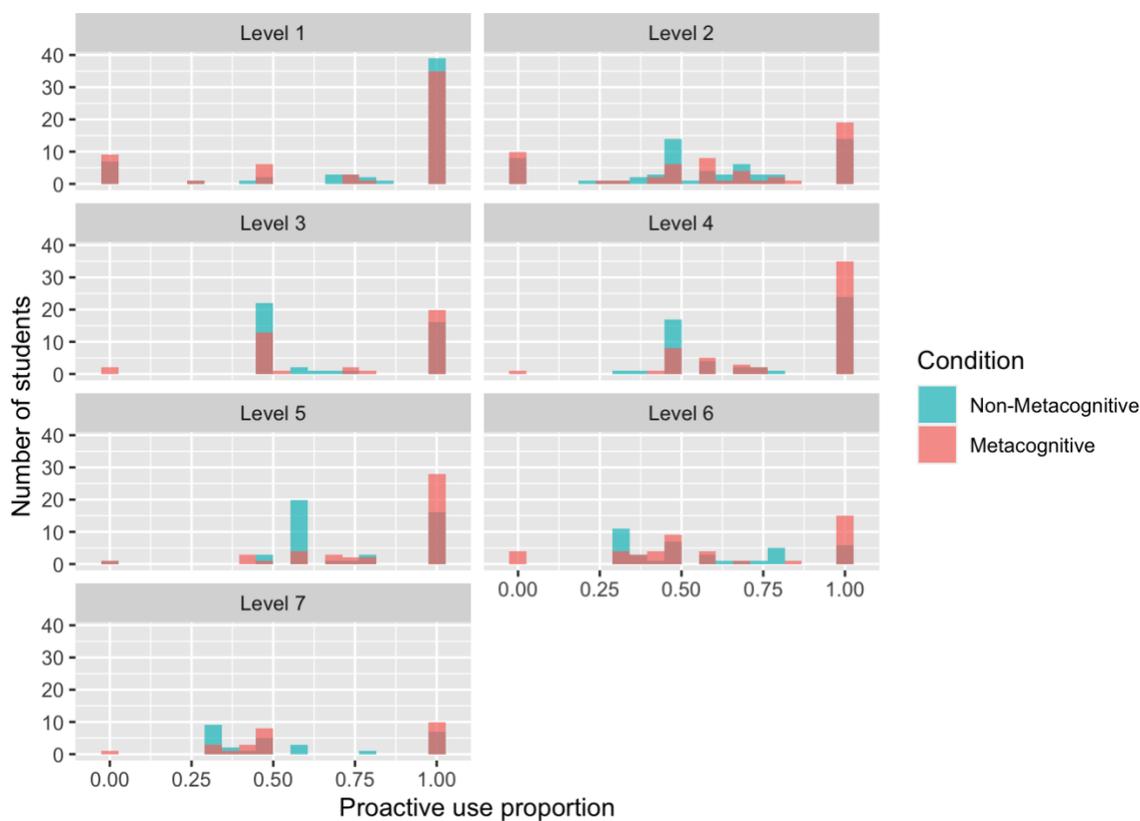
I first took a detailed look at patterns of diagram use by students in each problem set (and in each condition) separately. Figure 46 is a histogram of spontaneous diagram use rate in each problem set. As a general pattern, it is observed that many students in both conditions used diagrams on 50% or less of the given number of steps where they could have used diagrams in the tutor. Also, in both conditions, for the first two levels, the histograms are rather *flat*, indicating that there are not many users who very frequently used diagrams (although there is one student who used diagrams 100% of the time in the Non-Metacognitive condition). From Level 3 up to and including Level 5, a distinctive pattern emerges where about 1/4 of students in the Non-Metacognitive condition ( $n = 15$ ) used diagrams more often than 50% of the time (about 60%). No such pattern was observed for students in the Metacognitive condition (except some students in Level 3).



**Fig. 46.** A histogram of spontaneous diagram use in each problem level. We only examined Levels 1-7 because Levels 8-12 were extra problem sets with problems from Levels 1-7.

When it comes to proactive use proportion (i.e., “how much of diagram use is proactive?”), I found another interesting pattern. Figure 47 shows the same graph as Figure 46 but the x-axis is set differently. In this graph (Figure 47), the x axis represents proactive use proportion (i.e., proactive diagram use divided by spontaneous diagram use). Therefore, students who are at the very far end on the right (proactive use proportion = 1.0) always used diagrams proactively when they used diagrams. Students on the left side of the graph tended to use diagrams reactively. As shown in Figure 47, across all problem levels except Level 1, more students in the Metacognitive condition used diagrams proactively all the time (proactive use proportion = 1.0) whereas some students in the Non-

Metacognitive condition (about 1/4 of the students) used diagrams proactively only about 50% of the time (the other 50%, for these students, is reactive diagram use).

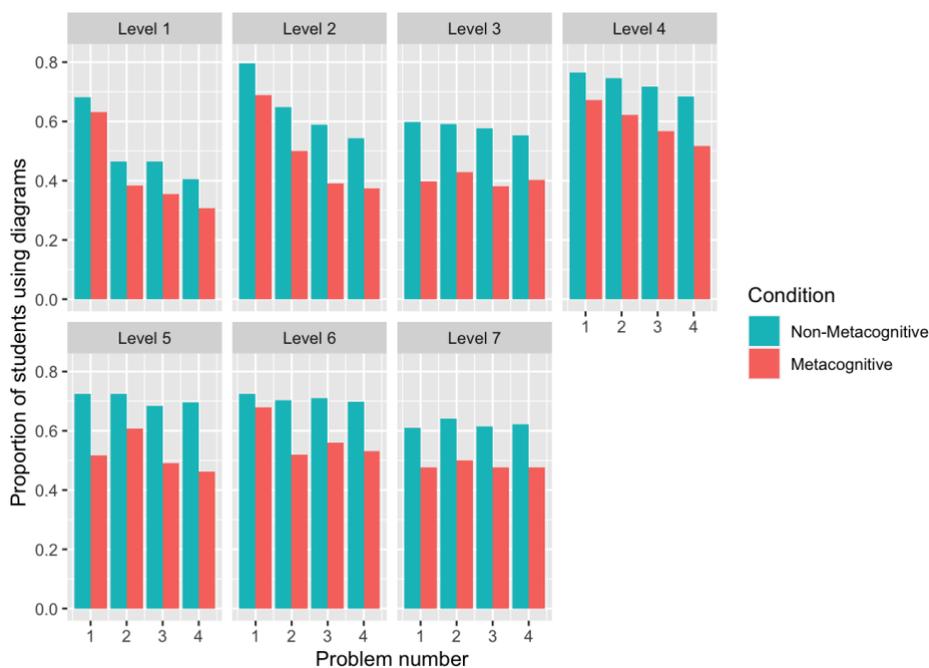


**Fig. 47.** A histogram of proactive use proportion in each problem level. We only examined Levels 1-7 as Levels 8-12 were extra problem sets with problems from Levels 1-7.

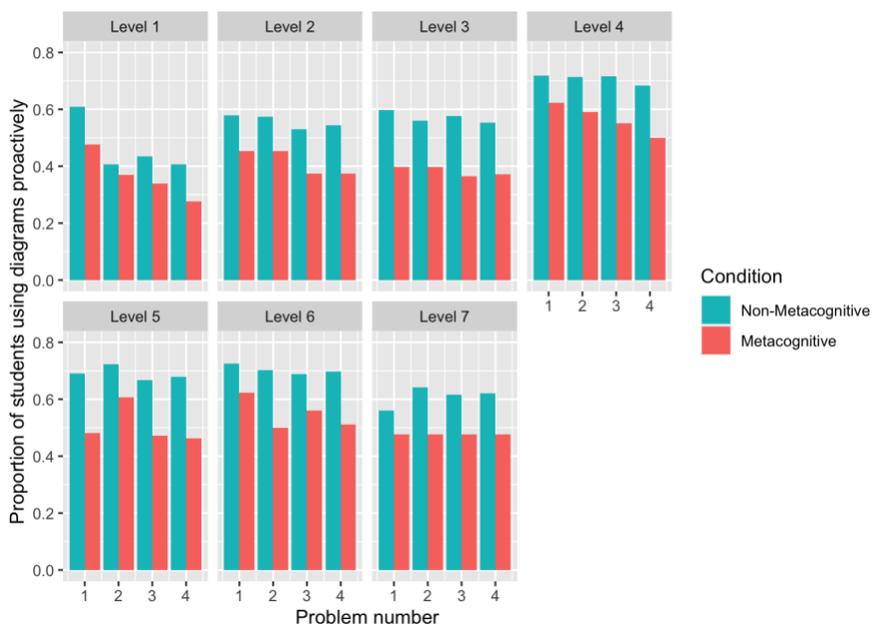
Further, I also investigated, within each level, how students' use of diagrams changed over time (on successive problems) and how the change patterns differ between the conditions, a critical aspect in examining self-regulated learning (Ben-Eliyahu & Bernacki, 2015). Recall, in each problem set, there were four problems, assigned in the same order for all students. These four problems in each problem set had the same difficulty (in terms of the structure of the equation) and therefore assumed to be the same in terms of the difficulty that students would perceive. Figure 48 shows how students'

spontaneous diagram use changed over time (i.e., over the four successive problems within each level) and Figure 49 shows how students' proactive diagram use changed over time. Figures 48 and 49 both illustrate the general trend (reported above) that more students in the Non-Metacognitive condition used diagrams than in the Metacognitive condition (and proactively).

However, students in the Metacognitive condition seemed to choose to use diagrams *strategically*, or choose to use diagrams when the use of diagrams would be the most helpful. Of the Levels 1-7 in the tutor, Levels 1, 2, 4, and 6 were the problem sets where a new problem type was introduced (i.e., Levels 3, 5, and 7 were the repetition of the Levels 2, 4, 6, respectively, with the same equation type but different numbers). In both Figures 48 and 49, a trend is observed that students in the Metacognitive condition tended to use diagrams (both spontaneously and proactively) at the beginning of Levels 1, 2, 4, and 6 whereas they did not request diagrams much on Levels 3, 5, and 7. Students in the Non-Metacognitive condition do not show the same pattern.



**Fig. 48.** Patterns of students' spontaneous diagram use within each problem level. It shows the number of students (in proportion) who used diagrams at least once for each problem in each problem set.



**Fig. 49.** Patterns of students' proactive diagram use within each problem level. It shows the number of students (in proportion) who proactively used diagrams at least once for each problem in each problem set.

Furthermore, in both graphs (but more pronounced in Figure 49 for proactive use), students in the Metacognitive condition showed a decreasing pattern of diagram use in Levels 1, 2, 4, and 6 but not in other levels. This trend may indicate that many students with the metacognitive intervention used diagrams when they saw the new problem type, but they gradually chose not to use the visual scaffolding. Because Levels 3, 5, and 7 had the problem types that they had practiced already, many of the students did not use the visual scaffolding for the first problem. On the other hand, those in the Non-Metacognitive condition do not show such a decreasing trend. Although there is a decreasing use pattern in Level 1, students consistently requested the visual scaffold within each problem level in other problem levels.

## Discussion

In Study 8, I conducted an experimental study with 140 students at two schools in the U.S, investigating the effects of a metacognitive intervention, designed with input from students from the

target population, on students' strategic choice-making behaviors, domain-level learning, performance in the tutor, and on learning transfer.

First, for our RQ1 (i.e., *will the metacognitive intervention help learners use diagrams proactively during learning, when it is provided as an optional resource?*), the results revealed that students who received the metacognitive intervention tended to use diagrams less frequently overall compared to those who did not receive any interventions. On proactive use, there was no difference between the conditions. However, correlational analyses conducted separately for each condition revealed that the metacognitive intervention affected students' diagram use differently; we found that, in the Metacognitive condition, students' pretest score was negatively associated with more frequent and proactive use of diagrams (i.e., students with lower prior knowledge tended to use diagrams frequently and proactively). Such associations were not observed for students in the Non-Metacognitive condition. These findings may imply that the metacognitive intervention effectively facilitated students' diagram use depending on their prior knowledge level. Further, observational log data analyses showed that students with the metacognitive intervention seemed to choose to use diagrams *strategically*; specifically, analyses of students' choice-making processes in each problem level found that students with the metacognitive intervention tended to use diagrams when they were given a new problem type, but decreased diagram use when they saw the same problem type repeatedly. Considering that a strategic diagram user would use diagrams when they need it as scaffold (i.e., when they have low knowledge to try a given task, or when they see a new type of problem), the finding indicates that the metacognitive intervention helped students become more strategic. This view of the strategic use of diagrams is also well aligned with the theoretical view discussed in Chapter 2. In particular, it indicates that students with the metacognitive intervention engaged in self-assessment of their own knowledge and task (e.g., "Is this problem familiar?") to decide whether or not to use diagrams for each problem. Of course, it is possible that having the dashboard at the end of each problem level might have led students' propensity to use diagrams at the beginning of the next level (rather than careful self-assessment of the task), but it does not explain students' low diagram use when the problems of the same type as those in the previous level repeated. However, we acknowledge that these observational analyses are not as rigor as statistical significance testing.

*Why did students in the Metacognitive condition use diagrams less frequently?* It is interesting to ask why the metacognitive interventions did not simply increase the use of diagrams in the tutor. I argue that this result can be attributed to the design of the intervention elements. As described in the Metacognitive Behavior Change Model, the intervention elements (i.e., the tutorial, the adaptive recommendations, and the dashboard) were designed to promote deeper thinking about whether and

when to use the target strategy (i.e., whether or not to use diagrams in specific situations), instead of merely to increase the use of the strategy, which is the main target of many behavior change models. For instance, for some students in the study, the dashboard may have shown a graph suggesting that the use of diagrams is not helpful for that student. The dashboard would not encourage students to use diagrams but rather encouraged students to think about the diagram use (e.g., “Do you think that diagrams are helpful?”). This descriptive (as opposed to prescriptive) nature of the dashboard may have contributed to these students’ strategic decisions regarding when *not to use* diagrams in the tutor, and, as discussed in the previous paragraph, may have led to the strategy of “using diagrams when I need help.” In contrast, students without the metacognitive intervention were not prompted to think about diagram use, which may have led to the use of diagrams that had no consistent patterns.

For RQ2 (i.e., *will the metacognitive intervention help learners achieve greater performance and learning gains?*), we found that students who received the metacognitive intervention achieved better conceptual and procedural knowledge gains from pretest to posttest. The study did not find any effects of the metacognitive intervention on students’ equation-solving performance in the tutor (i.e., hint use and error rate). There was also no significant difference in the number of problems solved by the students in the Metacognitive condition ( $M = 32.64$  problems in the tutor  $SD = 13.52$ ) compared to the Non-Metacognitive condition ( $M = 31.25$ ,  $SD = 13.10$ ). This lack of difference (regarding the number of problems solved) is notable, especially if one considers that the students in the metacognitive condition had to additional mandatory work on the dashboard in each level and the tutorial unit (which we did not count as problems in the analysis, also see Figure 42).

*Why did students in the Metacognitive condition achieve higher learning gains?* A critical question given the findings on learning is how the intervention helped students in learning conceptual and procedural knowledge. The only association found between diagram use and learning was the correlation between students’ learning gains and students’ proactive diagram use proportion in the Metacognitive condition (i.e., students who used diagrams more proactively tended to learn more, regardless of how frequently they used, RQ5). However, descriptive differences discussed above suggest that students’ choice-making strategies (i.e., using diagrams more often when solving an equation of a new type and less often when seeing the same type of equation afterwards) may have contributed to their learning. For instance, it maybe that, by choosing to engage with diagrams when they need the scaffold, students engaged with the diagram steps more carefully. They might have read the conceptual feedback messages given for those steps as well, which students in the Non-Metacognitive condition might have missed (as they might have used diagrams more than the amount they would have *needed*). The decreasing pattern in diagram use also implies that students were correctly recognizing the type of equations each time (even in Levels 3, 5, and 7, which, by their

names, might indicate new, more challenging problem types) and decided whether to request diagrams accordingly. It is possible that students with the metacognitive intervention attended to the problem structure more carefully and correctly assessed it, and therefore were able to make conceptual connections between diagrams and symbolic representations (if they did not correctly assess problem difficulty, error rate would go higher, given results from prior studies, Nagashima, Tseng et al., 2021). For procedural knowledge, an interpretation could be made that students with the metacognitive intervention sometimes proactively chose to practice their procedural skills without the visual scaffolding in the Choice-based Diagram Tutor, which many students did not do in the Non-Metacognitive condition. In other words, students in the Non-Metacognitive condition might have relied too much on the help of diagrams rather than appropriately choosing *not to use* diagrams when they could have tried solving equations without the visual help (Binbarasan-Tüysüzoglu & Greene, 2015).

Also, the fact that we found no differences on students' performance measures in the tutor illustrates that, even with less use of diagrams, students in the Metacognitive condition did not make significantly more errors than those in the Non-Metacognitive condition and the number of problems they solved over the duration of the study was also not reliably different between the conditions. This finding may imply that students with the metacognitive intervention were able to appropriately make use of the visual scaffolding (i.e., used when needed, and tried solving equations without diagrams when they correctly thought they could).

Our RQ3 asked if *the metacognitive intervention helps learners perform better on a future learning task in a transfer environment with diagrams*. The results did not show any significant difference between the conditions on any of the transfer tasks. However, as mentioned previously, the problem text had a misleading statement. Hence, the experiment does not give a conclusive finding for RQ3.

For RQ4 (i.e., *will the metacognitive intervention help learners perceive the usefulness of using diagrams?*), we did not find a statistically significant difference between the conditions. This result might have some connections with the strategic use of diagrams in the Metacognitive condition; it might be that students in the Metacognitive condition proactively chose to use and *not to use* diagrams depending on the situation, therefore the behavior might not have led to a higher score on the usefulness of using diagrams (i.e., it may be that they also appreciated the opportunities where practicing equation solving without diagrams can be beneficial).

In sum, this study offers important insights into students' strategic choice-making behaviors in using visual representations during problem solving. Most importantly, the study shows that an intervention aimed at promoting self-regulated learning can facilitate domain-level learning. This

effect on domain-level learning has been rarely achieved even when the intervention is focused on fostering domain learning (but see Rittle-Johnson et al., 2016), and more so when the intervention is focused on self-regulation. Although how that happened is still to be explored, there were notable differences in how and when students used and did not use diagrams, which might have contributed to the strategic use of diagrams in the tutor, as discussed above. This finding illustrated an importance of examining student behavior of *choosing not to use* visual representations during students' self-regulated learning. This finding adds new knowledge to the literature in the area of learning with visual representations; as opposed to a previously-known principle that argues for the importance of spontaneous use of visuals, it exemplifies the importance of conducting a deeper investigation of students' choice-making behaviors regarding when to use and when not to use visuals. In this study, such an investigation was possible because of the tutor log data from the Choice-based Diagram Tutor. Past research on diagram use has rarely conducted such deeper investigations, mostly because of the lack of appropriate assessment environment for that purpose (Uesaka et al., 2010) but even with a technology-mediated assessment environment, a choice of not using a strategy has rarely received attention (e.g., Cutumisu et al., 2019).

Some limitations of the study are worth noting: First, the study tested the effect of the metacognitive intervention on a specific strategy of anticipatory diagrammatic self-explanation. It is unclear whether the findings from this study would be generalized in other environments and other types of visual use. Also, the study does not help us understand individual effects of the intervention elements separately. I plan to conduct further analyses to examine how individual components associate with students' diagram use and learning outcomes (e.g., coding students' responses to the dashboard prompts). Finally, how the effect of the intervention on domain-level learning was mediated is still unclear. Future analyses will make further deeper investigations.

## 5 Conclusion

### 5.1 Concluding Thoughts

One ultimate goal of education is to foster autonomous, self-regulated learners who can learn on their own. Such learners would strategically use available resources around them to keep updating their knowledge and skills. The importance of fostering such self-regulated learners is growing rapidly in our current society with its massive flow of many types of information (Pendleton-Jullian & Brown, 2018). One would hope, of course, that such self-regulated behaviors and dispositions will also lead to greater domain-level learning. In other words, we want to foster learners who can autonomously learn on their own so that they can learn the most out of the given opportunities and seek new opportunities. Through a series of design research and experiments in US schools, my dissertation makes an attempt at investigating and supporting this challenging goal in the context of using visual representations during math problem solving practice.

Initially, my research focused on learning *with* visual representations. This is an important cognition-focused research topic that has been researched in many relevant studies in various domains. My entire work from a design study with teachers (i.e., Pedagogical Affordance Analysis) to multiple experiments examining the effect of visual scaffolding (i.e., confirmatory and anticipatory diagrammatic self-explanation) on students' cognitive learning informed many insights into how to scaffold algebra learning. Specifically, these studies show that interactive visual scaffolding designed with users can support efficient and effective learning in the domain of early algebra, a challenging subject known as a “gatekeeper” subject. Across my studies, it has been observed and suggested that the interactive diagrammatic self-explanation, through its scaffolded interactions (e.g., contrasting cases), help students visually and conceptually understand how algebraic equations transform and associated conceptual principles (“conceptual knowledge that underlies procedures”, Crooks & Alibali, 2014). Even though diagrammatic self-explanation creates additional steps (almost doubling the number of steps required in each problem), this interactive scaffold made problem solving efficient with fewer hint requests and lower error rate. These studies overcame challenges of designing self-explanation activities with visual representations by providing an interactive, effective learning experience with diagrams in a way that is not cognitively demanding (Wu & Rau, 2018).

In the later studies, the focus of the intervention shifted to supporting not only domain-level learning (i.e., conceptual and procedural knowledge) but also metacognitive processes that may play an important role in learning from visual representations. Using the Metacognitive Choice Behavior

Model as a design guide, I designed an intervention package that was aimed at supporting important stages of students' choice making in learning with my intelligent tutor. This metacognitive intervention helped students use diagrams strategically, and led to enhanced conceptual and procedural learning, important yet challenging dual goal in algebra instruction (Rittle-Johnson, 2016; Rittle-Johnson & Alibali, 1999). How could the metacognitive scaffold *indirectly* address this challenging goal of enhancing both conceptual and procedural knowledge?

An important perspective to consider may be that learning is a multi-faceted human activity. Studies on algebra instruction typically test an instructional strategy and technology of interest against others, and sometimes with flexible mixture of multiple strategies (Fyfe et al., 2015; Ottmar & Landy, 2017; Rittle-Johnson & Star, 2007). Discussion from such studies (including my prior studies) tend to focus on further supporting learners by manipulating factors associated with cognitive aspects of learning, such as increasing intervention time so that learners can be scaffolded for a longer time, changing the order of providing scaffolding, how much scaffolding to give, and when to fade the scaffold. However, as my last study showed, students with less use of the visual scaffolding could learn better conceptual and procedural knowledge with no detrimental effects on problem-solving performance during a learning activity. This effect could hardly be attributed only to the amount of dose of the scaffolding. Rather, it could be due to students' enhanced self-regulated learning processes, metacognition, engagement, and/or agency that affected their domain-level learning. For instance, it could be that students with the intervention were more engaged in their learning experience (e.g., by seeing their own performance with a graph with badges), which may have helped them attend to conceptual aspects of diagrams (with focused practice) instead of just trying to *get by* through shallow processing of diagrams (e.g., selecting diagrams that *look* right). Such non-cognitive activities might have been enabled by the combination of the learning environment where students had the control over choice and the metacognitive intervention (both of these components are needed because, if there was no choice, there would be little that they could adjust after reflecting on their own performance, Long & Alevan, 2017). However, of course, this is just an example; other interpretations would be possible. For instance, students in the Metacognitive condition may have used diagrams when the use of diagrams was the most helpful.

This complexity around how the indirect effect occurred needs to be investigated further, and I do not have a clear answer for it yet. Follow-up experiments could show interesting insights that may help us understand the mechanism better. For example, it would be interesting to test whether strictly structuring the visual scaffolding (with no choice and the metacognitive intervention) in a way that was suggested effective in Study 8 (i.e., use visuals at the beginning for each equation type, and gradually fade the scaffolding) would still help students learn (Fyfe et al., 2015). Such a study would

help us understand if there is any “optimal” sequencing of scaffolding support (Renkl et al., 2004) in the context of early algebra. Or, we may better forget the idea of prescribing “optimal” scaffolding to learners and instead focus our efforts on designing for multi-faceted aspects of learning so that learners can be effectively supported to find their own optimal learning path.

## **5.2 Limitations**

I acknowledge that my studies have several major limitations. First, all of my studies used a specific type of visual representations (i.e., tape diagrams). As other types of visual representations would have different pedagogical affordances, it is unclear whether findings would generalize to other types of visual representations. Second, relatedly, the design of confirmatory and anticipatory diagrammatic self-explanation activities might look different if they are designed with other types of representations and in other domains. Third, the choice pattern analysis reported the last study (Study 8) only offers observational insights and it needs to be validated with statistical analyses. Fourth, other interpretations are possible for the results of my studies. For instance, in Study 6, I argued that proactive diagram use represents students’ self-regulated diagram use whereas reactive diagram use would indicate un-planned diagram use; however, it is possible that reactive diagram use would involve meaningful self-regulated learning processes (e.g., realizing their own knowledge gap by making mistakes). Finally, Studies 3-8 were conducted remotely during the COVID-19 pandemic between 2020 and 2022 (Nagashima, Yadav, & Alevan, 2021). Therefore, it is unclear if the main findings would hold in *normal* classroom settings. For instance, the way teachers interact with students would differ between in-person and remote teaching, and it might alter the way interventions influence student learning and interactions with my tutors.

## **5.3 Contributions**

My dissertation makes several contributions in the broad field of the learning sciences. In the following, I provide major contributions that this entire dissertation makes to this field.

### **Contributions to research on instructional design**

Instructional design is a field of study that designs and tests instructional activities, materials, and technological programs towards the goal of supporting student learning. However, recent critiques

argue that traditional models, such as the ADDIE model (Croxtton and Chow 2015; Trust and Pektas 2018) and 4C/ID model (Güney 2019), have been used to *prescribe* what to do during instructional design processes, causing users' over-reliance on instructional models (e.g., "I am following the model, so I should be doing the right thing") (Stefaniak & Xu, 2020). Such over-reliance does not help users incorporate contextual factors or other practitioners' perspectives into the design process (Stefaniak & Xu, 2020).

My dissertation addresses this problem. Specifically, Pedagogical Affordance Analysis (PAA) provides one approach to overcoming this issue in the field. PAA offers a *bottom-up* approach in (re)designing instructional tools and activities through systematically inviting and integrating educators' viewpoints and pedagogical practices. The idea of "pedagogical affordance" had also gained little attention in the field in the past, until PAA. My dissertation offers both theoretical and practical contributions to the area of instructional design: It provides 1) new understanding of the notion of pedagogical affordance as a relational concept (rather than a fixed concept that can be derived from an instructional tool, which has been the main conceptualization of pedagogical affordance) and 2) a systematic method that can be applied when examining or re-designing an instructional tool.

### **Contributions to research on Intelligent Tutoring Systems**

To date, a number of Intelligent Tutoring Systems (ITSs) have been designed, developed, and evaluated at scale (Aleven et al., 2009; Koedinger & Corbett, 2006; Kulik & Fletcher, 2016). Through their step-by-step guidance, ITSs have not only achieved effective learning in many subject domains (Graesser et al., 2001; Melis & Siekmann, 2004; Steenbergen-Hu & Cooper, 2013) but also supported metacognitive learning or self-regulated learning (Azevedo, 2005; Long & Aleven, 2017; Roll et al., 2011). However, to further enhance the effectiveness of ITSs, it is important to design and develop ITSs following user-centered design approaches. In fact, it is not very common that ITSs are designed with users, especially when it comes to metacognitive features in ITSs (but see Long et al., 2015).

My dissertation shows that an intelligent tutor and metacognitive support in the tutor designed with users (students and educators, Studies 1 and 7) can help achieve challenging instructional goals. Specifically, I followed user-centered design approaches in designing diagrammatic self-explanation and the metacognitive intervention. The design of these interactions and features that I had in the tutor would not have existed if I had not started the design with users. These designs, implemented in the equation solving tutor, effectively supported student learning. Therefore, my work makes a

contribution that user-centered design approaches can be effectively used to achieve challenging learning problems (supporting cognitive and metacognitive learning).

### **Contributions to diagram research**

My dissertation offers contributions to the literature on diagram research. In diagram research, past studies have focused on examining and supporting spontaneous use of diagrams during problem solving (Ainsworth et al., 2011; Manalo & Uesaka, 2012; Uesaka et al., 2012; Uesaka et al., 2010; Wu & Rau, 2018). A main approach employed in these past studies to assess students' spontaneous use of diagrams was to examine whether or not and how frequently students drew diagrams on paper or through self-report data. Due to the limited nature of these assessments, prior work on diagram use was not able to get data on *when* and *how* students use diagrams during problem solving.

Through the design and implementation of the Choice-based Diagram Tutor, my research addressed this challenge. Using tutor log data, Studies 6 and 8 investigated not only students' spontaneous use of diagrams but also proactive and reactive use of diagrams, as well as how their use patterns changed over time. This rich, fine-grained data suggested that self-regulated, strategic users of diagrams would choose to use the visual scaffold *when they need it*. My dissertation offers theoretical and technical contributions of 1) providing theoretical knowledge that not only a mere use of diagrams (i.e., frequency of diagram use) but how students use diagrams (i.e., choose not to use when unnecessary) is important when students use diagrams during problem solving practice, and 2) developing Choice-based Diagram Tutor, a choice-based learning environment (Cutumisu et al., 2019; Schwartz & Arena, 2013) that captures students' choice making behaviors in using diagrams.

### **Contributions to research on mathematical cognition**

It is essential that students learn both conceptual and procedural knowledge in mathematics, especially in early algebra which is often called a “gatekeeper” course to many advanced STEM (Science, Technology, Engineering, and Mathematics) courses (Spielhagen, 2006). Instructional principles have been developed towards this goal. For instance, studies have shown that comparing between multiple solution strategies (Rittle-Johnson & Star, 2009), self-explanation (Booth et al., 2013), and practice on “non-traditional” problems (e.g.,  $7 = 4 + 3$ ) (McNeil et al., 2011) can foster conceptual knowledge in algebra. Recent research also shows that playful learning technologies can contribute to students' conceptual learning in algebra (Chan et al., 2022; Nagashima et al., 2022). For procedural knowledge, intelligent tutoring software has been shown effective (e.g., Long &

Aleven, 2017). Yet, studies have not been successful at fostering both conceptual and procedural knowledge in algebra, except Rittle-Johnson et al. (2016), in which they found that direct instruction on mathematical equivalence had significant effect on both conceptual and procedural learning, compared to combining conceptual and procedural instruction. However, it has been challenging to understand the mechanism through which conceptual learning and procedural learning is supported (Rittle-Johnson, 2019).

My dissertation achieved the challenging goal of supporting both conceptual and procedural knowledge in algebra. It also shows results that suggest how such a learning effect was achieved. In particular, as discussed in the previous section, my study provides a hint suggesting that it may not necessarily be the *dose* of a scaffold that matters; rather, the use of a scaffold when it is needed (e.g., when solving a new type of problem) may be the key to achieving both conceptual and procedural learning through focused practice. This was only possible with detailed investigations using the log data.

### **Contributions to research on self-explanation**

My studies established new interactive forms of self-explanation with visual representations, namely, confirmatory diagrammatic self-explanation and anticipatory diagrammatic self-explanation. In these self-explanation activities, students would *explain* their problem-solving steps through the tape diagram representation. Previously, self-explanation with visual representations was limited to either referring to visual representations to aid problem solving (e.g., referring to diagram of the human circulatory system, in Ainsworth & Iacovides, 2005) or drawing visual representations when solving a problem (Wu & Rau, 2018, Zhang & Fiorella, 2021). The former type does not fully engage students with the visual scaffolding (students do not interact directly with the visual provided) and the latter type can be too cognitively demanding (Nagashima, Bartel et al., 2020; Scheiter et al., 2017). Such poor designs of self-explanation with visual representations pose a challenge to some groups of students (e.g., students with low spatial ability), failing to provide effective scaffolding for those who need support.

The interactive diagrammatic self-explanation approaches described in my dissertation address these challenges. In both confirmatory and anticipatory diagrammatic self-explanation, students are asked to select diagrams to solve problems with the target representation they are supposed to learn to use, and students are scaffolded well through the use of established strategies such as contrasting cases (Schwartz et al., 2013), hints, and feedback. Further, anticipatory diagrammatic self-explanation, through its unique, anticipative way of using visual scaffolding (c.f., Renkl, 1997), helped students both learn and perform well, a fundamental instructional design issue that previous

self-explanation studies have rarely addressed (Bisra et al., 2018; Rittle-Johnson et al., 2017, but see Alevén & Koedinger, 2002). Therefore, my interactive diagrammatic self-explanation activities offer both theoretical and practical contributions of 1) understanding whether and how interactive diagrammatic self-explanation activities help students learn and perform well and 2) providing practical instructional strategies that can be used (e.g., through Intelligent Tutoring Systems and potentially in other software) in classrooms.

### **Contributions to self-regulated learning and behavior change literature**

Studies on self-regulated learning have shown its importance in fostering autonomous learners. Empirical studies on self-regulated learning, particularly those conducted using an interactive learning environment, have provided detailed insights into how students make self-regulated decisions in a learning environment (Azevedo et al., 2009; Leelawong & Biswas, 2008; Roscoe et al., 2013). They also explore how students' self-regulated behaviors may or may not help with domain-level learning and performance (Kramarski & Gutman, 2005; Roll et al., 2011). Even with the abundant amount of research on this topic, however, it is still unclear how to promote both students' self-regulated behaviors and, as a result of the behaviors, students' domain-level learning and performance.

Behavior change research, originally from domains such as healthcare and criminology, offers similar but different approaches in assessing and supporting human behaviors (Consolvo et al., 2009). Many behavior change models propose a linear path from an unideal behavior to an ideal behavior, reducing their applicability to more complex behavior changes that are of interest in other fields, such as in education (Heimlich & Ardoin, 2008). Prior work has made attempts to apply behavior change models to learning research (e.g., Opezzo & Schwartz, 2013) but such new models still are not applicable to learning behaviors that involve complicated choice-making processes.

My dissertation research proposed the Metacognitive Behavior Change Model, a model that integrates behavior change models (Prochaska et al., 1992; Opezzo & Schwartz, 2013) and a theoretical framework on self-regulated learning (Zimmerman & Campillo, 2003) to allow for non-linear metacognitive behavior changes. Such a model did rarely exist in the literature, and it had been challenging to appropriately assess and support complicated choice-making processes that are important aspect of self-regulated learning. My research contributes a theoretical understanding and framework that illustrates the importance of investigating both people's choices of using and not using a target strategy during complex learning processes.

### **Contributions to research on choice making in learning**

As discussed in Chapter 2, in the field of the learning sciences (and several relevant fields), research has been conducted to assess and support students' choice making (Lust et al., 2011; Winters et al., 2008). Such work covers a wide variety of choices, including a choice of whether to start learning with a worked example (van Harsel et al., 2021), a choice of asking for help (Roll et al., 2011), a choice of using visual representations (Uesaka et al., 2007), and a choice of using design-thinking strategies (Chin et al., 2019). Although these studies offer interesting insights into whether and when learners make certain choices and some of them show how their choices relate to learning and performance, they fail to provide a detailed account of how students' choice behaviors may or may not change over time. As learners develop skills and knowledge while learning, it is reasonable to think that their use of certain strategies may also change over time (Azevedo et al., 2010; Greene et al., 2021; Roscoe et al., 2013). This view is also aligned with the view of self-regulated learning, in which learners would reflect on their own behavior and adjust their learning strategies. Previous studies, however, typically only use aggregated data (e.g., average or sum of the number of strategy use) to examine choice behaviors and therefore miss any such changes in students' choices. Part of this issue is due to technological capabilities of research (e.g., without appropriate technology, it is difficult to assess such changes).

My research looked at how students' choice behaviors changed over time in the Choice-based Diagram Tutor and found that students who received the metacognitive intervention showed a strategic choice behavior that suggests that they used diagrams when they needed it. This difference was masked when I only looked at the aggregated data; examining temporal data allowed me to capture this trend. Thus, my research makes theoretical and technological contributions: theoretically, it gives initial evidence that changes in how students make choices may be a key to better student learning. My research also offers the Choice-based Diagram Tutor, an environment that can assess students' choices and how the choices change over time through tutor log data.

### **Contributions to educational practice**

As a learning science researcher, it has always been my hope that my research offers something good for educational practices. The field of the learning science has long been exploring various ways to work with and support practitioners. For example, new methodological approaches such as Design-Based Research (Barab & Squire, 2004; Brown, 1992) and in-vivo classroom research (Koedinger et al., 2009) have been developed. These methodologies enable researchers and practitioners (e.g., educators, administrators) to work together to solve both scientific and real-world problems in educational settings. Others from relevant fields have tried different approaches, such as to offer

practical instructional guidelines on how to use certain scientifically-proven instructional strategies in teaching (Rittle-Johnson et al., 2020) and to use participatory approaches to actively involve practitioners in the design process (Ahn et al., 2019; Holstein et al., 2019; Prather et al., 2022).

Using these methodologies and through other ways, I have made several practical contributions during my PhD. First, through employing in-vivo classroom research (Koedinger et al., 2009) in my experiments, my research established scientific evidence in a classroom environment (as opposed to a more refined setting such as in a lab). In my studies, I have treated any classroom routines (e.g., teachers' interaction with students, students' seating locations) that would not directly affect research questions differently between conditions as part of the classroom environment. I have worked closely with teachers before, during, and after a study to reach a mutual agreement on who (e.g., which grade levels) is appropriate for the study, what content (e.g., types of equations) is both appropriate and beneficial, and what logistical procedure (e.g., how to structure a study schedule) would be the most preferred. In other words, to conduct a study in an environment that is as similar as possible to daily classroom settings, I value and prioritize teachers' (who often decides what to teach and how to teach in everyday classes) opinions when designing a study. This flexibility towards how to structure a study helped me conduct a study in a less-refined classroom environment. Rigorous studies conducted in such an environment have meaningful insights to offer for practitioners.

Still, research articles and findings are not accessible to many teachers for a lot of reasons. Teachers may not have time or knowledge (e.g., terminologies) to read scientific articles (Lortie-Forgues et al., 2021), they may not know where to find such articles, and most importantly, it would be difficult to come up with a clear idea of what they could do in their classroom on the following day after learning some new research evidence (Higgins et al., 2022; Rittle-Johnson et al., 2020). To provide established principles in an accessible form, several initiatives have been developed (e.g., What Works Clearinghouse's Practice Guides). Instead of merely sharing principles, I have developed and distributed research materials as Open Educational Resources (OER) that can be directly and freely used in everyday instruction. OER are teaching, learning, and research materials that are provided under an open license (e.g., Creative Commons licenses) (Hilton, 2016; Nagashima & Hrach, 2021; The William and Flora Hewlett Foundation, n.d.) to foster creative reuse, revision, and redistribution of educational content and practices. For instance, I created a "[tape diagram for equations](#)" template for teachers so that they could integrate tape diagrams I designed in their teaching materials (distributed using Google Slides – a tool that teachers use to share their instructional ideas and content, which I learned by working with teachers). Several teachers have used this template to create their own tape diagrams and teach with them. Furthermore, to provide a quick means to producing tape diagrams, research staffs and I developed, based on our diagram

tutors, [an intelligent tutor that can produce a tape diagram based on any equation that users input](#) (with some restrictions, such as not capable of creating tape diagrams with negatives). Users (teachers and students) could use the tool to quickly visualize any equations and visually understand how the equations are structured. These materials and tools offer immediate resources for teachers to use without spending a lot of time to read scientific articles and reports.

Finally, working with practitioners during the COVID-19 made me confirm that working with practitioners is a core part of my work as a learning science researcher. Reports have been published illustrating what it was like to suddenly switch to remote instruction and how difficult it has been for teachers to communicate with children and parents during the pandemic (Middleton, 2020; Patrick et al., 2020; Stelitano et al., 2021). By establishing a guideline for conducting classroom research remotely (Nagashima, Yadav, & Alevan, 2021), which focused on ensuring that practitioners (i.e., teachers and students) have a meaningful study participation experience, I have been able to partner with many schools and teachers across the country. This guideline focused on aspects of classroom research that have often been unnoticed or undiscovered by researchers, such as to provide opportunities for school students to interact with researchers (which many students would never have a chance to do) (Laursen et al., 2007). In other words, when I worked with practitioners, I carefully designed a “study participation experience” for the practitioners so that it would not end up being merely a data collection activity but rather a mutual learning experience where both researchers and practitioners can maximally benefit from classroom research. This initiative helped reach many practitioners who would otherwise not be able to experience study participation, such as those in low-income districts and a school located on an island in Hawaii. By remotely connecting with classrooms, I have been able to provide meaningful experiences to teachers and students in five states in the US while advancing the science of learning.

### **Broader implications**

With the increasing use of artificial intelligence technology in everyday life today, it has become more critical than ever that “whitewater kayakers” skillfully make good choices in their lives. Learners of the 21st century will need to discern between correct information and misinformation, strategically respond to various types of recommendations provided by technological systems, and also proactively seek information that they need to keep moving forward in the desired direction. In such a society, investigating and supporting people’s choices becomes one of the central goals of research on human behavior and learning. Deep investigations on people’s choices have a great impact and potential in informing future designs of choice architecture embedded in technological systems. They will also inform designs of instructional strategies and learning technologies that

prepare novice “kayakers” for their future. My dissertation, which examined learning with visual representations and learning to self-regulate choices involving the use of visual representations from various angles, offers one of the very few rich investigations on how people make choices, and how to support strategic choice making and meaningful learning processes.

## References

1. Ahn, J., Campos, F., Hays, M., & DiGiacomo, D. (2019). Designing in context: Reaching beyond usability in learning analytics dashboard design. *Journal of Learning Analytics*, 6(2), 70-85.
2. Ainsworth, S., & Iacovides, I. (2005). Learning by constructing self-explanation diagrams. In: *11th Biennial Conference of European Association for Research on Learning and Instruction (EARLI 2005)*, Nicosia, Cyprus.
3. Ainsworth, S., & Loizou, A. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science*, 27(4), 669-681.
4. Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to learn in science. *Science*, 333(6046), 1096-1097
5. Ainsworth, S. E., & Scheiter, K. (2021). Learning by drawing visual representations: Potential, purposes, and practical implications. *Current Directions in Psychological Science*, 30(1), 61-67.
6. Airey, J., & Erikson, U. (2019). Unpacking the Hertzprung-Russell diagram: A social semiotic analysis of the disciplinary and pedagogical affordances of a central resource in astronomy. *Designs for Learning*, 11(1), 99-107.
7. Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179-211.
8. Alevan, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-tracing tutors. *International Journal of Artificial Intelligence in Education*, 19(2), 105-154.
9. Alevan, V., & Koedinger, K. R. (2001). Investigations into help seeking and learning with a cognitive tutor. In *Papers of the AIED Workshop on Help Provision and Help Seeking in Interactive Learning Environments* (pp. 47-58).
10. Alevan, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, 26(2), 147-179.
11. Alevan, V., McLaren, B., Roll, I., & Koedinger, K. (2006). Toward meta-cognitive tutoring: A model of help seeking with a Cognitive Tutor. *International Journal of Artificial Intelligence in Education*, 16(2), 101-128.
12. Alevan, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive, action-by-action assessment of self-regulation during learning with an intelligent tutoring system. *Educational Psychologist*, 45(4), 224-233.
13. Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, 52(3), 215-241.

14. Arroyo, I., Woolf, B. P., Burelson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education*, 24(4), 387-426.
15. 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.
16. Ayabe, H., Manalo, E., Fukuda, M. & Sadato, N. (2021). What diagrams are considered useful for solving mathematical word problems in Japan? In A. Basu, G. Stapleton, S. Linker, C. Legg, E. Manalo, & P. Viana. (Eds.) *International Conference on theory and Application of Diagrams (Diagrams 2021)*, 79-83.
17. Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2010). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational psychologist*, 45(4), 210-223.
18. Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., & Fike, A. (2009). MetaTutor: A metacognitive tool for enhancing self-regulated learning. In R. Pirrone, R. Azevedo, & G. Biswas (Eds.) *Proceedings of the AAAI Fall Symposium on Cognitive and Metacognitive Educational Systems (pp. 14–19)*. Menlo Park, CA. Association for the Advancement of Artificial Intelligence (AAAI) Press.
19. Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why students engage in “gaming the system” behavior in interactive learning environments. *Journal of Interactive Learning Research*, 19(2), 185-224.
20. Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions in Psychological Science*, 9(3), 75-78.
21. Bano, M., Zowghi D., & Kearney, M. (2017). Feature based sentiment analysis for evaluating the mobile pedagogical affordances of apps. In A. Tatnall & M. Webb (Eds.), *Tomorrow's Learning: Involving Everyone. Learning with and about Technologies and Computing*. WCCE2017. IFIP Advances in Information and Communication Technology.
22. Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, 13(1), 1-14.
23. Barbieri, C. A., Miller-Cotto, D., & Booth, J. L. (2019). Lessening the load of misconceptions: Design-based principles for algebra learning. *Journal of the Learning Sciences*, 28(3), 381-417.
24. Bell, A., & Gresalfi, M. (2017). Teaching with videogames: How experience impacts classroom integration. *Technology, Knowledge and Learning*, 22(3), 513-526.

25. Belland, B. R., Walker, A. R., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, 87(2), 309-344.
26. Ben-Eliyahu, A., & Bernacki, M. L. (2015). Addressing complexities in self-regulated learning: A focus on contextual factors, contingencies, and dynamic relations. *Metacognition and Learning*, 10(1), 1-13.
27. Bernacki, M. L. (2017). Examining the cyclical, loosely sequenced, and contingent features of self-regulated learning: Trace data and their analysis. In *Handbook of Self-regulation of Learning and Performance* (pp. 370-387). Routledge.
28. Bernacki, M. L., Nokes-Malach, T. J., & Aleven, V. (2015). Examining self-efficacy during learning: Variability and relations to behavior, performance, and learning. *Metacognition Learning*, 10, 99-117.
29. Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, 13(2), 221-252.
30. Binbasaran Tüysüzoglu, B., & Greene, J. A. (2015). An investigation of the role of contingent metacognitive behavior in self-regulated learning. *Metacognition and Learning*, 10(1), 77-98.
31. Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review*, 30, 703-725.
32. Bjork, E. L. & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, 2, 59-68.
33. Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, 10(4), 405-418.
34. Booth, J. L., Barbieri, C., Eyer, F., & Paré-Blagoev, E. J. (2014). Persistent and pernicious errors in algebraic problem solving. *The Journal of Problem Solving*, 7(1), 3.
35. Booth, J., & Koedinger, K. R. (2010). Facilitating low-achieving students' diagram use in algebraic story problems. In *Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci 2010)* 32, 32.
36. Booth, J. L., & Koedinger, K. R. (2012). Are diagrams always helpful tools? Developmental and individual differences in the effect of presentation format on student problem solving. *British Journal of Educational Psychology*, 82(3), 492-511.

37. 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
38. Bower, M. (2008). Affordance analysis - matching learning tasks with learning technologies. *Educational Media International, 45*(1), 3-15.
39. Bransford, J. D., & Schwartz, D. L. (1999). Chapter 3: Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24*(1), 61-100.
40. Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1-13.
41. Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141-178.
42. Chan, J. Y. C., Lee, J. E., Mason, C. A., Sawrey, K., & Ottmar, E. (2022). From Here to There! A dynamic algebraic notation system improves understanding of equivalence in middle-school students. *Journal of Educational Psychology, 114*(1), 56-71.
43. Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145-182.
44. Chin, D. B., Blair, K. P., Wolf, R. C., Conlin, L. D., Cutumisu, M., Pfaffman, J., & Schwartz, D. L. (2019). Educating and measuring choice: A test of the transfer of design thinking in problem solving and learning. *Journal of the Learning Sciences, 28*(3), 337-380.
45. Cho, M. H., & Shen, D. (2013). Self-regulation in online learning. *Distance Education, 34*(3), 290-301.
46. Chu, J., Rittle-Johnson, B., & Fyfe, E. R. (2017). Diagrams benefit symbolic problem-solving. *British Journal of Educational Psychology, 87*(2), 273-287.
47. Clarebout, G., & Elen, J. (2006). Tool use in computer-based learning environments: Towards a research framework. *Computers in Human Behavior, 22*(3), 389-411.
48. Clark, R. E., & Estes, F. (1996). Cognitive task analysis for training. *International Journal of Educational Research, 25*(5), 403-417.
49. Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*(5), 537-550.
50. Clements, D. H., & McMillen, S. (1996). Rethinking "concrete" manipulatives. *Teaching Children Mathematics, 2*(5), 270-279.

51. Coburn, C. E., & Penuel, W. R. (2016). Research-practice partnerships in education: Outcomes, dynamics, and open questions. *Educational Researcher*, 45(1), 48-54.
52. Common Core State Standards Initiative. (2010). *Common Core State Standards for Mathematics*. [http://www.corestandards.org/wp-content/uploads/Math\\_Standards.pdf](http://www.corestandards.org/wp-content/uploads/Math_Standards.pdf)
53. Consolvo, S., Landay, J. A., & McDonald, D. W. (2009). Designing for behavior change in everyday life. *Focus*, 405, 414.
54. Cooper, J. L., Sidney, P. G., & Alibali, M. W. (2018). Who benefits from diagrams and illustrations in math problems? Ability and attitudes matter. *Applied Cognitive Psychology*, 32(1), 24-38.
55. Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, 34(4), 344-377.
56. Croxton, R. A., & Chow, A. S. (2015). Using ADDIE and systems thinking as the framework for developing a MOOC: A case study. *Quarterly Review of Distance Education*, 16(4), 83.
57. Cutumisu, M., Blair, K. P., Chin, D. B., & Schwartz, D. L. (2015). Posterlet: A game-based assessment of children's choices to seek feedback and to revise. *Journal of Learning Analytics*, 2(1), 49-71.
58. Cutumisu, M., Chin, D. B., & Schwartz, D. L. (2019). A digital game-based assessment of middle-school and college students' choices to seek critical feedback and to revise. *British Journal of Educational Technology*, 50(6), 2977-3003.
59. Davenport, J., Yaron, D., Klahr, D., & Koedinger, K. (2008). When do diagrams enhance learning? A framework for designing relevant representations. In G. Kanselaar, V. Jonker, P. A. Kirschner, & F. J. Prins (Eds.) *International Perspectives in the Learning Sciences: Creating a learning world, Proceedings of the International Conference of the Learning Sciences (ICLS2008)* (pp. 191-198). Utrecht The Netherlands: International Society of the Learning Sciences.
60. de Koning, B. B., Tabbers, H. K., Rikers, R. M., & Paas, F. (2010). Learning by generating vs. receiving instructional explanations: Two approaches to enhance attention cueing in animations. *Computers & Education*, 55(2), 681-691.
61. Dickey, M. D. (2003). Teaching in 3D: Pedagogical affordances and constraints of 3D virtual worlds for synchronous distance learning. *Distance Education*, 24(1), 105-121.
62. diSessa, A. A. (2004). Metarepresentation: Native competence and targets for instruction. *Cognition and Instruction*, 22(3), 293-331.

63. Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior, 52*, 338-348.
64. Faraj, S., & Azad, B. (2012). The materiality of technology: An affordance perspective. In P. M. Lornardi, B. A. Nardi, & J. Kallinikos (Eds.), *Materiality and organizing: Social interaction in a technological world* (pp. 237-258).
65. Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review, 30*(3), 1115-1137.
66. Foster, A., Mishra, P., & Koehler, M. (2011). Digital game analysis: Using the Technological Pedagogical Content Knowledge framework to determine the affordances of a game for learning. In M. S. Khine (Ed.), *Learning to play: Exploring the future of education with video games*. New York: Peter Lang Publications.
67. 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)
68. Fukuda, M., Manalo, E., & Ayabe, H. (2021). The presence of diagrams and problems requiring diagram construction: Comparing mathematical word problems in Japanese and Canadian textbooks. In Basu, A., Stapleton, G., Linker, S., Legg, C., Manalo, E., Viana, P. (Eds.) *International Conference on theory and Application of Diagrams (Diagrams 2021)*. 353-357.
69. Fyfe, E. R., Matthews, P. G., Amsel, E., McEldoon, K. L., & McNeil, N. M. (2018). Assessing formal knowledge of math equivalence among algebra and pre-algebra students. *Journal of Educational Psychology, 110*(1), 87.
70. Fyfe, E. R., McNeil, N. M., & Borjas, S. (2015). Benefits of “concreteness fading” for children's mathematics understanding. *Learning and Instruction, 35*, 104-120.
71. Gibson, J. J. (1977). The theory of affordances. In R. E. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67-82). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
72. Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine, 22*(4), 39-51.
73. Grawemeyer, B., & Cox, R. (2004). The effect of knowledge-of-external-representations upon performance and representational choice in a database query task. In A. F. Blackwell, K. Marriott, Shimojima, A. (Eds.) *In Proceedings of the International Conference on Theory and Application of Diagrams* (pp. 351-354). Springer, Berlin, Heidelberg.

74. Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist, 45*(4), 203-209.
75. Greene, J. A., Plumley, R. D., Urban, C. J., Bernacki, M. L., Gates, K. M., Hogan, K. A., Demetriou, C., & Panter, A. T. (2021). Modeling temporal self-regulatory processing in a higher education biology course. *Learning and Instruction, 72*, 101201.
76. Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 335-342.
77. Greeno, J. G., & Gresalfi, M. S. (2008). Opportunities to learn in practice and identity. In P. A. Moss, D. C. Pullin, J. P. Gee, E. H. Haertel, & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn* (pp. 170–199). Cambridge University Press.
78. Gresalfi, M. S. (2013). Technology in mathematics education: A discussion of affordances. In *Proceedings of the 35th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*.
79. Gresalfi, M. S., Barnes, J., & Cross, D. (2012). When does an opportunity become an opportunity? Unpacking classroom practice through the lens of ecological psychology. *Educational Studies in Mathematics, 80*(1-2), 249-267.
80. Güney, Z. (2019). Four-component instructional design (4C/ID) model approach for teaching programming skills. *International Journal of Progressive Education, 15*(4), 142-156.
81. Hadwin, A. F., & Winne, P. H. (2001). CoNoteS2: A software tool for promoting self-regulation. *Educational Research and Evaluation, 7*(2-3), 313-334.
82. Hansen, A., Mavrikis, M., & Geraniou, E. (2016). Supporting teachers' technological pedagogical content knowledge of fractions through co-designing a virtual manipulative. *Journal of Mathematics Teacher Education, 19*(2), 205-226.
83. Hedges, L. V., Tipton, E., & Hedberg, E. (forthcoming). *Title pending*.
84. Heimlich, J. E., & Ardoin, N. M. (2008). Understanding behavior to understand behavior change: A literature review. *Environmental Education Research, 14*(3), 215-237.
85. Higgins, S., Katsipataki, M., Villanueva Aguilera, A. B., Dobson, E., Gascoine, L., Rajab, T., ... & Uwimpuhwe, G. (2022). The teaching and learning toolkit: Communicating research evidence to inform decision-making for policy and practice in education. *Review of Education, 10*(1).
86. Hilton, J. (2016). Open educational resources and college textbook choices: A review of research on efficacy and perceptions. *Educational Technology Research and Development, 64*(4), 573-590.

87. Holstein, K., McLaren, B. M., & Alevan, V. (2019). Co-designing a real-time classroom orchestration tool to support teacher-AI complementarity. *Journal of Learning Analytics*, 6(2), 27-52.
88. Huang, Y., Lobczowski, N. G., Richey, E., McLaughlin, E. A., Asher, M. W., Harackiewicz, J. M., Alevan, V., & Koedinger, K. R. (2021). A general multi-method approach to data-driven redesign of tutoring systems. In *Proceedings of the 11<sup>th</sup> International Learning Analytics and Knowledge Conference (LAK21)*, 161-172.
89. Hubber, P., Tytler, R., & Haslam, F. (2010). Teaching and learning about force with a representational focus: Pedagogy and teacher change. *Research in Science Education*, 40(1), 5-28.
90. Koedinger, K. R. & Alevan, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review* 19, 3. 239-264.
91. Koedinger, K., Alevan, V., Roll, I., & Baker, R. (2009). In vivo experiments on whether supporting metacognition in intelligent tutoring systems yields robust learning. In *Handbook of Metacognition in Education* (pp. 395-424). Routledge.
92. Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science*, 32(2), 366-397.
93. Koedinger, K. R., & Corbett, A. (2006). Cognitive tutors: Technology bringing learning sciences to the classroom. In K. Sawyer (Ed.) *The Cambridge Handbook of the Learning Sciences*, 61-78. New York: Cambridge University Press.
94. Koedinger, K., Cobett, A., & Perfett, C. (2012). The Knowledge-Learning-Instruction Framework: Bringing the science-practice chasm to enhance robust student learning. *Cognitive Science* 36(5), 757-798.
95. Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60-70.
96. Kozma, R. (2003). The material features of multiple representations and their cognitive and social affordances for science understanding. *Learning and Instruction*, 13(2), 205-226.
97. Kramarski, B., & Gutman, M. (2006). How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning*, 22, 24-33.
98. Krauskopf, K., Zahn, C., & Hesse, F. W. (2012). Leveraging the affordances of Youtube: The role of pedagogical knowledge and mental models of technology functions for lesson planning with technology. *Computers & Education*, 58(4), 1194-1206.

99. Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M., & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology, 100*(3), 716.
100. Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: a meta-analytic review. *Review of Educational Research, 86*(1), 42-78.
101. Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*(1), 65-100.
102. Laursen, S., Liston, C., Thiry, H., Graf, J. (2007). What good is a scientist in the classroom? Participant outcomes and program design features for a short-duration science outreach intervention in K–12 classrooms. *CBE—Life Sciences Education, 6*(1), 49-64.
103. Lee, K., Khng, K. H., Ng, S. F., & Ng Lan Kong, J. (2013). Longer bars for bigger numbers? Children's usage and understanding of graphical representations of algebraic problems. *Frontline Learning Research, 1*(1), 81-96.
104. Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: The Betty's Brain system. *International Journal of Artificial Intelligence in Education, 18*(3), 181-208.
105. Lim, C. P. (2007). Effective integration of ICT in Singapore schools: Pedagogical and policy implications. *Educational Technology Research and Development, 55*(1), 83-116.
106. Long, Y. & Aleven, V. (2014). Gamification of joint student/system control over problem selection in a linear equation tutor. In *Proceedings of the International Conference on Intelligent Tutoring System (ITS 2014)*, 378-387.
107. Long, Y., & Aleven, V. (2017). Enhancing learning outcomes through self-regulated learning support with an open learner model. *User Modeling and User-Adapted Interaction, 27*(1), 55-88.
108. Long, Y., Aman, Z., & Aleven, V. (2015). Motivational design in an intelligent tutoring system that helps students make good task selection decisions. In *International Conference on Artificial Intelligence in Education* (pp. 226-236). Springer, Cham.
109. Long, Y., Holstein, K., & Aleven, V. (2018). What exactly do students learn when they practice equation solving? Refining knowledge components with the additive factors model. In *Proceedings of the 8<sup>th</sup> Learning Analytics and Knowledge conference (LAK2018)*. 399-408.
110. Lortie-Forgues, H., Sio, U. N., & Inglis, M. (2021). How should educational effects be communicated to teachers?. *Educational Researcher, 50*(6), 345-354.
111. Lucero, A. (2015). Using affinity diagrams to evaluate interactive prototypes. In *IFIP conference on Human-Computer Interaction* (pp. 231-248). Springer, Cham
112. Lurie, N. H., & Mason, C. H. (2007). Visual representation: Implications for decision making. *Journal of Marketing, 71*(1), 160-177.

113. Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., & Clarebout, G. (2011). Tool-use in a blended undergraduate course: In search of user profiles. *Computers & Education*, *57*, 2135–2044.
114. Magner, U. I., Schwonke, R., Alevan, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, *29*, 141-152.
115. Manalo, E., & Uesaka, Y. (2012). Elucidating the mechanism of spontaneous diagram use in explanations: how cognitive processing of text and diagrammatic representations are influenced by individual and task-related factors. In P. Cox, B. Plimmer, P. Rodgers (Eds.) In *Proceedings of the International Conference on Theory and Application of Diagrams* (pp. 35-50). Springer, Berlin, Heidelberg.
116. Martin, N. D., Gnesdilow, D., & Puntambekar, S. (2018). Teachers' mediation of students' interactions with physical and virtual scientific models in biology. In J. Kay, & R. Luckin (Eds.) *Rethinking Learning in the Digital Age: Making the Learning Science Count, 13<sup>th</sup> International Conference of the Learning Sciences (ICLS2018)*. London, UK: International Society of the Learning Sciences.
117. Masoudi, N., Fadel, G. M., Pagano, C. C., & Elena, M. V. (2019). A Review of Affordances and Affordance-Based Design to Address Usability. In *Proceedings of the Design Society: International Conference on Engineering Design* (Vol. 1, No. 1, pp. 1353-1362). Cambridge University Press.
118. Matcha, W., Gašević, D., & Pardo, A. (2019). A systematic review of empirical studies on learning analytics dashboards: A self-regulated learning perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226-245.
119. Mavrikis, M., Holmes, W., Zhang, J., & Ma, N. (2018). Fractions lab goes east: Learning and interaction with an exploratory learning environment in China. In *International Conference on Artificial Intelligence in Education* (pp. 209-214). Springer, Cham.
120. Mawasi, A., Nagy, P., Finn, E. et al. (2021). Narrative-based learning activities for science ethics education: an Affordance perspective. *Journal of Science Education and Technology*, *31*(1), 16-26.
121. McGrenere, J., & Ho, W. (2000). Affordances: Clarifying and evolving a concept. In *Proceedings of the Graphics Interface Conference* (Vol. 2000, pp. 179-186).
122. McNeil, N. M., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Brletic-Shiple, H. (2011). Benefits of practicing  $4 = 2 + 2$ : Nontraditional problem formats facilitate children's understanding of mathematical equivalence. *Child Development*, *82*, 1620-1633.

123. Melis, E., & Siekmann, J. (2004). ActiveMath: An intelligent tutoring system for mathematics. In *International Conference on Artificial Intelligence and Soft Computing* (pp. 91-101). Springer, Berlin, Heidelberg.
124. Middleton, K. V. (2020). The longer-term Impact of COVID-19 on K–12 student learning and assessment. *Educational Measurement: Issues and Practice*, 39(3), 41-44.
125. Morin, L. L., Watson, S. M., Hester, P., & Raver, S. (2017). The use of a bar model drawing to teach word problem solving to students with mathematics difficulties. *Learning Disability Quarterly*, 40(2), 91-104.
126. Mullis, I. V., Martin, M. O., Foy, P., Kelly, D. L., Fishbein, B. (2019). TIMSS 2019 International Results in Math and Science: Highlights. Retrieved from Boston College, *TIMSS & PIRLS International Study Center website*: <https://timss2019.org/reports/>
127. Murata, A. (2008). Mathematics teaching and learning as a mediating process: The case of tape diagrams. *Mathematical Thinking and Learning*, 10(4), 374-406.
128. Nagashima, T., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevin, V. (2020). Enhancing conceptual knowledge in early algebra through scaffolding diagrammatic self-explanation. In M. Gresalfi & I. S. Horn (Eds.), *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences* (pp. 35-42). Nashville, TN: International Society of the Learning Sciences.
129. Nagashima, T., Bartel, A. N., Tseng, S., Vest, N.A., Silla, E. M., Alibali, M. W., & Alevin, V. (2021). Scaffolded self-explanation with visual representations promotes efficient learning in early algebra. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *43rd Annual Meeting of the Cognitive Science Society* (pp. 1858-1864). Cognitive Science Society.
130. Nagashima, T., Bartel, A. N., Yadav, G., Tseng, S., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevin, V. (2021). Using anticipatory diagrammatic self-explanation to support learning and performance in early algebra. In E. de Vries, J. Ahn, & Y. Hod (Eds.), *15th International Conference of the Learning Sciences* (pp. 474–481). Bochum, Germany: International Society of the Learning Sciences.
131. Nagashima, T., Britti, J., Wang, X., Zheng, B., Turri, V., Tseng, S., & Alevin, V. (2022). Designing playful intelligent tutoring software to support engaging and effective algebra learning. In *Proceedings of the 17th European Conference on Technology Enhanced Learning (EC-TEL2022)*.
132. Nagashima, T., & Hrach, S. (2021). Motivating Factors among University Faculty for Adopting Open Educational Resources: Incentives Matter. *Journal of Interactive Media in Education*, 2021(1).

133. Nagashima, T., Ling, E., Zheng, B., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Aleven, V. (2022). How does sustaining and interleaving visual scaffolding help learners? A classroom study with an Intelligent Tutoring System. In *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*. Cognitive Science Society.
134. Nagashima, T., Tseng, S., Ling, E., Bartel, A. N., Vest, N. A., Silla, E. M., Alibali, M. W., & Aleven, V. (2022). Students' self-regulated use of diagrams in a choice-based Intelligent Tutoring System. In *Proceedings of the Annual Meeting for the International Society of the Learning Sciences (ISLS2022)*, Hiroshima, Japan: International Society of the Learning Sciences.
135. Nagashima, T., Yadav, G., & Aleven, V. (2021). A framework to guide technology-based educational studies in the evolving classroom environment. In T. De Laet T, R. Klemke, C. Alario-Hoyos, I. Hilliger I, & A. Ortega-Arranz. (Eds.), *Proceedings of the 16th European Conference on Technology Enhanced Learning (EC-TEL2021)* (pp. 207-220).
136. Nagashima, T., Yang, K., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Aleven, V. (2020). Pedagogical Affordance Analysis: Leveraging teachers' pedagogical knowledge for eliciting pedagogical affordances and constraints of instructional tools. In M. Gresalfi & I. S. Horn (Eds.), *The Interdisciplinarity of the Learning Sciences, 14th International Conference of the Learning Sciences* (pp. 1561-1564). Nashville, TN: International Society of the Learning Sciences.
137. Nguyen, H., Wang, Y, Stamper, J., & McLaren, B. M. (2019). Using knowledge component modeling to increase domain understanding in a digital learning game. In *Proceedings of the International Conference on Educational Data Mining*.
138. Opezzo, M., & Schwartz, D. L. (2013). A behavior change perspective on self-regulated learning with teachable agents. In R. Azevedo, & V. Aleven. (Eds), *International Handbook of Metacognition and Learning Technologies* (pp. 485-500). Springer, New York, NY.
139. Ottmar, E., & Landy, D. (2017). Concreteness fading of algebraic instruction: Effects on learning. *Journal of the Learning Sciences*, 26(1), 51-78.
140. Panadero, E., & Alonso Tapia, J. (2014). How do students self-regulate?: Review of Zimmerman's cyclical model of self-regulated learning. *Anales de psicología*, 30(2), 450-462.
141. Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261-284.
142. Patrick, S. W., Henkhaus, L. E., Zickafoose, J. S., Lovell, K., Halvorson, A., Loch, S., Letterie, M., Davis, M. M. (2020). Well-being of parents and children during the COVID-19 pandemic: A national survey. *Pediatrics*, 146(4).

143. Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and Educational Considerations*, 11, 47-87.
144. Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*, 13(3), 423-451.
145. Pendleton-Jullian, A. M., & Brown, J. S. (2018). *Design unbound: Designing for emergence in a white water world*. Cambridge, MA: MIT Press.
146. Pintrich, P. R. (1995). Understanding self-regulated learning. *New Direction for Teaching and Learning*, 63, 3-12.
147. Pintrich, P. R. (1999) The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459-470.
148. Prather, R. W., Benitez, V. L., Brooks, L. K., Dancy, C. L., Dilworth-Bart, J., Dutra, N. B., Faison, M., Figueroa, M., Holden, L. R., Johnson, C., Medrano, J., Miller-Cotto, D., Matthews, P. G., Manly, J. J., & Thomas, A. K. (2022). What can Cognitive Science do for people? *Cognitive Science*, 46(6),
149. Prochaska, J. O., DiClemente, C. C., & Norcross, J. C. (1992). In search of how people change: Applications to the addictive behaviors. *American Psychologist*, 47, 1102-1114.
150. Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher*, 29(1), 4-15.
151. Rau, M. A. (2017). Conditions for the effectiveness of multiple visual representations in enhancing STEM learning. *Educational Psychology Review*, 29(4), 717-761.
152. Rau, M. A., Aleven, V., & Rummel, N. (2010). Blocked versus interleaved practice with multiple representations in an Intelligent Tutoring System for fractions. In *Proceedings of the International Conference on Intelligent Tutoring Systems (ITS2010)*, 413-422.
153. Rau, M. A., Aleven, V., & Rummel, N. (2013). Interleaved practice in multi-dimensional learning tasks. Which dimension should we interleave? *Learning and Instruction* 23. 98-114.
154. Rau, M. A., Aleven, V., & Rummel, N. (2015). Successful learning with multiple graphical representations and self-explanation prompts. *Journal of Educational Psychology*, 107(1), 30.
155. Rauth, I., Köppen, E., Jobst, B., & Meinel, C. (2010). Design thinking: An educational model towards creative confidence. In *Proceedings of the 1st International Conference on Design Creativity (ICDC 2010)*.
156. Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading worked solution steps works—a cognitive load perspective. *Instructional Science*, 32(1), 59-82.

157. Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science*, 21(1), 1-29.
158. Renkl, A. (1999). Learning mathematics from worked-out examples: Analyzing and fostering self-explanations. *European Journal of Psychology of Education*, 14(4), 477-488.
159. Rittle-Johnson, B. (2019). Iterative development of conceptual and procedural knowledge in mathematics learning and instruction. In J. Dunlosky & K. A. Rawson (Eds.), *The Cambridge handbook of cognition and education* (pp. 124–147). Cambridge University.
160. Rittle-Johnson, B., Fyfe, E. R., & Loehr, A. M. (2016). Improving conceptual and procedural knowledge: The impact of instructional content within a mathematics lesson. *British Journal of Educational Psychology*, 86(4), 576-591.
161. Rittle-Johnson, B., Loehr, A. M., & Durkin, K. (2017). Promoting self-explanation to improve mathematics learning: A meta-analysis and instructional design principles. *ZDM*, 49(4), 599-611.
162. Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, 103(1), 85.
163. Rittle-Johnson, B., Fyfe, E. R., & Loehr, A. M. (2016). Improving conceptual and procedural knowledge: The impact of instructional content within a mathematics lesson. *British Journal of Educational Psychology*, 86(4), 576-591.
164. 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.
165. Rittle-Johnson, B., Star, J. R., & Durkin, K. (2020). How can cognitive-science research help improve education? The case of comparing multiple strategies to improve mathematics learning and teaching. *Current Directions in Psychological Science*, 29(6), 599-609.
166. Rivers K., Harpstead, R. & Koedinger, K. R. (2016). Learning curve analysis for programming: Which concepts do students struggle with? In *Proceedings of the International Computing Education Conference*, 143-151.
167. Rohrer, D., Dedrick, R. F., & Stershic, S. (2015). Interleaved practice improves mathematics learning. *Journal of Educational Psychology* 107(3), 900-908.
168. Roscoe, R. D., Segedy, J. R., Sulcer, B., Jeong, H., & Biswas, G. (2013). Shallow strategy development in a teachable agent environment designed to support self-regulated learning. *Computers & Education*, 62, 286-297.

169. Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction, 21*(2), 267-280.
170. Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics, 2*(1), 7-12
171. Rothman, A. J., Baldwin, A. S., Hertel, A. W., & Fuglestad, P. T. (2011). Self-regulation and behavior change: Disentangling behavioral initiation and behavioral maintenance. In K. D. Vohs & R. F. Baumeister (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 106–122). Guilford Press.
172. Rummel, N., Mavrikis, M., Wiedmann, M., Loibl, K., Mazziotti, C., Holmes, W., & Hansen, A. (2016). Combining exploratory learning with structured practice to foster conceptual and procedural fractions knowledge. In C. K. Looi, J. L. Polman, U. Cress, & P. Reimann. (Eds.), *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS)*, Singapore: International Society of the Learning Sciences.
173. Sabourin, J.L., Shores, L.R., Mott, B.W., & Lester, J. (2013). Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education, 23*, 94-114.
174. Scheiter, K., Schleinschok, K., & Ainsworth, S. (2017). Why sketching may aid learning from science texts: Contrasting sketching with written explanations. *Topics in Cognitive Science, 9*(4), 866-882.
175. Schraw, G. (2007). The use of computer-based environments for understanding and improving self-regulation. *Metacognition Learning 2*, 169–176.
176. Schunk, D. H. (2008). Attributions as motivators of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-Regulated learning: Theory, research, and applications* (pp. 245–266).
177. Schwartz, D. L., & Arena, D. (2013). *Measuring what matters most: Choice-based assessments for the digital age*. The MIT Press.
178. Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759.
179. Schwonke, R., Ertelt, A., Otieno, C., Renkl, A., Aleven, V., & Salden, R. J. (2013). Metacognitive support promotes an effective use of instructional resources in intelligent tutoring. *Learning and Instruction, 23*, 136-150.

180. Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
181. Sidney, P. G., Hattikudur, S., & Alibali, M. W. (2015). How do contrasting cases and self-explanation promote learning? Evidence from fraction division. *Learning and Instruction*, 40, 29-38.
182. Spielhagen, F. R. (2006). Closing the achievement gap in math: The long-term effects of eighth-grade algebra. *Journal of Advanced Academics*, 18(1), 34-59.
183. Stamper, J., Koedinger, K., & McLaughlin, E. (2013). A comparison of model selection metrics in DataShop. In *Proceedings of the 6<sup>th</sup> International Conference on Educational Data Mining*. International Educational Data Mining Society.
184. Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970.
185. Stefaniak, J., & Xu, M. (2020). An examination of the systemic reach of instructional design models: A systematic review. *TechTrends*, 64(5), 710-719.
186. Stelitano, L., Doan, S., Woo, A., Diliberti, M., Kaufman, J. H., & Henry, D. (2020). The digital divide and COVID-19: Teachers' perceptions of inequities in students' internet access and participation in remote learning. Data Note: Insights from the American Educator Panels. Research Report. *Rand Corporation*.
187. Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Sage Publications.
188. Suyatna, A., Anggraini, D., Agustina, D., & Widyastuti, D. (2017). The role of visual representation in physics learning: Dynamic versus static visualization. *Journal of Physics: Conference Series*, 909(1), 12-48.
189. Tan, J., & Biswas, G. (2006). The role of feedback in preparation for future learning: A case study in learning by teaching environments. In *International Conference on Intelligent Tutoring Systems* (pp. 370-381). Springer, Berlin, Heidelberg.
190. Tan, T. H., Lin, M. S., Chu, Y. L., & Liu, T. Y. (2012). Educational affordances of a ubiquitous learning environment in a natural science course. *Journal of Educational Technology and Society*, 15(2), 206-219.
191. Tippett, C. D. (2016). What recent research on diagrams suggests about learning with rather than learning from visual representations in science. *International Journal of Science Education*, 38(5), 725-746.

192. Trust, T., & Pektas, E. (2018). Using the ADDIE model and universal design for learning principles to develop an open online course for teacher professional development. *Journal of Digital Learning in Teacher Education*, 34(4), 219-233.
193. Uesaka, Y., & Manalo, E. (2006). Active comparison as a means of promoting the development of abstract conditional knowledge and appropriate choice of diagrams in math word problem solving. In D. Barker-Plummer, R. Cox, N. Swoboda. (Eds.), In *Proceedings of the International Conference on Theory and Application of Diagrams* (pp. 181-195). Springer, Berlin, Heidelberg.
194. Uesaka, Y., & Manalo, E. (2007). Peer instruction as a way of promoting spontaneous use of diagrams when solving math word problems. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 29, No. 29).
195. Uesaka, Y., Manalo, E., & Ichikawa, S. I. (2007). What kinds of perceptions and daily learning behaviors promote students' use of diagrams in mathematics problem solving? *Learning and Instruction*, 17(3), 322-335.
196. Uesaka, Y., & Manalo, E. (2012). Task-related factors that influence the spontaneous use of diagrams in math word problems. *Applied Cognitive Psychology*, 26(2), 251-260.
197. Uesaka, Y., Manalo, E., & Ichikawa, S. I. (2010). The effects of perception of efficacy and diagram construction skills on students' spontaneous use of diagrams when solving math word problems. In A. K. Goel, M. Jamink, N. H. Narayan. (Eds.), In *Proceedings of the International Conference on Theory and Application of Diagrams* (pp. 197-211). Springer, Berlin, Heidelberg.
198. Van Garderen, D., & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research & Practice*, 18(4), 246-254.
199. Van Harsel, M., Hoogerheide, V., Janssen, E., Verkoeijen, P., & van Gog, T. (2022). How do higher education students regulate their learning with video modeling examples, worked examples, and practice problems? *Instructional Science*, 1-26
200. Van Harsel, M., Hoogerheide, V., Verkoeijen, P., & van Gog, T. (2022). Instructing students on effective sequences of examples and problems: Does self-regulated learning improve from knowing what works and why? *Journal of Computer Assisted Learning*, 38(1), 19-39.
201. Weinstein, C. E., Acee, T. W., & Jung, J. (2011). Self-regulation and learning strategies. *New Directions for Teaching and Learning*, 126, 45-53.
202. West, R. (2005). Time for a change: Putting the Transtheoretical (Stages of Change) Model to rest. *Addiction*, 100, 1036-1039.
203. The William and Flora Hewlett Foundation (n.d.) Open Education. Retrieved from: <https://hewlett.org/strategy/open-education/>.

204. Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277-304). Mahwah, NJ: Erlbaum.
205. Wisittanawat, P., & Gresalfi, M. S. (2021). The “tricky business” of genre blending: Tensions between frames of school mathematics and video game play. *Journal of the Learning Sciences, 30*(2), 240-278.
206. Wood, H., & Wood, D. (1999). Help seeking, learning and contingent tutoring. *Computers & Education, 33*(2-3), 153-169.
207. Wu, H., & Puntambekar, S. (2012). Pedagogical affordances of multiple external representations in scientific processes. *Journal of Science Education and Technology, 21*(6), 754-767.
208. Wu, S. P., & Rau, M. A. (2018). Effectiveness and efficiency of adding drawing prompts to an interactive educational technology when learning with visual representations. *Learning and Instruction, 55*, 93-104.
209. Wu, S. P., Van Veen, B., & Rau, M. A. (2020). How drawing prompts can increase cognitive engagement in an active learning engineering course. *Journal of Engineering Education, 109*(4), 723-742.
210. Yung, H. I., & Paas, F. (2015). Effects of cueing by a pedagogical agent in an instructional animation: A cognitive load approach. *Journal of Educational Technology & Society, 18*(3), 153-160.
211. Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. J. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology, 107*(4), 954.
212. Zhang, Q. & Fiorella, L. (2021). Learning by drawing: When is it worth the time and effort? *Contemporary Educational Psychology, 66*.
213. Zhou, M., & Winne, P. H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction, 22*(6), 413-419.
214. Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational psychologist, 48*(3), 135-147.
215. Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166-183.
216. Zimmerman, B. J. (1990) Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3-17.

217. Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. Sternberg (Eds.), *The nature of problem solving*. New York: Cambridge University Press.
218. Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614-628.

**Appendix 1. Tests and surveys used in experiments**

*Sample pretest and posttest items used in Study 2.2. 1-6, 10-11 are conceptual items whereas 7-9 are procedural items.*

1. One of the following statements is FALSE. Which one? Circle your answer.

- a.  $3 + 5 = 8$
- b.  $7 + 6 = 6 + 6 + 1$
- c.  $5 + 5 = 5 + 6$
- d. All of the above
- e. None of the above

2 – 1. Look at the following step for the equation  $2x + 4 = 8$ :

$$2x + 4 - 4 = 8 - 8$$

Which of the following is the appropriate description of this step? Circle one answer.

- a. Correct and helpful (it preserves the solution and it gets you closer to the solution)
- b. Correct but not helpful (it preserves the solution but does not get you close to the correct solution)
- c. Incorrect (it does not preserve the solution)
- d. Helpful though not correct
- e. None of the above

2 – 2. Look at the following step for the equation  $2x + 4 = 8$  (same equation as above):

$$2x + 4 - 4 = 8 - 4$$

Which of the following is the appropriate description of this step? Circle one answer.

- a. Correct and helpful (it preserves the solution and it gets you closer to the solution)
- b. Correct but not helpful (it preserves the solution but does not get you close to the correct solution)
- c. Incorrect (it does not preserve the solution)
- d. Helpful though not correct
- e. None of the above

2 – 3. Look at the following step for the equation  $2x + 4 = 8$  (same equation as above):

$$2x + 4 + 4 = 8 + 4$$

Which of the following is the appropriate description of this step? Circle one answer.

- a. Correct and helpful (it preserves the solution and it gets you closer to the solution)
- b. Correct but not helpful (it preserves the solution but does not get you close to the correct solution)
- c. Incorrect (it does not preserve the solution)
- d. Helpful though not correct
- e. None of the above

2 – 4. Look at the following step for the equation  $2x + 4 = 8$  (same equation as above):

$$2x - 6 + 4 + 6 = 8$$

Which of the following is the appropriate description of this step? Circle one answer.

- a. Correct and helpful (it preserves the solution and it gets you closer to the solution)
- b. Correct but not helpful (it preserves the solution but does not get you close to the correct solution)
- c. Incorrect (it does not preserve the solution)
- d. Helpful though not correct
- e. None of the above

3. The arrow below is pointing to a symbol.

$$2x + 6 = 10$$

↑

3 – 1. What is the name of the symbol?

3 – 2. What does the symbol mean?

4. Consider the equation,  $3x + 2 = 8$ . What value of  $x$  makes the equation true? Circle one answer.

- a. 1
- b. 2
- c. 3
- d. 4
- e. All of the above
- f. None of the above
- g. Any value of  $x$

5. Consider the equation,  $251x + 1769 = 78x + 5748$ . This equation is true for  $x = 23$ .

Larry (who was not told what value  $x$  would have) decides to solve for  $x$ . He subtracts 1769 from both sides. That gives:

$$251x = 78x + 3979$$

No need to check if the numbers are correct. We already did for you! ( $5748 - 1769 = 3979$ )

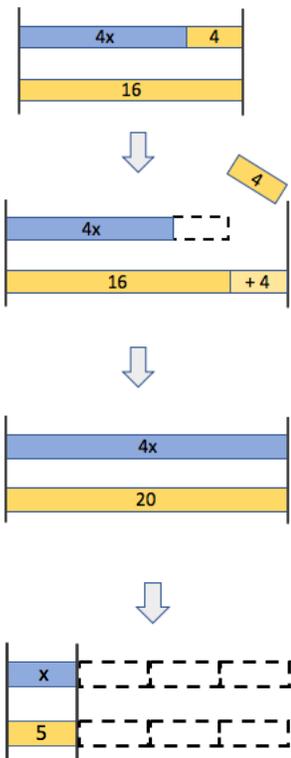
Without plugging 23 into the equation, can you tell if this new equation is true for  $x = 23$ ? Please explain why or why not.

6. Which of the following is a like term to  $8x$ ? Circle one answer.

- a. 1
- b.  $5y$
- c.  $2x$
- d.  $-20$
- e. All of the above
- f. None of the above



11-2. For the same solution steps above by Sonia, explain the error and why it is an error using the diagrams below, which illustrate her steps.



*Sample pretest and posttest items used in Studies 3 and 4. 1-8 are conceptual items whereas 9-12 are procedural items.*

1. Debby is solving an equation. Here is her first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8 - 8$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid but it gets closer to the solution
- NOT valid and it does NOT get you closer to the solution
- None of the above

2. Christopher is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8 - 4$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid but it gets closer to the solution
- NOT valid and it does NOT get you closer to the solution
- None of the above

3. Joseph is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$
$$2x + 4 + 4 = 8 + 4$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid but it gets closer to the solution
- NOT valid and it does NOT get you closer to the solution
- None of the above

4. Tiana is solving an equation. Here is her first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid but it gets closer to the solution
- NOT valid and it does NOT get you closer to the solution
- None of the above

156

5. Ali is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$

$$2 + 4 - 2 = 8 - 2$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid but it gets closer to the solution
- NOT valid and it does NOT get you closer to the solution
- None of the above

6. Jill is solving an equation. Here is her first solution step:

$$7n - 6 = 15$$

$$\mathbf{7n = 21}$$

What did she do?

- She added 6 to both side
- She subtracted 6 from both side
- She multiplied both sides by 6
- She divided both sides by 6
- None of the above

7. Marcel is solving an equation. Here's his first solution step:

$$3x + 2 = 10$$

$$5x = 10$$

Is Marcel's step correct? Why or why not?

- Correct, because he subtracted 2 from both sides
- Correct, because he added  $2x$  to both sides
- Correct, because he subtracted 2 from one side
- Incorrect, because he tried to add  $3x$  and 2
- Incorrect, because he did not do the same thing to both sides
- Incorrect, because he did not isolate the  $x$

8.a. Below is Chloe's solution steps to the equation,  $3x + 9 = 18$ . Click on the area where you see a mistake (click on only one area).

$$3x + 9 = 18$$

$$3x + 9 - 9 = 18 + 9$$

$$3x = 27$$

$$x = 9$$

8.b. What is the mistake and why is it a mistake? Please explain.

9. Solve for x. Please show your work:

$$3x + 4 = 16$$

10. Solve for x. Please show your work:

$$9x = 2x + 14$$

11. Solve for x. Please show your work:

$$2x - 3 = 9$$

12. Solve for x. Please show your work:

$$2x = -3x + 10$$

*Sample pretest and posttest items used in Study 5. 1-5 are conceptual items whereas 6-12 are procedural items.*

1. Debby is solving an equation. Here is her first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8 - 8$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

2. Christopher is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8 - 4$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

160

3. Joseph is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$

$$2x + 4 + 4 = 8 + 4$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

4. Tiana is solving an equation. Here is her first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

5.a. Below is Sonia's solution steps to the equation,  $4x + 4 = 16$ . Click on the area where you see a mistake (click on only one area).

$$\begin{aligned}4x + 4 &= 16 \\4x + 4 - 4 &= 16 + 4 \\4x &= 20 \\x &= 5\end{aligned}$$

5.b. What exactly is the mistake and why is it a mistake? Please explain.

6. Solve for  $x$ . Please show your work:

$$3x + 4 = 16$$

7. Solve for  $x$ . Please show your work:

$$9x = 2x + 14$$

8. Solve for  $x$ . Please show your work:

$$11x + 6 = 8x + 21$$

9. Solve for  $x$ . Please show your work:

$$2x - 3 = 9$$

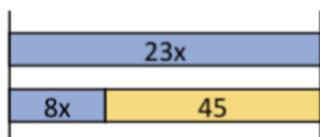
10. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$7 + 8x = 71$$



11. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$23x = 8x + 45$$



12. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$14x + 5 = 7x + 19$$



*Sample pretest and posttest items used in Study 6. 1-5 are conceptual items whereas 6-9 are procedural items. At the end of the test (only on the posttest), it had a survey asking for their perception of diagram use.*

1. Debby is solving an equation. Here is her first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8 - 8$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

2. Christopher is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$

$$2x + 4 - 4 = 8 - 4$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

3. Joseph is solving an equation. Here is his first solution step:

$$2x + 4 = 8$$

$$2x + 4 + 4 = 8 + 4$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

4. Tiana is solving an equation. Here is her first solution step:

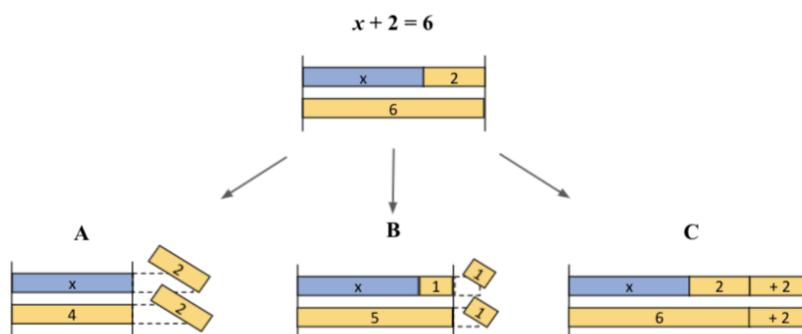
$$2x + 4 = 8$$

$$2x + 4 - 4 = 8$$

This step is:

- Valid and it gets you closer to the solution
- Valid but it does NOT get you closer to the solution
- NOT valid
- None of the above

5. These are called "tape diagrams." Tape diagrams visualize equations. The tape diagram on the top shows " $x + 2 = 6$ " and A, B, and C show possible next steps. Please answer the following questions.



5a. Which of the three next steps (A, B, and C) are **valid** (Choose all that apply)?

A

B

C

None of these is valid

5b. Which of the three next steps (A, B, and C) are **valid but do not get you closer to the solution** (Choose all that apply)?

A

B

C

None of these is "valid and does not get me closer to the solution"

5c. What is the value of  $x$  in this equation?

 2 3 4 5

6. Solve for  $x$ . Please show your work:

$$4 + x = 9$$

7. Solve for  $x$ . Please show your work:

$$3x + 4 = 16$$

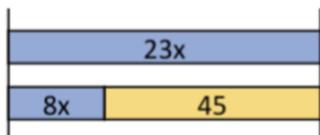
8. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$7 + 8x = 71$$



9. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$23x = 8x + 45$$



**Please let us hear about your honest thoughts about the following statements:**

Strongly disagree      Somewhat disagree      Neither agree nor disagree      Somewhat agree      Strongly agree

I liked the diagrams in the software



I think that the diagrams helped me learn to solve algebra problems



I think that the diagrams helped me solve problems faster and more accurately



I think I am good at using the diagrams in solving algebra problems



I think I am good at solving algebra problems (without diagrams)



Sample pretest and posttest items used in Study 8. 1-4 are conceptual items whereas 5-9 are procedural items. At the end of the test (only on the posttest), there was a survey asking for students' perception regarding diagram use and the transfer task.

1. If  $4x + 10 = 54$  is true, state whether each of the following must also be **true/valid**:

	Yes/True/Valid	No/False/Not Valid
$4x + 10 - 3 = 54 - 3$	<input type="radio"/>	<input type="radio"/>
$4x + 10 - 10 = 54$	<input type="radio"/>	<input type="radio"/>
$4x + 10 + 10 = 54 + 10$	<input type="radio"/>	<input type="radio"/>
$4x + 10 - 10 = 54 - 54$	<input type="radio"/>	<input type="radio"/>
$4x + 10 - 10 = 54 - 10$	<input type="radio"/>	<input type="radio"/>
$4x + 10 - 10 = 54 + 10$	<input type="radio"/>	<input type="radio"/>

2. For  $8x + 21 = 61$ , state whether each of the following will **get you closer to the solution** (will these help you isolate the "x"?):

	Yes, it will get me closer to the solution	No, it will NOT get me closer to the solution
$8x + 21 - 3 = 61 - 3$	<input type="radio"/>	<input type="radio"/>
$8x + 21 + 21 = 61 + 21$	<input type="radio"/>	<input type="radio"/>
$8x + 21 - 61 = 61 - 61$	<input type="radio"/>	<input type="radio"/>
$8x + 21 - 21 = 61 - 21$	<input type="radio"/>	<input type="radio"/>

3. Sophie, Kingdom, and Kenji are discussing what would be a good step to take for the equation below. Read each one's statement and decide whether each statement is correct or not.

$$10x - 40 = 3x - 5$$

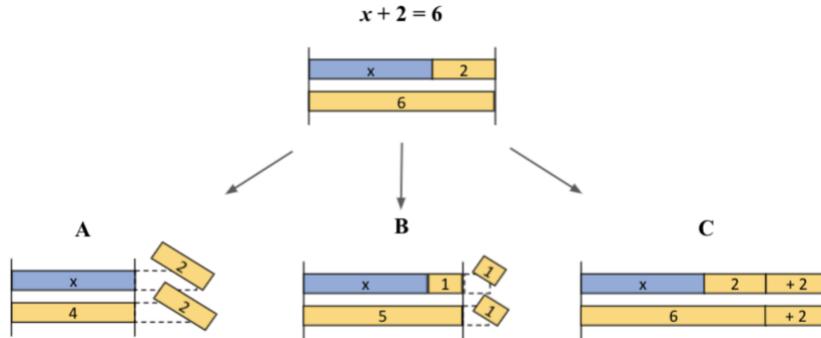
Sophie  ? We can remove -40 and -5 by adding these numbers to the equation. So, it will be  $10x - 40 + 40 = 3x - 5 + 5$

Kingdom  ? We should add 40 to the both sides of the equation, so it would be  $10x - 40 + 40 = 3x - 5 + 40$

Kenji  ? No no, you need to add 40 to the left-hand side only. So, it would be  $10x - 40 + 40 = 3x - 5$

	Correct	Not Correct
Sophie	<input type="radio"/>	<input type="radio"/>
Kingdom	<input type="radio"/>	<input type="radio"/>
Kenji	<input type="radio"/>	<input type="radio"/>

4. These are called "tape diagrams." Tape diagrams visualize equations. The tape diagram on the top shows " $x + 2 = 6$ " and A, B, and C show possible next steps. Please answer the following question.



4a. State whether each of the three next steps (A, B, and C) is valid/true

	Yes/True/Valid	No/False/Not Valid
A	<input type="radio"/>	<input type="radio"/>
B	<input type="radio"/>	<input type="radio"/>
C	<input type="radio"/>	<input type="radio"/>

5. Solve for  $x$ . Please show your work:

$$7 + 3x = 16$$

6. Solve for  $x$ . Please show your work:

$$8x + 4 = 76$$

7. Solve for  $x$ . Please show your work:

$$25x = 11x + 42$$

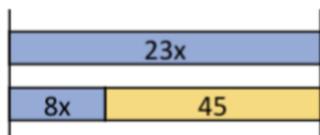
8. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$7 + 8x = 71$$



9. Solve for  $x$ . Please show your work. You can use the diagram to help your thinking.

$$23x = 8x + 45$$





**If the sum of the smallest and largest of three consecutive even numbers is 45, what is the value of the second largest number in the series?**

You can solve this problem by making a simple equation that you practiced in the software.

**HINT: If we let the smallest, unknown number be "x", we can write down the second largest number as "x + 2" and the largest number as "x + 4"!**

Now, you can choose to solve this problem without any help or you can choose to use a new kind of diagram. Would you want to use the diagram?

I will use the diagram to solve the problem

I will solve the problem without the diagram

Please say a bit more about your decision above about using diagrams. Why or why not?



## Appendix 2. A permission letter from the International Society of the Learning Sciences



International Society of the Learning Sciences

[www.isls.org](http://www.isls.org)

June 17, 2022

Tomohiro Nagashima  
Carnegie Mellon University

Tomohiro Nagashima,

On behalf of the ISLS, I am pleased to grant you, and your colleagues, permission to reuse material from the International Conference of the Learning Sciences, the International Conferences of Computer Supported Collaborative Learning, and ISLS annual meeting proceedings below:

- Nagashima, T., Tseng, S., Ling, E., Bartel, A. N., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevan, V. (2022, to appear). Students' self-regulated use of diagrams in a choice-based Intelligent Tutoring System. In *Proceedings of the Annual Meeting for the International Society of the Learning Sciences (ISLS2022)*, Hiroshima, Japan.
- Nagashima, T., Bartel, A. N., Yadav, G., Tseng, S., Vest, N. A., Silla, E. M., Alibali, M. W., & Alevan, V. (2021). Using anticipatory diagrammatic self-explanation to support learning and performance in early algebra. In E. de Vries, J. Ahn, & Y. Hod (Eds.), *15th International Conference of the Learning Sciences – ICLS 2021* (pp. 474–481). International Society of the Learning Sciences. [\[URL\]](#)
- Nagashima, T., Yadav, G., & Alevan, V. (2021). Rethinking technology-based educational studies in the evolving classroom environment: An interview study with US teachers. In E. de Vries, J. Ahn, & Y. Hod (Eds.), *15th International Conference of the Learning Sciences – ICLS 2021* (pp. 933–934). International Society of the Learning Sciences. [\[URL\]](#)
- Nagashima, T., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2020). Enhancing conceptual knowledge in early algebra through scaffolding diagrammatic self-explanation. In M. Gresalfi & I. S. Horn (Eds.), *14th International Conference of the Learning Sciences* (pp. 35-43). Nashville, TN: International Society of the Learning Sciences. [\[URL\]](#)
- Nagashima, T., Yang, K., Bartel, A. N., Silla, E. M., Vest, N. A., Alibali, M. W., & Alevan, V. (2020). Pedagogical Affordance Analysis: Leveraging teachers' pedagogical knowledge for eliciting pedagogical affordances and constraints of instructional tools. In M. Gresalfi & I. S. Horn (Eds.), *14th International Conference of the Learning Sciences* (pp. 1561-1564). Nashville, TN: International Society of the Learning Sciences. [\[URL\]](#)

This permission is provided free of charge. We kindly ask that you note that the ISLS owns the copyright of these short papers. Subsequent bibliographic details should include a notice to that effect. In addition, please include the link to the respective conference proceedings.

With my sincere regards,

Anna Keune  
Co-Chair, Publications Committee, International Society of the Learning Sciences