

Boosting Knowledge-Building with Cognitive Dialog Games

Jonathan S. Herberg (jonathan.herberg@gmail.com)
Ilker Yengin (yengini@ihpc.a-star.edu.sg)
Praveena Satkunarajah (praveenas@ihpc.a-star.edu.sg)
Margaret Tan (mtan@ihpc.a-star.edu.sg)

Institute of High Performance Computing, A*STAR
1 Fusionopolis Way, #16-16 Connexis, Singapore 138632

Abstract

Dialog game tools are text chat applications which aim to structure and promote students' collaborative learning by having them select a label and sentence-opener for each message they type to their learning partner. In this experiment, we compared students' learning from discussions via a dialog game tool to their learning via a standard freechat application. Students discussed topic questions with a learning partner. They then individually completed a multiple choice test, for assessing knowledge-gain, and a short-answer test, to assess readiness for knowledge-building. Results suggest that dialog games applications lead to increased readiness for knowledge-building, in the form of integrating distinct pieces of learned knowledge, than freechat applications. Follow-up analyses suggest that the degree of concept overlap between students' dialog messages and topic keywords, as measured by a "semantic fingerprint" system, is a potentially useful metric for predicting students' knowledge-building. Implications and potential applications of our findings are discussed.

Keywords: collaborative learning; generative learning; knowledge-building; metacognition; dialog games

Introduction

One technique that aims to enhance collaborative learning activities among students, and to promote their communicative interaction skills, is to employ the dialog games approach. Dialog-game applications are computerized education-tools that structure students' interactive text chats by having them select the function of each dialog act they make. For each dialog act they also choose a sentence opener "scaffold" from a set of options available for the dialog act type. Such applications have been demonstrated to facilitate construction of structured communication behavioral patterns such as helping, information-seeking, probing, and instructing, between online learners (e.g., Ravenscroft, Wegerif, & Hartley, 2007; Wells, 2014).

Analyses of learners' dialog patterns in their use of dialog games applications suggest several avenues by which they potentially may lead to more effective collaborative learning. In particular, the structure of communication promoted by dialog games implementations may improve common understanding of the knowledge perspective of one's dialog partners, more effective and coherent

argumentation, and more critical thinking (e.g., Carlson, 2012; Weigand, 2016).

Along these lines, one possibility is that dialog games applications may encourage more metacognition. Metacognition in this context refers to thinking about knowledge states, including insufficient knowledge, whether one's own or one's learning partner. It is a core factor for self-regulated learning patterns, which involve targeting one's misconceptions and effectively integrating newly learned information with prior knowledge (Azevedo et al., 2009). In a collaborative learning context, in addition to metacognition encouraging self-correction of one's misconceptions, it may elicit explanation and re-representation of one's knowledge to one's learning partners that in turn may support the construction of more robust knowledge-representations.

In other words, several patterns of behavior encouraged by dialog games applications may align with those that promote generative learning. Generative learning is learning which goes beyond mere memorization, involving deeper cognitive processing, manipulation, and restructuring of information (e.g., Fiorella & Mayer, 2015). The outcome is new knowledge that can be applied in novel situations. Self-explaining and re-representing information in order to teach others are examples of learning strategies which can lead to generative learning. Experimental evidence supports the notion that self-explaining can increase one's integration of learned knowledge and inferring of new knowledge (e.g., Ainsworth & Burcham, 2007).

Tied to the notion of generative learning are the levels of learning in Bloom's taxonomy that go beyond remembering and understanding learning-domain information (Bloom, 1956). In particular, the "apply" and "analyze" levels involve transferring learned knowledge in order to solve problems and infer new knowledge. Related to this notion, in tutor-learner dialogs, tutor behaviors that encourage knowledge-building, or inference of new knowledge from existing knowledge, rather than shallow knowledge-telling behaviors (e.g., when the tutor immediately jumps to correct a learner's misconception, rather than eliciting the learner to figure out his or her own misconception) entail more generative learning (Roscoe & Chi, 2007). The analysis of tutor-learner dialogs by Chi et al. (2001) indicates that certain dialog patterns, namely those which are interactive in nature, (which means that they contain joint-actions), encourage more generative learning, whereas dialogs that

are dominated by the tutor lead to more shallow learning. Whereas self-explaining is a *constructive* learning activity, i.e. one that encourages knowledge inference, it is not an *interactive* constructive activity. According to Chi et al., (2001) behaviors that are at-once *both* interactive and constructive are the core drivers of effective tutor-learner interactions. That is, the most effective tutor-learner dialogs are ones in which new knowledge is *jointly* constructed for the learner. In particular, their extensive analysis of tutor-learner dialogs suggests distinct interactive patterns that define effective knowledge-construction. An example of such a pattern is a tutor providing scaffold prompts (e.g., hints and highlights of relevant information) for the learner to figure out the solution to a learning-domain question or problem. Chi et al. (2001) further crafted categories of questions intended to assess whether a learner has acquired information laid out in a learning-text (text-explicit questions), has effectively integrated information from different places in the learning text (text-implicit questions), or has successfully constructed and applied a mental model for the learning domain, not explicitly described in the learning text (model-implicit questions). Ainsworth et al. (2007) in their self-explanation learning studies have adopted some of these questions, referring to the latter two categories as “implicit” and “knowledge-inference” questions. However, we would argue that the integration of disparate pieces of domain knowledge toward figuring out the answer to a question, as opposed to arriving at the answer by mere recall, is itself also a form of knowledge-inference, even if it does not involve an implicit mental model. Thus, we regard successful answering of both text-implicit and model-implicit questions as entailing some form of knowledge-building.

Our hypothesis for the study was that the patterns of communicative interaction promoted by a dialog-games application would elicit more generative learning among peer-learners than a free chat application. We developed a dialog-games application and a control free chat application and designed an experiment to evaluate students’ collaborative learning outcomes. This included evaluating students’ basic knowledge-gain through their performance on multiple-choice items assessing (text-explicit) recall and understanding of the learning material. Critically, to test our generative learning hypothesis we assessed participants’ readiness for knowledge-building, through their performance on short-answer items that required them to either integrate pieces of existing knowledge (recalled from the learning-material) or to infer the answer by applying an accurate implicit mental model based on recalled learning-material information. We utilized three text-implicit questions for the first category of knowledge-building questions, and three model-implicit questions for the second kind. Our study thus extends prior research by investigating whether specifically the scaffold functions of dialog game applications enhance collaborative learning and increase the potential for knowledge-building. Additionally, we also explored the possibility of applying a natural-

language processing system to obtain dialog metrics that effectively predict better knowledge-building from students, in order to investigate the feasibility of integrating such features into dialog-games tools.

Method

Participants

Participants included in the analyses were 56 9th grade students across three secondary schools in Singapore. Signed parental consent was obtained for these students to participate at a pre-scheduled school-day time in school classrooms or computer labs that had been made available for the experiment, with laptops set up at desks in the rooms.¹

Materials

The learning domain was the human circulatory system that was adapted from the Chi et al. (2001) peer-tutor dialog study. Each of the 13 subsections was designed on the computer screen to describe each topic (e.g., “The Blood Flow in the Heart”). Diagrams were added to facilitate comprehension. Bullet points under the diagrams described the main concepts.

The topic questions that students discussed via text chat are shown in Table 1. Topic 1 corresponds to questions in the Chi et al. taxonomy which require integration of distinct pieces of explicitly learned knowledge. That is, the two sub-questions for Topic 1 are text-implicit knowledge-building questions. Topic 2 question, in addition, requires the correct mental model of circulation (as a double-loop) to answer the question effectively. Thus, it is a model-implicit knowledge-building question. Topic 3 provided students a general discussion about the learning-domain concepts.

There were pretest and posttest multiple choice questions to gauge students’ prior knowledge, and their recall and understanding of the learning material. Also for the posttest, students received six knowledge-building questions. They are shown in Table 2. The first 3 are text-implicit questions, and the latter 3 are model-implicit questions, as developed and utilized by Chi et al. (2001) and Ainsworth (2007).

There were two conditions of the dialog games text chat tool employed for this experiment. For the Scaffold condition, the application included dialog act labels for

¹ Three of the 62 students who initially participated were excluded from the analyses due to a technical error (one leading to a posttest log not being created, and the other to one of the topics between a pair not being discussed). There was a procedure error for two additional participants (i.e. they had kept their learning-material window open and used it for the posttest). Lastly, one student withdrew participation assent during the posttest. In addition, of the remaining 56, due to an ID entry error one participant lacked a pretest log, and therefore was excluded from the multiple-choice question analysis, and for two participants we could not link their short-answer log IDs to their dialog screen IDs; they were excluded from the dialog metric analyses.

students to select, and corresponding sentence openers. These message types were based on speech-act theory and were adopted from those used in other dialog game implementations (Weigand, 2016). Students were also provided with a sheet that defined the different dialog acts to guide them (see Table 3). Table 3 also shows examples of sentence-openers that students could choose for each dialog act. Figure 1 illustrates the design of the dialog game window, with labels numbered indicating the steps for entering and sending a dialog message, as follows: (1) The topic question that defines the parameters of a given dialog is at the top of the screen. (2) Users may click on a bubble next to a dialog message to make a reply to the specific message (which can also be used to reply to earlier messages in the chat history). Reply messages are indented relative to the original message. If no reply bubble is clicked, the entered text message will appear below all the text messages in the chat window, with no indentation. (3) Users select one of the six communicative act labels, and then select a linked sentence opener from a dropdown menu. The selected sentence opener appears at (4). (5) Users type in the rest of their message into the text box. Note that only one user may type into his or her text entry box message at a time. If it is the other user's turn, the shadow text in the box says "Please wait your turn." If it is the given user's turn, it says "Enter your text." In addition, if a user has failed to first select a dialog act label and sentence opener, on clicking the text entry box a reminder message will appear, and the user is unable to type into the box until making these selections. (6) When a user has completed a message, he or she clicks the "Send" button.

The Freechat application was of similar design and appearance as the Scaffold application, and included the turn-by-turn use features, but did not feature the dialog act label buttons and sentence-opener display. Thus, users took turns simply entering messages, without the scaffold steps.



Figure 1: Dialog Game Screen.

Table 1: Topic discussion questions

Topic No.	Discussion Question(s)
1	a) Why do we have valves in veins, but not in arteries and capillaries? b) Why don't we have valves in pulmonary veins?
2	Why do we sometimes refer to the heart as a "double pump"?
3	What do you think are the most interesting aspects of the structure and function of the human circulatory system? Please discuss.

Table 2: Posttest questions to assess Knowledge-Building

Item No.	Short-Answer Question
1	Why is there an artery that carries deoxygenated blood?
2	Why do vessels get increasingly smaller as they get close to the body cells, and increasingly larger as they get nearer to the heart?
3	In which kind of blood vessels (arteries, veins, or capillaries) is the blood pressure the lowest? Why?
4	Why is your right ventricle less muscular than your left ventricle?
5	The artery that carries blood from the right side of the heart to the lungs (the pulmonary artery) carries about the same amount of blood as the artery that carries blood from the left side of the heart to the rest of the body (aorta). Why do they carry the same amount of blood?
6	What would happen if the valves between the atria and the ventricles got stuck open and wouldn't close?

Table 3: Descriptions of communicative act labels, with example sentence-opener choices (Scaffold condition)

Dialog Act	Description	Example Sentence Opener Choices
Information	To provide or describe relevant facts or knowledge.	Let me explain... Some facts are... My understanding is that...
Propose	To bring up a new idea to consider.	I suggest that... Let us focus on... I think it makes sense to...
Challenge	To argue against, or provide evidence against a dialog statement.	I disagree because... A counter-argument is... An alternative view is...
Question	To ask your dialog partner about something you don't know.	Why is it... Can you explain... What do you think about...
Agreement	To agree with a statement made by your dialog partner.	I agree, ... Good point, ...
Support	To argue for, or provide evidence for a dialog statement.	I think this view is supported by, ... To give an example, ...

Procedure

Students were randomly assigned to the Scaffold condition, involving the text chat application that required them to select dialog act labels and sentence openers, or to the

Freechat condition. There were 28 students for each condition. Within each condition, the students were again randomly assigned to dialog-discussion pairs. Students in each condition were taken to separate rooms for the study (Scaffold or Freechat). To minimize verbal and indirect interaction, no student sat next to any other student. Each of two experimenters was also randomly assigned to conduct the session for each condition.

In each study session room, pre-arranged laptops were placed on the desks. The experimenter overviewed the session, which consisted of the following tasks:

- 1) Students were given up to 7 minutes to individually complete the multiple-choice pretest (could click “submit” if they finished early). (The timer for all tasks was viewable at the top of the application window).
- 2) Following the pretest, students were taken to the learning material screen where they had 15 minutes to read and study the learning material.
- 3) Then the experimenter went over how to use the system. For the Scaffold condition, the experimenter went over the different communicative act labels, and the steps for entering in a message including a sentence opener. Students also received a dialog act description sheet (Table 3).
- 4) The students were given a five-minute demo dialog session to help them get accustomed to the application.
- 5) Next, the students (with their randomly assigned learning partner) discussed the dialog questions for the 3 topics. For both conditions, students took turns entering in a dialog message. They could also open a pop-up window that contained the learning-material, which they could refer to for the discussions. For each topic, students had a 10 minute dialog discussion.
- 6) Following the end of their dialogs discussion, the students completed the post-test individually. These consisted of the same multiple-choice questions as in the pre-test (6 minutes). In addition, they had to answer the short-answer questions (as in Table 2) to assess knowledge-building, for which they were given 25 minutes. For each portion of the posttest, students could click a “submit” button if they finished early.

Measures

Knowledge-gain. To assess students’ knowledge gain from reading the learning material and engaging in the dialog discussions, their scores on the posttest multiple choice (out of 10 points) were compared to their pretest scores.

Knowledge-building. A scoring guide was developed that allowed for 2 points maximum on each of the three text-implicit questions, and 3 points maximum on each of the three model-implicit questions. Basically, a point was awarded for each piece of information relevant for inferring the answer to the question, and for each correct inference.

For example, for Question 4, one point would be awarded for an accurate description of the function of the left ventricle, one for the right ventricle, and one point for the inference that the right ventricle doesn’t need to pump blood with as great force as the left, as the blood travels less distance. Two raters, familiar with the scoring guide and the learning material and related concepts, scored participants’ answers to these questions. They were kept naïve to the experimental condition for all the short-answer logs. The scores were averaged across the two raters. The intraclass correlation for absolute agreement on the items was computed as $ICC(1, 128) = 0.87$ for the text-implicit items and $ICC(1, 128) = 0.96$ for the model-implicit items.

Topic adherence. We conducted exploratory follow up analyses that utilized the “semantic fingerprint” system developed by the Cortical.io Company (with the API available on their website). The goal was to assess the feasibility of utilizing natural-language processing methods to predict students’ capacity for knowledge-building (short-answer performance) from their dialog messages. Such functions, if predictive, could be useful to incorporate into dialog game applications, for teachers and students to track (in an automated fashion) learning outcomes implicitly from dialogs. The Cortical.io system represents the meaning of words in terms of their distributional overlap in a large linguistic corpus (i.e., Wikipedia). Its theoretical basis is the notion of distributional semantics, or “word spaces” (e.g., Sahlgren, 2006). The more frequently that words co-occur in near proximity in the corpus, the higher is their computed “semantic fingerprint overlap.” The metric can also be extended by the system to compute the degree of semantic fingerprint overlap among text segments and documents, rather than of single words. For implementation details, refer to De Sousa Webber (2015).

Dialog file inputs were first corrected for spelling errors and abbreviations. What we refer to as “topic adherence” is, for each topic dialog and participant, the semantic fingerprint overlap between the participant’s dialog messages (entered into the system as a single “document”) and pre-selected keywords intended to represent important concepts related to the topic question. Refer to Table 1 for the Topic questions. For Topic 1, the keywords were: “valves,” “veins,” “arteries,” “capillaries,” “pulmonary,” and “pressure.” For Topic 2, they were: “heart,” “lungs,” “oxygen,” “blood,” and “pump.” For Topic 3, they were “valves,” “veins,” “arteries,” “heart,” “lungs,” “oxygen,” “blood,” and “circulatory.” The additional dialog metrics of mean number of words-per-turn, and total number of turns, were used.

Findings and Discussion

Knowledge-gain scores

Figure 2, on the two pairs of bars on the left, shows the mean scores across the Scaffold and Freechat conditions on the pre-test and post-test multiple choice for assessing

students' level of recall and understanding of the domain material. It also displays the proportion-scores, so that tests with different scales can be displayed on the same chart. Participants did not differ significantly on their pretest scores, $t(53) < 1$. Across both conditions, participants showed improvement on their post-test multiple-choice scores relative to their pre-test scores, $t(55) = 5.22, p < .001$, with an effect size of $d = 0.76$. The knowledge-gain (post- minus pre- test score difference) in the Scaffold condition ($M = 1.07$) did not significantly differ from the Freechat knowledge-gain ($M = 1.14$), $t(53) < 1$. The two conditions also did not differ significantly on the mean post-test multiple-choice scores, $t(53) < 1$.

Knowledge-building scores

To assess our hypothesis of increased knowledge-building for the Scaffold condition, we first conducted a MANOVA on the text-implicit and model-implicit scores. There was an overall effect of condition, $F(2, 53) = 3.19, p < .05$. Figure 2, on the two pairs of bars on the right, shows the mean proportion-scores across the two sets of knowledge-building questions (text-implicit and model-implicit). The follow-up tests indicated no effect of condition on the model-implicit questions, $t(54) < 1$. However, for the text-implicit questions, the mean score was higher in the Scaffold than the Freechat condition, $t(54) = 2.39, p = .02$, with an effect size of $d = 0.64$.

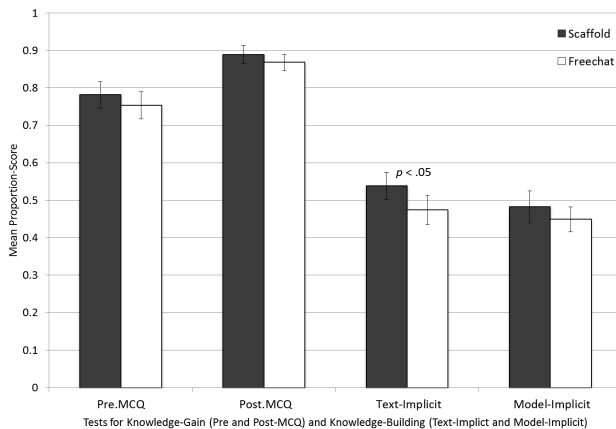


Figure 2: Mean scores (+/- SE) in the Scaffold and Freechat for the multiple-choice pretest and posttest, and the knowledge-building tests (text-implicit and model-implicit).

Dialog metrics for knowledge-building

We conducted follow-up multiple-regression analyses to explore whether the topic adherence scores obtained by the semantic fingerprint system, along with the metrics of words-per-turn and number of turns, could be of use for predicting students' readiness for knowledge-building (i.e., their short-answer scores). Scores were averaged for each participant across the three dialog topics. Tables 4 and 5

show the regressions separately on the Scaffold and the Freechat cases, respectively. For the Scaffold condition, the overall regression trends toward statistical significance, and the coefficient for topic-adherence reaches statistical significance. Total-turns trends in the direction of predicting increased knowledge-building scores. For the Freechat condition, the overall regression also trends toward statistical significance, but with a non-significant coefficient for topic adherence, and with the total-turns coefficient trending in the direction of predicting *reduced* knowledge-building. Overall, across both regressions words-per-turn appears to be a relatively weak predictor.

Table 4: Multiple regression for predicting knowledge-building (Scaffold condition)

Predictor	B	SE B	β	t	p
Topic adherence	19.53	7.64	0.55	2.56	0.02*
Words-per-turn	0.06	0.07	0.24	0.97	0.34
Total turns	0.42	0.21	0.52	1.96	0.06

$R^2 = 0.27, F(3, 22) = 2.72, p = 0.07$

Table 5: Multiple regression for predicting knowledge-building (Freechat condition)

Predictor	B	SE B	β	t	p
Topic adherence	4.62	6.09	0.16	0.76	0.46
Words-per-turn	-0.01	0.04	-0.08	-0.34	0.74
Total turns	-0.13	0.08	-0.41	-1.62	0.12

$R^2 = 0.23, F(3, 24) = 2.32, p = .10$

Discussion

Our hypothesis was partially supported. Namely, students in dialog-games interactions to discuss topic questions in the learning domain exhibited a higher readiness for knowledge-building, in the form of making text-implicit inferences, than students in the freechat discussions. There was no significant improvement on model-implicit questions. The increased knowledge-building readiness was also over and above any knowledge-gain, which did not significantly differ between conditions.

In addition, the multiple-regression results suggest that natural-language processing methods may hold some promise in producing dialog metrics with predictive utility for knowledge-building. The predictive value (in terms of the standardized Beta coefficient) of our topic-adherence metric was particularly more prominent for the Scaffold condition than for Freechat condition. Also of interest, though more caution is warranted for interpretation of non-statistically significant trends, is that the total-number of dialog turns went in the direction of predicting more knowledge-building for Scaffold condition, and less knowledge-building for Freechat condition. These trends

may be indicative of qualitative differences in the nature of dialogs with versus those without scaffolds, with a tendency for scaffolds to raise the potential learning value of each dialog turn, and to increase the potential knowledge-building when dialog partners jointly discuss core concepts in the learning domain. One possibility is that the dialog game scaffold functions in effect promote more self-explanation in the process of developing explanations and arguments to one's dialog partner. An extensive study of collaborative learning dialogs by Asterhan & Schwarz (2009), on the other hand, suggests that the process of argumentation itself may be essential for driving conceptual change from the joint construction of explanations. In the current context, if the dialog game scaffold functions encouraged more structured argumentation, this would open the door for dialogs that are more focused on the main topic concepts to generate improved conceptual understanding of the learning domain.

The conceptual foundation for applying the framework of dialog-games to learning is grounded in the notion of learning as a dialectical, social, and interactive process (cf. Mercer & Littleton, 2007). Structuring a learning-discussion as dialog-game is therefore seen as a means to encourage effective argumentation and critical thinking (e.g., McAlister, Ravenscroft, & Scanlon, 2004). In terms of Bloom's taxonomy, the potential, more immediate benefits of dialog-games can be viewed as focused on the application and analysis levels of learning. However, effective learning at these levels requires first a solid groundwork of basic understanding of concepts in a learning domain, and in turn takes time. Reaching even higher levels of learning, and unlocking creativity, is an ever increasing long-term process (cf. Bloom, 1956). Thus, dialog games may be beneficial for developing students' creativity, but this would need to be evaluated by an extended use of such applications for learning, e.g. over weeks or months.

Along these lines, one limitation of the current study is that it was a "single-shot" learning and evaluation session. For generative learning more time for absorbing, processing, and transforming information may be an essential element (Fiorella & Mayer, 2015). Thus, even on the text-implicit questions, for which there was a medium-sized effect for the difference between conditions, the mean proportion of total points obtained was for both conditions only about half of the total possible. In addition to being constrained by time for the current study, another note is that dialog-games are often applied for conversations among small-groups (Ravenscroft, 2007). It is possible that learning-dialogs for groups of 3 or 4 may allow for more argumentation and perspective-taking opportunities than two-way dialogs. Future research directions are indicated for "scaling" up dialog-games applications for knowledge-building, both in terms of time (over a long-term learning period) and in terms of group-size (e.g., from learning-pairs to learning-groups). Such extensions may lead to larger-scale knowledge-building effects, and increase

the predictive value of dialog metrics for knowledge-building.

References

- Ainsworth, S., & Burcham, S. (2007). The impact of text coherence on learning by self-explanation. *Learning and instruction, 17*(3), 286-303.
- Ainsworth, S., & Loizou, A.T. (2003). The effects of self-explaining when learning with text or diagrams. *Cognitive Science, 27*(4), 669-681.
- Asterhan, C. S., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications from protocol analyses of peer-to-peer dialog. *Cognitive Science, 33*(3), 374-400.
- Azevedo, R., Moos, D. C., Johnson, A. M., & Chauncey, A. D. (2009). Measuring cognitive and metacognitive regulatory processes during hypermedia learning: Issues and challenges. *Educational Psychologist, 45*(4), 210-223.
- Bloom, B. S., Engelhart, M. D. Furst, E. J., Hill, W. H.; Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. New York, NY: David McKay Company.
- Carlson, L. (2012). *Dialogue games: An approach to discourse analysis* (Vol. 17). Springer Science & Business Media.
- Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science, 25*(4), 471-533.
- De Sousa Webber, F. (2015). *Semantic folding theory and its application in semantic fingerprinting*. White paper retrieved from <http://www.cortical.io>.
- Fiorella, L. & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*(4), 717-741.
- Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge.
- Ravenscroft, A. (2007). Promoting thinking and conceptual change with digital dialogue games. *Journal of Computer Assisted Learning 23*(6), 453-465.
- Ravenscroft, A., Wegerif, R. & Hartley, R. (2007). Reclaiming thinking: dialectic, dialogic and learning in the digital age. *BJEP Monograph Series II, Number 5-Learning through Digital Technologies, 1*(1), 39-57.
- Roscoe, R. D. & Chi, M. T. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77*(4), 534-574.
- Weigand, E. (2016). *The dialogic principle revisited: Speech acts and mental states* (pp. 209-232). Springer International Publishing.
- Wells, S. (2014). Supporting argumentation schemes in argumentative dialogue games. *Studies in Logic, Grammar and Rhetoric, 36*(1), 171-191.