Failure to Replicate Using Dialogue Videos in Learning: Lessons Learned from an Authentic Course

Lu Ding, Joshua Adams, Michelle Stephens, Sara Brownell, and Michelene T. H. Chi
luding@asu.edu, jradam12@asu.edu, mstephens@asu.edu, sbrownel@asu.edu, mtchi@asu.edu
Arizona State University

Abstract: Previous laboratory studies have found that students learned significantly more from dialogue videos compared to monologue videos. However, these studies were conducted in laboratory conditions, which do not adequately reflect the characteristics of an authentic course environment. Therefore, we replicated the study in an actual blended biology course to test if the benefits of dialogue videos still remain. Contrary to laboratory studies, no significant difference was found in students' learning between the two groups. We then posited that the integrated content, chose of tutees, and students' preference to monologue videos may have caused the discrepancies in the findings. Relevant evidence was provided to support our hypotheses. Implications and limitations were also discussed.

Learning by observing others learn is a promising instructional approach for learning. Studies conducted on this approach have demonstrated the effectiveness with respect to learning gains of having students observe an expert tutoring a novice as they working through a problem compared to the conventional style of only an expert working through a solution in a monologue format (Driscoll et al., 2004). We henceforth call the first type of videos as dialogue videos, and the second type of videos as monologue videos.

Videos are the most common instructional approach in online learning or blended learning (Kay, 2012). With the dramatic growth of online or blended learning in higher education, it is imperative to investigate the potential benefits of the application of dialogue videos in an online/blended college-level course environment. However, most of the previous studies were conducted in lab conditions, which were not an accurate representation of an authentic learning environment. The current paper aims to fill the gap in the existing literature by presenting a replication of using dialogue videos in an authentic blended STEM course.

Relevant literature

Learning from observing dialogues

For decades, numerous studies have been carried out to prove that individuals can learn actions or behaviors from observing others. For instance, aggressive behaviors can be learned by observing others displaying them; novices can learn work-related skills via observing experts performing work activities; children can imitate behaviors by observing adults acting.

Can one also learn cognitive-related skills from observing? Observing just the overt outputs of learning alone presents many challenges. One would not see the underlying many-to-one cognitive processes that underpin learning (Chi & Bjork, 1991). For example, when an expert is solving a math problem, some covert reasoning processes cannot be demonstrated by her overt behaviors. To overcome the many-to-one cognitive processing issue, one possible solution is that the expert explains her thinking while solving the problem (Collins, Brown, & Holum, 1991). However, studies have shown that experts are notorious for not being able to convey their entire thinking process to novices (Nisbett & Wilson, 1977), and novices cannot learn very well from hearing experts' didactic-like explanations (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). One reason why a novice cannot learn well from an experts' didactic-like explanations is because the expert cannot gauge the novice's understanding level (Chi, Siler, & Jeong, 2004).

Observing dialogues between an expert tutoring a novice to solve problems poses a possible solution. Engaging a novice in dialogue with expert can force the expert to bring their cognitive level down to level of the novice as they exchange ideas, thus reaching a common ground by which a student can observe both the outputs and cognitive processes that accompany expert-level thinking. During the dialogues, the misconceptions expressed by the novices and the correcting information provided by the experts enable observing students reflect on their understandings, and thus learning occurs (Chi, Kang, & Yaghmourian, 2017).

Previous studies of observing dialogue videos

Dialogues between a teacher and a student are critical in observers' learning (Cox, McKendree, Tobin, Lee, & Mayes, 1999). Observers who watched the dialogue videos in dyads learned substantially more than those who

watch monologue videos in solo or in dyads (Chi et al., 2017). Several studies have shown that even when the observers watch the dialogue videos in solo, they still can learn more in comparison than the observers who watched the monologue videos individually. For example, Driscoll and colleagues (2004) found that students who overheard the dialogues between a virtual tutor and a tutee wrote significantly more content than those students who were in monologue condition. Similar results were found in two other studies (Muller, Bewes, Sharma, & Reimann, 2008; Muller, Sharma, Eklund, & Reimann, 2007), in which observers watched dialogue videos outperformed those observers who watched monologue videos. However, all of these studies were conducted in laboratory conditions. It is necessary to replicate these studies in an authentic course, and to investigate if the benefits extend to the online and blended learning environments.

Methods

Participants and study site

This study was conducted in an upper level biology class. Two-hundred and ten students consented to be enrolled in the study, in which 49.1% of the students were Caucasian, and the following were Hispanic (19%), Asian (18.1%), and African American (4.0%). The rest of the students indicated their ethnicity as other. The majority (87.2%) of the students declared their major as Biology (Medical) Science, with 3.5% students in Genetics, and the rest of the students were either from Psychology or the Public Health department.

Video creation and procedures

Eight dialogue- and eight monologue-videos were filmed. In the dialogue videos, the course instructor tutors a tutee who works through a set of biology problems, asking questions. Four students who completed the course in the previous year were recruited as tutees for creating dialogue videos. Each tutee filmed two videos: one from a first set of four and one from a second set of four. The monologue videos only present the instructor solving the same set of biology questions. Therefore, both dialogue- and monologue videos use the same set of biology questions. Each week, a worksheet containing the same questions in the videos was handed out to each observer student. Observer students watched the assigned videos and completed the worksheets on Thursday as pre-class assignments, and attended a recitation on Fridays. Students were randomly assigned to two groups, each observing one type of video for the first four weeks and the other for the second four weeks. That is, one group of students watched dialogue videos for the first four weeks, and monologue videos for the second four weeks. The other group of students first watched monologue videos and then switched to dialogue videos. The topics covered in the videos including Homeostasis, Information Flow, and Adipocyte Cells, etc.

Measures

Eight quizzes were created by the instructor to measure students' learning from the videos. Each quiz consisted of 10 to 12 multiple-choice questions, totaling 89 questions. Each week, the first ten minutes of Friday recitation were devoted to in-class quizzes. At the end of week 8, students were asked to indicate their preferences for the videos, and explain why.

Results

Because all students viewed both monologue and dialogue videos, we used a repeated-measures model to compare students' aggregate dialogue scores to their aggregate monologue scores. To address the issue of missing data, we used a mixed model to estimate parameters with maximum likelihood estimation and a compound symmetry covariance structure specified to make the model an analogue of a repeated measures ANOVA. We found no significant difference between student performance when using monologue videos ($M = .80$, $SD = .09$) and dialogue videos ($M = .790$, $SD = .10$); $F = 2.26$, $p > .05$.

Prior laboratory studies suggested that the dialogue videos were more beneficial to learning the tasks that require higher-order thinking skills (Muller et al., 2008, 2007). Therefore, to further test if the dialogue videos influenced on students' performance on higher-order cognitive tasks, 16 questions that required students to transfer their knowledge to new scenarios were identified by two experts together in biology field. However, using the same statistical model specified above, the results revealed a significant trend, in which the monologue group ($M = .88$, $SD = .12$) performed better than the dialogue group ($M = .86$, $SD = .13$); $F = 3.90$, $p = .05004$.

Possible reasons of a failure and implications

Why would the studies in laboratory conditions favor the dialogue videos but in an authentic course yield a failure? Our discrepant findings may be attributable, in part, to several factors, including the nature of authentic courses,

design of the dialogue videos, and students' reactions to the videos. In this section, we attempt to discuss the possible reasons and their implications.

Integrated content

Unlike a laboratory experiment, the learning content in an authentic course is always more likely to be integrated. That is, the content presented in later videos was more likely to be built upon previous videos. Students may have already established clear understandings of the concepts from other learning materials or the previous videos, thus diminishing the main benefits of the dialogue videos (i.e., reflect on misconceptions).

In comparison to Muller and colleagues' (2007, 2008) studies, in which the dialogue videos were used for Newton's First and Second Laws which are notoriously difficult to learn, the content that covered in our videos was relatively less difficult. Due to the difficulty, Newton's First and Second Laws easily evoke misconceptions (Mayer, 2004), and a clear understanding cannot be easily established. The dialogue videos provided great opportunities to the students in Muller and colleagues' (2007, 2008) studies to reflect on their misconceptions. Therefore, we proposed that dialogue videos should be used for introducing novel concepts that are more likely to be misunderstood and require higher-order cognitive processes to maximize their benefits.

Video design

Including tutoring in videos aims to elicit more dialogues between experts and novices, and hence scaffolding the cognitive process for problem solving. Observing students can thus learn from the process of an expert correcting errors made by a novice. However, the tutees in the dialogue videos for the current study completed the course in the previous year. Compared to the observing students, the student tutees in the videos were more familiar with the concepts covered in the videos. Therefore, the number of questions asked by the tutees were less than what one would expect the observing students would ask. We thus proposed that the discrepancy in the findings of this study and the literature may have arisen from the dialogue level between the instructor and the students in the videos. To test our hypothesis, we analyzed the dialogue videos based on the instructor-student interaction levels. Each of the videos was segmented into episodes based on the questions (i.e., each video consisted of two to six episodes). The episodes were coded as no interaction, less interaction, and rich interaction. Only one question and one feedback between the instructor and the student was coded as less interaction, whereas the episodes contained at least two questions and feedback were coded as rich interaction.

Two researchers coded the first two dialogue videos independently, and then met to discuss any discrepancies in their coding until 100% agreement was achieved. The rest of the six videos were analyzed by the second researcher. A total of 36 episodes were identified from the videos. Seventeen episodes (47.2%) were categorized as rich interaction, 17 episodes (47.2%) were less interaction, and two episodes (5.6%) had no interaction. We, therefore, posited that the benefits of dialogue videos may have been minimized due to only less than half of the videos containing rich interactions. In future research, to maximize the advantages of dialogue videos, two approaches can be taken. First, students who have never taken the course and are not familiar with the topics in the course should be recruited as tutees for the videos to promote richer interaction that better matches the mastery levels of the observing students enrolled in the course. Second, the dialogue between instructors and students in videos can be specifically scripted so that more rich interactions can be included in the videos (Muller et al., 2007).

Preference of monologue videos

Among the students who indicated a preference for the types of videos they observed, 59.9% (N = 124) of the students reacted favorably to the monologue videos, and only 20.3% (N = 42) of the students favored dialogue videos. For decades, a substantial body of studies have evidenced that positive affect improves students' performance on higher-order cognitive tasks (see Isen, 2008, for a review). We posited that the strong preference of the monologue videos may have influenced students' cognitive processing. Therefore, no learning differences between the two groups were detected.

To find out if this is true, we analyzed students' responses to the "why" question for their preferences. Two researchers applied thematic analysis on the responses independently (Braun & Clarke, 2006), and met to discuss the discrepancies in their codes until 100% agreement was reached. The two main factors that influenced students' preferences were: 1) the monologue videos were more straightforward (N = 69), and 2) the dialogue videos were more likely to create confusion when students in the videos made mistakes (N = 42). However, confusions can actually inspire greater depth of cognitive processing if appropriately regulated (Craig, Graesser, Sullins, & Gholson, 2004). A failed confusion regulation is more likely to yield negligible learning (D'Mello, Dale, & Graesser, 2012). In our study, when confusion occurred, there was no direct channel for the students to immediately resolve it. A possible solution to overcome this issue can be allowing students to observe the dialogue

videos in dyads. Therefore, when confusion occurs, students will have peers to discuss and to resolve the confusion immediately.

Conclusion and limitations

The goal of this paper was to replicate laboratory studies of using dialogue videos to enhance observing students' learning. However, no significant difference in learning was found between the dialogue group and the monologue group in this replication. We therefore provided three possible reasons to explain why there was a discrepancy in findings between prior laboratory studies and our replication study. We proposed possible solutions to overcome these issues for future research.

In addition, we need to highlight some limitations of this study. One pressing concern is that there is no log data of the students. No instructional material can be beneficial if students do not utilize it. Due to the lack of log data, the analyses for student learning outcomes were based on the suppositions that all students have watched the videos. Second, since students were enrolled in the same course, there are chances that students from the two groups discussed their video types. Thus, a threat to internal validity of the study may exist.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chi, M. T. H., & Bjork, R. A. (1991). Modeling expertise. In D. Druckman & R. A. Bjork (Eds.), *In the mind's eye: Enhancing human performance* (pp. 57–79). Washington, DC: National Academy Press.

Chi, M. T. H., Kang, S., & Yaghmourian, D. L. (2017). Why students learn more from dialogue- than monologue-videos: Analyses of peer interactions. *Journal of the Learning Sciences*, *26*(1), 10–50. https://doi.org/10.1080/10508406.2016.1204546

Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition and Instruction*, *22*(3), 363–387. https://doi.org/10.1207/s1532690xci2203_4

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533. https://doi.org/10.1016/S0364-0213(01)00044-1

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, (12), 6–46.

Cox, R., McKendree, J., Tobin, R., Lee, J., & Mayes, T. (1999). Vicarious learning from dialogue and discourse: A controlled comparison. *Instructional Science*, *27*, 431–458. https://doi.org/10.1023/A:1003489631631

Craig, S. D., Graesser, A., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media*, *29*(3), 241–250. https://doi.org/10.1080/1358165042000283101

D'Mello, S., Dale, R., & Graesser, A. (2012). Disequilibrium in the mind, disharmony in the body. *Cognition and Emotion*, *26*(2), 362–374. https://doi.org/10.1080/02699931.2011.575767

Driscoll, D. M., Craig, S. D., Gholson, B., Ventura, M., Hu, X., & Graesser, A. C. (2004). Vicarious learning: Effects of overhearing dialog and monologue-like discourse in a virtual tutoring session. *Journal of Educational Computing Research*, *29*(4), 431–450. https://doi.org/10.2190/Q8CM-FH7L-6HJU-DT9W

Isen, A. M. (2008). Some ways in which positive affect influences decision making and problem solving. In M. Lewis, J. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (pp. 548–573). New York, NY: Guilford Press.

Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, *28*(3), 820–831. https://doi.org/10.1016/j.chb.2012.01.011

Mayer, R. E. (2004). Teaching of subject matter. *Annual Review of Psychology*, *55*, 715–744.

Muller, D. A., Bewes, J., Sharma, M. D., & Reimann, P. (2008). Saying the wrong thing: Improving learning with multimedia by including misconceptions. *Journal of Computer Assisted Learning*, *24*(2), 144–155. https://doi.org/10.1111/j.1365-2729.2007.00248.x

Muller, D. A., Sharma, M. D., Eklund, J., & Reimann, P. (2007). Conceptual change through vicarious learning in an authentic physics setting. *Instructional Science*, *35*(6), 519–533. https://doi.org/10.1007/s11251-007-9017-6

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259.

Acknowledgements

We are grateful for funding for this project provided by the National Science Foundation: Grant No. 1504893.