Comparing the Effectiveness of Supports for Collaborative Dialogic Sense-Making with Agent-Based Models

Ilana Dubovi, Ben-Gurion University of the Negev, dubovi@post.bgu.ac.il
Victor, R. Lee, Utah State University, victor.lee@usu.edu

Abstract: Collaborative dialogue has been classified by the ICAP framework (Chi & Wylie, 2014) as the highest interactive level of cognitive engagement. In this study we evaluate the balance between type of guidance and learner characteristics while collaborative learning with agent-based models takes place. Participants were students from health-care programs and were randomly assigned to one of two conditions: one group learned with agent-based models using the vicarious approach, where pairs viewed and discussed recordings of others using agent-based models; the other group explored agent-based models while discussing in pairs a set of text-based prompts. The results reveal that the vicarious-learning approach was superior to agent-based models' exploration. More detailed evaluation shows that nursing students' performance wasn't affected by the type of guidance, while less experienced in the content, non-nursing students, benefited more from vicarious guidance. Findings suggest that dialoguing while observing the dialogue of others can maximize immediate learning gains.

Keywords: agent-based models, vicarious learning, collaboration, ICAP framework, dialogue

Introduction

According to Chi's ICAP framework (Interactive, Constructive, Active, and Passive), collaborative dialoguing is the highest interactive level of cognitive engagement (Chi et al., 2018; Chi & Wylie, 2014). While differentiating modes of cognitive engagement on the basis of students' overt behaviors, dialogue can be considered truly interactive only when it consists of mutual exchanges of ideas between two (or more) individuals, resulting in new ideas that neither individual knew initially nor could generate alone. Computer-based simulations and models can provide excellent opportunities to generate constructive dialogue with the aim to promote more in-depth conceptual understanding of science content, especially when complex systems are involved (D'Angelo et al., 2014). Of special interest to this study is agent-based modeling (ABM). Joint exploration of agent-based models allows students to test their ideas, negotiate, argue a position and provide justifications, compare, and revise one another's understandings based on the dynamic feedback provided by the running simulation model (Wilensky & Reisman, 2006). Yet, to achieve these potential benefits of learning with agent-based models, there is a need for learning activities that provide appropriate guidance and support interactive dialoguing (see meta-analysis by Lazonder & Harmsen, 2016).

The current study seeks to compare types of support that facilitate constructive joint dialogue by accounting for some learner characteristics within the context of an instructional unit about diabetes mellitus pharmacology for health-care professions education. We compare a vicarious learning approach of observing collaboratively a dialogue between a learner and an instructor with one involving guided exploration and explicit written prompts to have dialogue related to a set of agent-based models. This study is intended to question the efficacy of vicarious learning with agent-based models by comparing it with guided exploration.

Vicarious learning

Learning from observing, or learning vicariously, was first proposed by social psychology research (Bandura, 1969) to evaluate learning by imitating and modeling someone's else behaviors and actions. Bruner (1986) stated that "most of our encounters with the world, are not direct encounters" (p. 122), which would seem to imply that it is possible to learn through mechanisms other than primary or firsthand experience. Such social learning is effective without the need for the observer to experience feedback directly. Bandura's introduction of the idea – vicarious reinforcement – was based on classic studies in which children were seen to imitate the aggressive behavior modeled by adults as they assaulted Bobo dolls (Bandura, Ross, & Ross, 1963).

Vicarious learning has been also explored in area of learning of cognitive behaviors as asking questions (Rummel & Spada, 2005), doctor–patient communications skills (Stegmann, Pilz, Siebeck, & Fischer, 2012), and acquiring complex cognitive skills (Chi, Roy, & Hausmann, 2008). Initial research on the differences between direct participation in a task and simple observation of it being performed was conducted by Schober and Clark (1989), in which pairs of participants communicated about a task that involved sequencing figures correctly. Simply overhearing the dialogue, rather than participating directly, resulted in poorer performance, leading

Schober and Clark to conclude that the direct participants had an advantage by being able jointly to construct a common ground of comprehension. However, in a replication of Schober and Clark's study (Chi et al., 2008), in which the vicarious leaners were required to be active, by self-explaining aloud while overhearing a tutor–tutee interaction, vicarious leaners did just as well as tutees who interacted with a tutor.

To further explore vicarious learning, Chi and her colleagues tested different conditions involving different vicarious aspects – observing monologue videos alone / observing dialogue videos alone / observing monologue videos collaboratively / observing dialogue videos collaboratively (Chi et al., 2008). The results suggest that vicarious learning through observation doesn't have to be passive but might instead be considered interactive when the conditions of *collaboratively* overhearing and observing *dialogue* between a learner and an instructor are taking place. Dialogue that incorporates mutual exchanges of ideas between individuals will result in new ideas that neither individual knew initially nor could generate alone. Therefore, the ICAP theoretical framework, which categorizes different modes of active learning, defines joint dialoguing as the highest interactive level of cognitive engagement (Chi & Wylie, 2014). The advantage of designing vicarious learning with the dual discourse approach, collaboratively observing a dialogue, is twofold. First, observing collaboratively facilitates interactions between peers that offer an opportunity for exchanges of ideas and argumentation processes which foster learning (Schwarz, Neuman, & Biezuner, 2000). Second, observation of dialogues over monologues makes it possible for vicarious learners to overhear the tutee's questions and struggles and to reflect on their own mental models using the tutor's feedback as well. As a result, learning gains of collaborative observations of dialogues were similar to learning gains following face-to-face human tutoring (Muller, Sharma, & Reimann, 2008). This study is aimed at adding to the existing domain of knowledge on the effectiveness of vicarious learning (Chi et al., 2018) by comparing pairs collaboratively observing videos vicariously with pairs following guided exploration coupled with explicit prompts to dialogue about agent-based models within the antidiabetic pharmacology education context.

Agent-Based Modeling environment

ABM is a computational modeling paradigm that emphasizes multi-level examination of complex, multi-agent systems. The ABM paradigm encodes the behavior of individual agents in a small set of simple rules so that we can specify and observe the results of these agents' individual actions and interactions. Learning through this approach focuses on entities and their actions (also called the micro level of the system), such as movement, interactions, and global flows (also called the macro level of the system), and allows students to comprehend parallel processes by which emergent phenomena form (Wilensky & Resnick, 1999). Exploration of agent-based models encourages causal emergent thinking in connecting individual behaviors with systemic patterns, thus helping students learn various scientific concepts effectively (i.e., Samon & Levy, 2017).

Yet, to achieve these potential benefits of agent-based models exploration, research has provided conclusive evidence that students need support and guidance (D'Angelo et al., 2014; Kirschner et al., 2006). However, the educational literature presents varied types and levels of guidance, from prompts and cues to specific and direct explanations of how to perform an action (Lazonder & Harmsen, 2016). There is an ongoing debate, the "assistance dilemma", about what type of guidance is adequate and whom it fits (Koedinger & Aleven, 2007, p. 261). On one side of this argument are those advocating direct instructional guidance as beneficial for novice learners (e.g., Kirschner et al., 2006); on the other side are those suggesting that less heavily guided instructions may be more productive, encouraging concept learning and transfer (e.g., Koedinger & Roll, 2012). Hmelo-Silver and her colleagues challenged this notion by suggesting that instead of contrasting these dichotomous positions on the continuum of guidance level, the questions to be asked are "under what circumstances do these guided inquiry approaches work, what are the kinds of outcomes for which they are effective . . . and what kinds of support and scaffolding are needed for different populations and learning goals?" (Hmelo-Silver, Duncan, & Chinn, 2007, p. 105).

In fact, regarding the exploration of agent-based models, it is still unclear how to balance the instructional guidance. On the one hand, agent-based models are exploratory models that are designed so that students can discover things by experimenting within a microworld; on the other hand, without providing explanations, understanding of complex systems won't become explicit but might remain implicit. Vitale, McBride, and Linn (2016) compared two forms of automated guidance – direct explanations and prompts to motivate learners (knowledge integration guidance) – to support learning of complex systems with agent-based models. They found that directed guidance for agent-based models exploration produced higher immediate learning gains than minimal guidance did. However, recent studies by Jacobson et al. (2017) and Levy et al. (2018, p. 2) on agent-based models' exploration for complex systems learning suggested that minimal guidance, or a "constrained discovery", can be valuable to conceptual learning.

The current study focuses on science learning, particularly pharmacology learning, among health-care professionals in a higher education setting. Moreover, we aimed to support science learning within health-related education, which has so far occupied only a modest space in the learning sciences. Therefore, we present the Deep Dive into Diabetes (DDD) agent-based learning environment, which was constructed with the NetLogo modeling platform (Wilensky, 1999). The proposed agent-based models, which simulate biochemical processes that represent the relevant anatomy of glucose equilibrium and the mechanisms of drug action, were constructed in a previous study (Dubovi, Dagan, Mazbar, Nassar & Levy, 2018). To explore the effect of guidance level, the DDD learning environment was constructed with two different levels of learning activities that support learning with agent-based models (Figure 1): (1) collaborative observation of a video tutorial, a vicarious approach; and (2) collaborative exploration and experimentation with agent-based models, a guided-exploration approach.

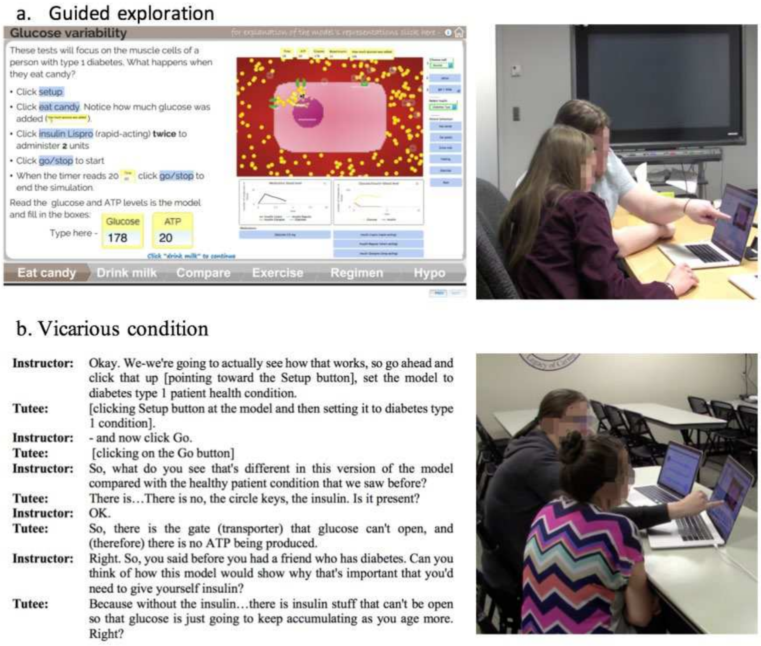

Figure 1. DDD learning environment: (a) Guided-exploration condition: screenshot, and photo of students learning with the environment; (b) Vicarious condition: example of conversation that students were overhearing while learning with the video tutorial, and photo of students learning with the environment.

Methods

Research design
This study employed a quasi-experimental pre- and post-test design with two comparison conditions to explore the effect of guidance level on inquiry learning with agent-based models. We conducted a quantitative analysis of questionnaires of all participants (70) and video analysis of 11 randomly selected dyads of students (22).

Participants and procedure
The participants were 70 health-care undergraduate students (40 students from a nursing program, 23 from a nutrition program, and 7 from a health education program) who volunteered to participate in this study. Most participants were females (61), and the average age was 23±4.85 years. Participants were randomly assigned to participate in pairs in one of two conditions: a guided-exploration (34) or a vicarious condition (36). Both conditions involved active and continuous dialogue in pairs as participants tried to interpret what was happening in the DDD learning environment. In the first condition, students were required to learn about diabetes through guided exploration of agent-based models (the guided-exploration condition); in the second condition, students learned about diabetes by observing instructor–learner dialogue and by exploring agent-based models (the vicarious condition).

There were no statistically significant differences in demographic characteristics such as grade point average (GPA) (χ^2=2.58, p = 0.46) or prior knowledge about diabetes (t=1.36, p=0.17) between the two conditions.

On average, students spent about 50 minutes learning with the vicarious-instruction and 52 minutes learning with the guided-exploration-instruction method.

Instruction design and materials

While learning with the DDD environment, students in the both groups were asked to use information from an agent-based model to solve problems about diabetes and antidiabetic pharmacology content. The *vicarious condition* group received guidance by observing a video tutorial of an instructor and a learner while exploring the agent-based models. A staff member at a university who was an advanced novice on diabetes and who was familiar with the NetLogo modeling system served as the instructor, and a graduate student who had passing familiarity with diabetes and the modeling system was the learner. This was an unscripted conversation. The study participants were able to overhear the dialogue between learner and instructor that included the instructor's explanations of agent-based models' representations; what to pay attention to while the model is running and how to interpret the model's output; and questions raised by the student which were followed by the instructor's feedback (Figure 1b). Similar to Chi et al.'s study, students could pause, reverse, and skip portions of the video; to enhance participants' active cognitive engagement while observing, we also followed Chi and her colleagues' vicarious learning design by asking learners to solve together several sub-problems which served as landmarks to the video and prompted interactive dialogue between the participants (Chi et al., 2008). The *guided exploration condition* group received less guidance while directly exploring the agent-based models. The scaffolds included information on how to set up the agent-models and on which graphs and monitors to pay attention to. As in the vicarious condition, to facilitate dialogue between pairs, students were asked to talk with one another and to solve sub-problems while exploring the agent-based models (Figure 1a). Their solutions to the problems were entered into a single shared computer. We used learning activities that we developed in a previous study (Dubovi et al., 2018).

Data collection instruments

Diabetes knowledge questionnaire

The Diabetes knowledge questionnaire was adapted from the Pharmacology Diabetes Mellitus questionnaire (PDM) developed in a previous study (Dubovi et al., 2018). The questionnaire consists of nine questions (7 multiple-choice, 2 open-ended), and evaluates understanding of biochemical glucose equilibrium, glucose disequilibrium (i.e., diabetes type 1 and diabetes type 2), and medications actions. Analysis of the Diabetes knowledge questionnaire using Cronbach's alpha yielded an internal consistency score of 0.68, which was similar to our previous report (a = 0.71) and can be considered acceptable.

Responses to the questionnaire were coded as either correct or incorrect, and the total score was calculated as the percentage of correct answers. In addition, the items on the Diabetes knowledge questionnaire were scaled by level of difficulty: four items were coded as the most difficult, and five items as the least difficult. Level of difficulty for each item was determined based on the percentage of students who correctly answered it on the post-test. Although students completed the pre- and post-tests questionnaires as individuals, their learning process was nested within their collaboration as a pair. Prior knowledge varied greatly between students as individuals as well as within pairs. To account for these variations, the analysis invoked multi-level modeling. Owing to the pre-test/post-test design used in this study, our data analysis encompassed repeated measures on individuals over time. Consequently, a three-level structure arose: both test times (Level 1) were clustered within students (Level 2), which were nested within dyads (Level 3). Although multi-level models quantified the variance across pairs, the focus of the study was on at the individual student level.

Video recordings

To assess the learning process, we recorded students' discourse and interactions with the agent-based models using screen-recording software and a separate standing video camera. For the analysis, we randomly chose 11 pairs (22 individual students); five pairs learned with the guided-exploration condition, and six pairs learned with the vicarious condition. We evaluated the frequency of accuracy of students' ideas and explanations as they learned with the DDD environment. To generate students' ideas accuracy, discourses and dialogues were carefully transcribed, iteratively reviewed, and coded in terms of the ideas, explanations, and statements expressed. This approach to the selection of knowledge elements and ideas is comparable to the approach documented in Sherin, Krakowski, and Lee (2012) and Minstrell (1982). From this, transcript excerpts were identified to illustrate some of the dialogues and ideas expressed.

Results

Diabetes knowledge questionnaire scores

Multilevel model analysis was conducted to determine the effect of pre-test scores, experimental condition (vicarious vs. guided exploration), and the student's program of study (nursing vs. other health-care field) on post-test outcomes, both independently and through examining the interactions between them. Moreover, multilevel model analysis also considers the random effect of individual and pair characteristics on factors' interactions. As shown in Table 1, after adjusting for differences between individuals and dyads, the overall post-test score for the Diabetes knowledge questionnaire was the sum of the intercept (37.318). A significant interaction of Time × Experimental condition indicates that the vicarious-condition pre-test to post-test learning gains were significantly higher than those for the guided condition (34.22 to 67.47 vs. 40.23 to 61.84).

Examining the level of the Diabetes knowledge questionnaire items independently, being in a nursing program versus another health-care program moderates the effect of experimental condition within the most difficult items (Table 1). More specifically, students from non-nursing programs made significantly higher learning gains when participating in the vicarious condition compared with the guided-exploration condition (8.93 to 60.71 vs. 18.53 to 38.84, respectively; Figure 2), whereas students from the nursing program gained knowledge from learning with both conditions, with no significant differences (Figure 2). This effect was true only for the most difficult questionnaire items; for the least difficult items, the interaction Time × Experimental condition was insignificant, meaning that students showed similar learning gains from learning with the vicarious (nursing: 51.60 to 65.39; non-nursing 57.84 to 71.63) and the guided-exploration conditions (nursing: 52.13 to 65.91; non-nursing 58.37 to 72.16).

Table 1: Three-level nested random-intercepts multilevel model predicting students' Diabetes knowledge questionnaire post-test scores

	Overall Questionnaire	The Most Difficult Items	The Least Difficult Items
Fixed Effects	**Estimate (SE)**	**Estimate (SE)**	**Estimate (SE)**
Intercept	35.279 (3.839) ***	18.535 (5.764)**	51.762 (4.389)***
Time: Post vs. Pre	21.609 (3.187) ***	20.313 (6.346)**	13.786 (2.771)*
Experimental condition: Vicarious vs. Guided	-6.010 (4.543)	-9.602 (8.428)	-0.941 (4.669)
Program: Nursing vs. Other	8.671 (3.859)*	-0.288 (7.817)	7.264 (4.706)
Interactions:			
Time × Experimental condition	11.647 (4.444) **	31.473 (9.290)***	-
Time × Program	-	28.298 (8.722)**	-
Experimental condition × Program	-	2.716 (11.029)	-
Time × Experimental condition × Program	-	-24.403 (12.304)*	-
Random Effects	**Var**	**Var**	**Var**
Dyads	79.364	65.575	231.849
Participants within Dyad	53.560	79.733	6.196
Residual	172.678	322.206	268.727

Note: Sample size is 35 dyads made up of 70 participants with 140 total observations.
Each model contains only significant interactions.
*p<0.05, ***p<0.01, ***p<0.001.

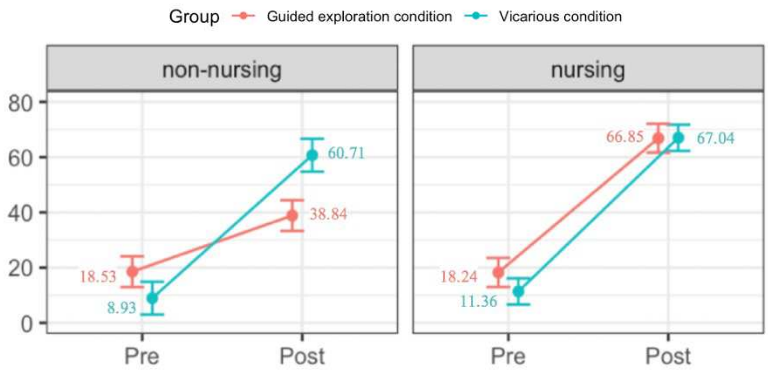

Figure 2. Visualization of significant interaction effect of Diabetes knowledge questionnaire's most difficult items, by experimental condition, among non-nursing students; and the non-significant interaction effect of the most difficult items by experimental condition, among nursing students. Depicted are estimated marginal means ± 1 standard error of the mean (SEM).

Learning process

To further examine the learning process that characterized learning with the DDD environment, we analyzed the video-recording data of 11 pairs to evaluate students' number and quality of different ideas exchanged between the partners. About 230 ideas and explanations were identified for five pairs who learned with the guided-exploration condition (average number of ideas for each pair was 48) and 305 for six pairs who learned with the vicarious condition (average number of ideas for each pair was 51). The following short discourse of a student pair assigned to the guided-exploration condition illustrates the expression of some accurate and less accurate ideas (All names are pseudonyms to protect participants' identities.):

Jessica:	Go. Now I click Go button. [setting the model to the diabetes type 1 condition]
Michelle:	So, we should have him eat candy.
Jessica:	Eat candy [clicking on the button which says "eat candy"]. It will be so much glucose.
Michelle:	So [observing the agent-based model], nothing is getting through [the muscle cell] at all.
Jessica:	Yeah, but, because there is no insulin.
Michelle:	Ahhh . . . there's no Insulin. So, like, if they drink milk or eat pasta, or fasting, it won't matter, doesn't it?
Jessica:	There's nothing. So, you do not need to exercise. There's no insulin. They still need insulin to be alive.
Michelle:	Yeah, [without insulin] they are unable to make ATP.
Jessica:	Mm-hmm. So, there's no energy and then what?!

During this episode, Jessica and Michelle articulate three basic ideas that are critical to understanding the pathological processes related to glucose equilibrium as part of the action of antidiabetic drugs. The first is insulin's role in regulating glucose metabolism: if there is no insulin, then "nothing is getting in", meaning that without insulin, glucose can't enter the GLUT4 transporters on muscle cells' membranes. This idea is correct for the muscle and lipid tissues, where glucose metabolism is mediated by insulin hormones. The second idea implicates the pathological definition of type 1 diabetes: "there is no insulin". Here Jessica expresses a nominal fact: that in type 1 diabetes there is no insulin. Next, the relationship between glucose molecules and ATP production is noted, namely the cellular respiration metabolic pathway (Glycolysis). Jessica summarizes here her third idea, that when there is no glucose within the cells, the ATP production process is impaired, which means that "there's no energy" production. This explanation is only partially correct because fatty acids and amino acids can be used for ATP production.

A close look at the correctness of students' ideas revealed that students who learned with the guided-exploration condition (five pairs) expressed 41 inaccurate ideas, whereas students who learned with the vicarious condition (six pairs) expressed only seven inaccurate explanations. Most of the inaccurate explanations with the guided-exploration condition resulted from misinterpreting what to pay attention to while exploring the agent-based models.

Discussion

The most interesting result is the advantage of vicarious learning for immediate learning achievements over the guided-exploration condition. One possible explanation for this effect is that learning with the guided-exploration condition triggered more inaccurate ideas and explanations, even though the number of ideas that students expressed was not affected by the experimental condition. Interestingly, a more precise analysis reveals that there are differential instructional effects across students' characteristics. Evaluation of the more difficult items of the Diabetes knowledge questionnaire shows that nursing students' performance wasn't affected by the type of guidance but instead benefited equally from both the vicarious-learning guidance and the guided-exploration support. However, students from the non-nursing programs, namely nutrition and health-education programs, made higher learning gains within the more difficult items of the Diabetic knowledge questionnaire following learning with the vicarious approach. Hence, non-nursing students benefited more from the vicarious approach than from the guided-exploration one. Although our results show no significant difference in prior diabetes knowledge between the nursing and non-nursing students, a plausible explanation for the difference is related to nursing students' prior experiences with the pharmacology topics from their clinical practices and curriculum. Whereas students from nutrition and health-education programs are focused on food science and on health promotion and less on pharmacology concerns, nursing training notably emphasizes nurses' need for accountable and responsible medication management, which is crucial for nursing practice (Khan & Hood, 2018). We propose that this acknowledgment of pharmacology importance as part of a professional identity better prepared nursing students for learning and for making meaning from any support that was available to them, whether vicarious or guided exploration.

As stated in the introduction, the ICAP framework predicts that the more active students are in their learning activities, the better their learning outcomes will be. According to this framework, both conditions of the current study evoked active engagement using both constructive activities (e.g., self-explanations, predictions, and model exploration) and interactive mode involving joint dialogue (Chi & Wylie, 2014). The main practical implication of this study is that learning with tutorials can be scaled up and maximized when visual displays encourage dialogue between instructor and students and when learners solve problems while observing and overhearing this dialogue collaboratively. As we showed, by evoking an active level of engagement using dialoguing, vicarious instructional design to support learning with agent-based models can be at least as efficient as the guided exploration of agent-based models. Vicarious learning provides a unique modeling opportunity for students to learn how to explore agent-based models and enables them to reflect on their own process of learning. Our results show that less-experienced students can especially benefit from the vicarious approach to support learning with agent-based models until they achieve a certain level of expertise.

This study has several limitations, specifically, the current study evaluated only immediate learning effects followed by one short intervention, using a single population. Therefore, the advantage of the vicarious approach over the guided exploration of modeling systems should be further explored. Open exploration of agent-based models is important for concept construction; our findings underline what makes working with agent-based models more or less challenging with different instruction types for different populations.

References

Bandura, A. (1969). *Principles of behavior modification.* New York, NY: Holt, Rinehart & Winston.

Bandura, A., Ross, D., & Ross, S. (1963) Vicarious reinforcement and imitative learning. *Journal of Abnormal and Social Psychology. 67*(6), 601–607.

Bruner, J. S. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., . . . & Wylie, R. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science,* doi:10.1111/cogs.12626.

Chi, M. T., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. Cognitive science, 32(2), 301–341.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist, 49(4), 219–243.

Chi, M. T., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science, 32*(2), 301–341.

D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., & Haertel, G. (2014). *Simulations for STEM learning: Systematic review and meta-analysis.* Menlo Park: SRI International.

Dubovi, I., Dagan, E., Mazbar, O. S., Nassar, L., & Levy, S. T. (2018). Nursing students learning the pharmacology of diabetes mellitus with complexity-based computerized models: A quasi-experimental study. *Nurse education today, 61*, 175-181.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and Clark. *Educational Psychologist, 42*(2), 99–107.

Khan, E. U., & Hood, P. A. (2018). Nurses' perspectives on pharmacology: why, what and at which point of the curricula should education be delivered? *British Journal of Nursing, 27*(10), 546–553.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239–264.

Koedinger, K. R., & Roll, I. (2012). Learning to think: Cognitive mechanisms of knowledge transfer. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 789–806).New York, NY: Oxford University Press.

Jacobson, M. J., Markauskaite, L., Portolese, A., Kapur, M., Lai, P. K., Roberts, G., (2017). Designs for learning about climate change as a complex system. *Learning and Instruction. 52,* 1–14.

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86*(3), 681–718.

Levy, S. T., Peleg, R., Ofeck, E., Tabor, N., Dubovi, I., Bluestein, S., & Ben-Zur, H. (2018). Designing for discovery learning of complexity principles of congestion by driving together in the TrafficJams simulation. *Instructional Science, 46*(1), 105–132.

Minstrell, J. (1982). Facets of students' knowledge and relevant instruction. In R. Duit, F. M. Goldberg, & H. Niedderer (Eds.), *Proceedings of Research in physics learning: Theoretical issues and empirical studies* (pp. 110–128). Keil, Germany: The Institute for Science Education (IPN).

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education, 92*(2), 278–296.

Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *Journal of the Learning Sciences, 14,* 201–241.

Samon, S., & Levy, S. T. (2017). Micro–macro compatibility: When does a complex systems approach strongly benefit science learning? *Science Education, 101*(6), 985–1014.

Sherin, B., Krakowski, M., & Lee, V. R. (2012). Some assembly required: How scientific explanations are constructed in clinical interviews. *Journal of Research in Science Teaching, 49*(2), 166–198. doi:10.1002/tea.20455

Schwarz, B. B., Neuman, Y., & Biezuner, S. (2000). Two wrongs may make a right . . . if they argue together! *Cognition and Instruction, 18*(4), 461–494.

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and observers. *Cognitive Psychology, 21,* 211–232.

Stegmann, K., Pilz, F., Siebeck, M., & Fischer, F. (2012). Vicarious learning during simulations: Is it more effective than hands-on training? *Medical Education, 46*(10), 1001–1008.

Vitale, J. M., McBride, E., & Linn, M. C. (2016). Distinguishing complex ideas about climate change: Knowledge integration vs. specific guidance. *International Journal of Science Education,* 38(9), 1548–1569.

Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston, IL.

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep, or a firefly: Learning biology through constructing and testing computational theories—an embodied modeling approach. *Cognition and Instruction, 24(*2), 171–209.

Wilensky, U., & Resnick, M. (1999). Thinking in levels: A dynamic systems approach to making sense of the world. *Journal of Science Education and Technology, 8*(1), 3–19.

Acknowledgments

This work was supported by AIS - Teaching with Technology Innovation Grant, Utah State University. We are especially grateful to Dr. Sarah Schwartz for her assistance with the statistical analysis.