Title
Learning by Observing Tutorial Dialogue versus Monologue Collaboratively or Alone

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN
1069-7977

Authors
Muldner, Kasia
Dybvig, Keith
Lam, Rachel
et al.

Publication Date
2011

Peer reviewed

Learning by observing tutorial dialogue versus monologue collaboratively or alone

Kasia Muldner, Keith Dybvig, Rachel Lam & Michelene T. H. Chi
(katarzyna.muldner,kdybvig,rachel.lam,mtchi@asu.edu)
Department of Psychology, Arizona State University Tempe, AZ 85287 USA

Abstract

We report on a study with 65 middle-school students who learned about the concept of *diffusion* through observation. We manipulated two factors: the number of observers, solo vs. dyad, and the type of video students observed, tutorial dialogue vs. monologue. Our findings show that dyad observers learn significantly better than solo observers, and that for certain types of questions, observing dialogue results in better learning gains, as compared to observing monologue.

Keywords: vicarious learning; collaboration; monologue vs. dialogue; emergent phenomena.

Introduction

Although there are many contexts in which learning can occur, traditionally, students learn by watching and listening to teachers and by doing various assignments. A less available, but more beneficial way to learn is through one-on-one tutoring, which has repeatedly been shown to be superior to group instruction (Cohen, Kulik & Kulik, 1982). Of course providing a tutor for every student is not feasible and so there have been various efforts to find other strategies that afford students the benefits of personalized tutoring. Recently, a promising new paradigm was proposed: learning by observing others learn, also referred to as *vicarious learning* (Chi, Roy & Hausmann, 2008).

In particular, Chi et al. (2008) showed that observing tutoring dialogues can be as powerful an intervention as being tutored. In that study, some students (referred to as *tutees*) were videotaped solving Newtonian physics problems while interacting with an expert tutor. Other students, in pairs or alone, then merely watched the videos of these tutoring sessions while solving the same physics problems as the tutees. One of the key findings was that there was no difference in learning outcomes between the tutees and the observers who watched the tutorial dialogue videos. As predicted, this finding was only true if pairs of students observed the videos collaboratively (students who observed alone did not learn as well as the tutees). Since dyad observers have opportunities to be interactive with each other, much as the tutees have opportunities to be interactive with the tutor, the interpretation made was that a large part of the tutoring advantage is due to interaction per se.

The benefits of collaboration are demonstrated by various studies (e.g., (Johnson & Johnson, 2009)), findings from which are encapsulated by the *active-constructive-interactive* framework (Chi, 2009). This framework differentiates learning activities according to levels of student engagement, and proposes that a student who collaborates with a peer will, in general, learn better than a student working alone. This is because collaboration offers the opportunity for joint construction, which requires, for instance, eliciting responses from a partner, integrating a partner's contribution, and explaining one's perspective. It is important to note, however, that the prediction regarding the beneficial impact of collaboration is based on the condition that students are interactive, i.e., do not merely sit quietly, each working on his or her own. In fact, some prior work on vicarious learning did not find that dyad observers learned better than solo ones (Craig, Driscoll & Gholson, 2004), because the observers were not very interactive.

The goal of this project is to both replicate and extend the original Chi et al. (2008) study. In replicating, we forgo comparing tutees to observers, and focus instead on comparing observing collaboratively versus observing alone, but in a new domain and with a new population. Specifically, while the original study used a procedural domain, we embed our target learning activity in a conceptual domain, and involve middle school instead of college students. Replicating in this new domain and age group will generalize and validate the Chi et al. (2008) findings.

In extending the 2008 study, we compare the effectiveness of dialogue versus monologue. Doing so can guide subsequent efforts on instructional material development. Currently, materials are often monologue-based, e.g., instructional videos with a "talking head" (Caspi, Gorsky & Privman, 2005; Zhang, Zhou, Briggs et al., 2006). However, prior work comparing observation of monologue and dialogue videos does provide some clues that dialogue may be better (Craig, Chi & VanLehn, 2009; Driscoll, Craig, Gholson et al., 2003; Fox Tree, 1999; Muller, Sharma, Eklund et al., 2007). For example, Craig et al. (2009) found some evidence that naturally-occurring tutorial dialogue fosters better learning. In that study, students observed a dialogue or a monologue video while working with an intelligent tutor that provided both feedback for correctness and hints. Thus, it is important to analyze if and how the benefits of naturally-occurring dialogue transfer to situations where students do not receive such additional scaffolding. There are, however, studies that did not find a difference between observing dialogue and monologue media (Fox Tree & Mayer, 2008; Muller, Bewes, Sharma et al., 2008), although some of these focused on simple puzzle tasks (Fox Tree & Mayer, 2008). In general, more work is needed for understanding the effect of each (monologue, dialogue).

The predominant approach for comparing dialogue to monologue has been to script the content, while varying a

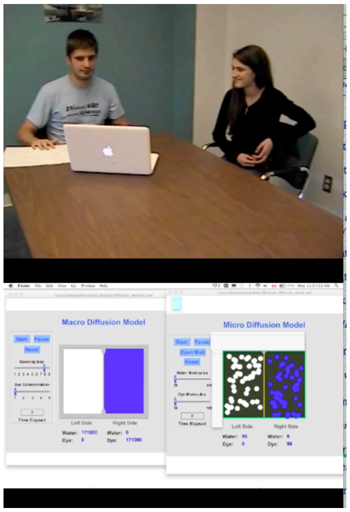

Figure 1: Dialogue video with tutor + tutee (top) and corresponding simulations (bottom)

factor of interest (e.g., (Craig et al., 2004; Craig, Sullins, Witherspoon et al., 2006; Driscoll et al., 2003; Muller et al., 2008; Muller et al., 2007)). For instance, the scripting approach was used to study the impact of including misconceptions (Muller et al., 2007) or questions that precede certain utterances (Craig et al., 2006). With the current very preliminary understanding of the benefits of tutorial dialogues, scripting may be premature since it may miss important nuances that occur in naturally-driven interactions. For this reason, we have chosen to use naturally-occurring tutorial dialogue and monologue.

The domain: emergent processes & diffusion

Our target domain corresponds to a conceptual science topic – *diffusion*. This is a highly misconceived and challenging topic for students (Chi, Roscoe, Slotta et al., in press; Meir, Perry, Stal et al., 2005), because it requires understanding of two very difficult concepts: *emergent processes* and *proportionality*. Emergent processes are defined through the attributes and features that characterize these processes (Chi, 2005). To illustrate, suppose a drop of dye is dropped into water. This diffusion of dye throughout the water is an emergent process because:

- (*disjoint* attribute) the dye and water molecules and/or their interactions and the visible flow pattern of the dye can behave in disjoint ways;
- (*collective* attribute) the flow pattern is caused by the collective summing of all the molecular interactions;
- (*random* feature) the molecular interactions are random.

There are a total of 10 emergent attributes and features - for a full list, see (Chi et al., in press). The two attributes listed above are classified as inter-level because they require students to reason about both the visible macro-level pattern (flow of dye) and the underlying micro-level interactions (movement of molecules). Doing so is very challenging and so students hold many misconceptions about diffusion, such as that the molecules stop moving when the solution appears a uniform unchanging color (Meir et al., 2005).

The *collective* attribute requires understanding of ratio and proportion. For instance, in the context of the dye example, the overall changes in concentration of dye relative to water cause the visible dye flow pattern. Thus, students need to understand proportion-related concepts to fully comprehend diffusion. There are numerous studies showing that these notions are very difficult for students (e.g., (Smith, Carey & Wiser, 1985)), further adding to the complexity of learning about diffusion.

Study details

Materials. The study involved the following materials related to diffusion: (1) a two-page text, (2) pre- and post-tests, (3) two simulations and (4) two instructional videos. Materials 1-3 were based on ones used in an earlier study (Chi et al., in press). The text was designed to provide the necessary foundations for diffusion-related concepts. The pre- and post-tests assessed students' diffusion knowledge, and did include some questions that probed understanding of emergent aspects of diffusion, but without explicitly mentioning emergence. For instance, to assess the inter-level *disjoint* attribute, one question asked "*As the dye diffuses away from where it was originally dropped into the water, can some dye molecules bounce back towards this original place?*" The pre-test included 19 multiple-choice questions, while the post-test included the same 19 questions and six extra questions for a total of 25 questions (the six extra questions were added to avoid the retest effect, i.e., increased learning due to identical pre and post tests).

To help students understand inter-level concepts, the simulations showed diffusion occurring on the visible level (*macro* simulation) and at the molecular level (*micro* simulation; see the left and right panel of the lower half of Figure 1 for the macro and micro simulation, respectively). The simulations were interactive (for instance, clicking the "start" button in the micro simulation resulted in molecules bouncing and colliding), and were used in both the instructional videos and by the observers (as described below).

Two instructional videos were created in our lab: a dialogue (tutor + tutee) and a monologue (tutor only). The tutor, used for both videos, had extensive tutoring experience and received domain training so that he was very familiar with the target concepts. An eighth-grade student was chosen to be the tutee in the dialogue video, based on the guidelines that (1) observers learn better when the tutee they are observing does not have ideal knowledge and so generates some errors (Schunk, Hanson & Cox, 1987), and (2) the tutee has *some* knowledge and so is able to answer a subset of the tutor's questions (Chi et al., 2008). This was the case with our tutee, who obtained a pre-test score of 61%. The tutee first read the diffusion text, took the pre-test and then discussed diffusion with the tutor for about 20 minutes.

During both the dialogue and monologue session, the tutor was asked to cover the key topic areas outlined before hand, including concentration and the 10 emergent features and attributes operationalized within the topic of diffusion. Both sessions (dialogue, monologue) involved the two simulations, which were used to illustrate the topics discussed and were shown on a laptop available to the tutor (and tutee for the dialogue). All laptop actions were recorded using screen capture software and the sessions were video taped. The final videos were a "split screen", where the tutor and tutee (dialogue) or tutor (monologue) were shown on the top portion of the screen, and the simulations and the users' actions in them were simultaneously shown on the bottom portion of the screen (see Figure 1).

In neither the dialogue or monologue session did the tutor adhere to a pre-defined script, because we wanted to keep the sessions as natural as possible. Moreover, scripting has the potential to miss key events since as mentioned above, it is not yet clear which dialogue or monologue features are needed for optimal observer learning. We did, however, aim to standardize a number of factors between the two videos. First, both were comparable in length (22:56 minutes and 21:10 minutes). Second, the tutor was instructed to cover the same concepts in both sessions.

Participants. The participants were local middle-school eighth-grade students, who engaged in the study as part of their standard classroom activities.

Design. The study included two independent variables: *video-type* (monologue, dialogue) and *number-observers* (solo, dyad); thus, there were four conditions. The participants came from four different classes. To avoid any class effects, students within each class were randomly assigned to evenly fill the four conditions (i.e., a given condition had subjects from all four classes).

Procedure. The study took place in the school. Students had not been taught about diffusion prior to the study. For each of the four classes participating, two 60 minute class periods were used on two consecutive school days (all classes were done with the study within three school days). On the first day, students were introduced to the research process (5 min.) signed the assent forms (10 min.), read the diffusion text (15 min.), and took the pre-test (15 min.). Students were also introduced to subsequent activities: a researcher used a smart board to explain the protocol for the next day (10 min.). On the second day, students watched an instructional video in a computer lab (*experimental intervention*, ~30 min.), and took the post-test (20 min.).

During the experimental intervention, each student (solo condition) or pair of students (dyad condition) used a computer to (1) watch an instructional video (the dialogue in Figure 1 or a monologue, not shown) and (2) interact with the micro and macro simulations. Thus, students could both see how the tutor (and tutee) interacted with the simulation in the video and could also interact with their own simulations. A subset of the dyad subjects were audio recorded and these recordings were transcribed.

Table 1: Subject information: pre-test % (# subjects) N=65

	dyad	solo
monologue	*50.4%* (N=19)	*47.4%* (N=16)
dialogue	*45.6%* (N=15)	*59.6%* (N=15)

Results

The results are based on the 65 students who completed all phases of the study and who each provided a parental consent and student assent form. Although subjects were randomly assigned to the study conditions, there were slight differences between the groups (see Table 1 - the difference in group sizes is the result of student absences and/or lack of consent). In such a situation, the appropriate analysis to use is an ANOVA with specially-adjusted gain scores, advocated in (Crouch & Mazur, 2001):

$$\text{adjusted-gain score} = \begin{cases} 0 \text{ if } \textit{pre-test \%} \text{ equals } 100\%, \text{ otherwise} \\ (\textit{post-test \%} - \textit{pre-test \%}) / (100 - \textit{pre-test \%}) \end{cases}$$

Thus, a student's gain score is adjusted according to the pre-test so that students who start out with a high pre-test score obtain an adjusted, higher gain than students who start with a lower pre-test score (e.g., a student who moves from 20% to 60% between pre and post test is assigned the same gain as a student who shifts from 80% to 90%). The rationale behind this adjustment is that it is more difficult to improve given a high pre-test score, as compared to a low pre-test score. In our study, the pre-test average for all the groups was below 60%, so students were not at ceiling.

A complimentary analysis to use is an ANCOVA with the *pre-test %* as the covariate and *post-test %* as the dependent variable. This analysis adjusts the post-test score through the covariate, thereby accounting for any differences in pre-test scores. While we conducted both types of analyses, the ANCOVA confirmed the adjusted-gain ANOVA, and so for brevity, we only report results from the former.

Analysis 1: effect of *number-observers* and *video-type* on adjusted-gain score

An ANOVA was conducted with the adjusted-gain score considering all pre- and post-test questions (19 and 25, respectively). As mentioned above, the post-test included an additional six questions to avoid the re-test effect. The post-test scores on these questions were very similar to the 19 matched post-test scores (62% vs. 59%); thus these questions were not significantly easier or harder. The mean adjusted-gain score for each group is shown in Figure 2. Overall, dyad observers learned significantly better than solo observers ($F(1,61)=5.9$, $p=0.018$; mean adjusted-gain score for dyad vs. solo observers: 21.0% vs. 5.5%; Cohen's effect is medium to large: $d=.6$). Moreover, we replicated the earlier result showing that dyad observers perform better than solo observers when given dialogue (Chi et al., 2008) (see Figure 2, $t(28)=28$, $p=0.02$; large effect, Cohen's $d=.9$).

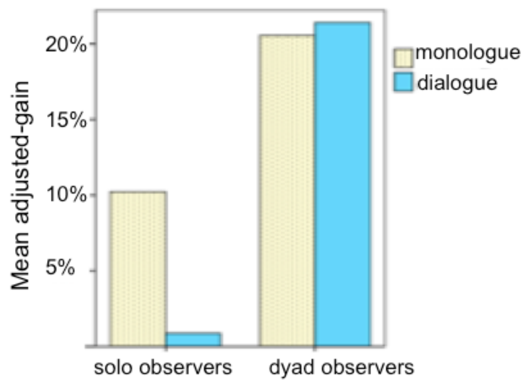

Figure 2: Adjusted-gain scores for each group (N=65)

The *active-constructive-interactive* framework (Chi, 2009) predicts that these findings are due to students being interactive and thus constructive.

In addition to examining the effect of *number-observers*, our second goal was to explore how middle-school students learn from dialogue and monologue. The ANOVA showed that *video-type* did not have a significant effect on adjusted-gain score (p=.5). This finding was somewhat unexpected, given that as we mentioned above, there is some evidence of dialogue being superior to monologue. There are a number of possibilities for why we did not find such an effect here. One is that the post-test contains some extremely difficult transfer questions assessing misconceptions related to proportionality and/or inter-level concepts. Accordingly, the next analysis addressed this issue.

Analysis 2: effect of *number-observers* and *video-type* on different types of post-test questions

To determine if the difficulty of some test questions was obscuring the results, we divided the questions into *explicit* and *implicit* ones. The explicit subset corresponded to post-test questions that the instructional materials addressed explicitly (there were nine such questions and six corresponding pre-test questions). To illustrate, consider the post-test question "*After the clamp is removed, the dye appears to flow from Beaker #1 to Beaker #2. Can a dye molecule that is now in Beaker #2 move backwards into Beaker #1?*". The fact that molecules can go "back" was explicitly addressed in both the monologue and dialogue videos, as illustrated below from an excerpt from the dialogue video:

T: which way is the dye overall going to end up moving?
S: to the left side [some discussion left out for brevity]
T: And what are they [molecules] going to sometimes do?
S: They'll bounce off and…
T: go to the …
S: other side
T: other side right. And once they go to the other side can they ever come back?
S: yes
T: yeah – they can come back.

Note that the explicit questions may still require students to abstract some details and so they certainly were not trivial.

The implicit subset included post-test questions requiring the observers to make inference(s) much beyond that what was stated in the videos (there were 15 such post-test questions and 12 corresponding pre-test questions). For 11 of the post-test questions, that inference involved switching contexts. For instance, one of the post-test questions asked students to select a choice that characterized the reason for oxygen and carbon molecules moving across a cell membrane. The correct choice reflected the fact that the cell does not pull in the beneficial oxygen and push out the harmful carbon dioxide, but rather that the corresponding molecules move randomly across the cell membrane. While this context was never mentioned explicitly in the text or the videos, these materials did describe random molecular movement, and so a student could answer this and other implicit questions by generating additional inferences. In a sense, 11 of the implicit test questions were transfer questions. The other four questions involved a context similar to one in the instructional materials, probing inter-level and proportionality concepts.

To label a question as explicit or implicit, two researchers coded the questions with respect to how the instructional materials addressed the two types of questions. Agreement was substantial (Kappa = .73); disagreements were resolved through discussion. There were no differences between the monologue and dialogue videos with respect to how they addressed the two types of questions except for one post-test question. This question was excluded from the subsequent analysis. Of the remaining 24 post-test questions, all but two questions were explicitly or implicitly addressed by the videos (two explicit post-test questions were addressed only in the diffusion text). The subsequent analysis computes the adjusted-gain score (see formula above) for the explicit and the implicit questions.

Although our primary goal with this current analysis was to determine whether a *video-type* effect exists for certain questions, we also included the *number-observers* factor in the ANOVA to check if the above finding that dyads perform better than solo applies to each question subset, and to account for any interaction effects.

For the implicit questions, the ANOVA did not find significant effects of *video-type* or *number-observers* on adjusted-gain score. This lack of significance is likely due to the fact that the implicit questions all required many additional inferences and/or transfer, something that is notoriously difficult for students to achieve. The difficulty of these questions is confirmed by the corresponding low adjusted-gain scores, as compared to the explicit questions (mean adjusted-gain score: implicit=2%, explicit=31%). Thus, the subsequent discussion is focused on the explicit questions.

Results from analyzing adjusted-gain score for the explicit questions. For the explicit questions, we first replicated the above finding that dyad observers gain more

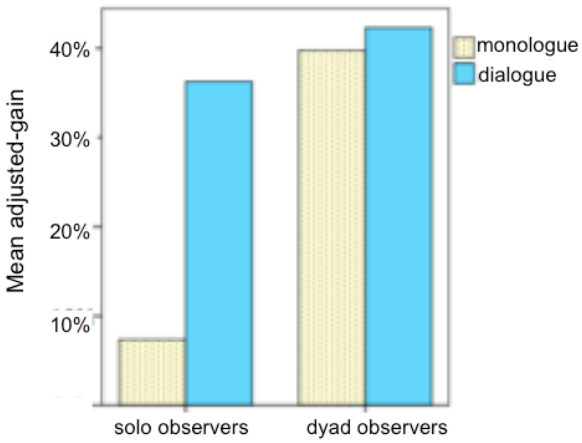

Figure 3: Adjusted-gain scores for each group for the explicit questions only (N=65)

than solo observers ((F(1,61)=6.8, p=0.01; mean adjusted-gain score for dyad and solo observers: 41% vs. 21.8%, respectively; medium to large effect: d=.6).

Moreover, *video-type* also had a significant effect on adjusted-gain score (F(1,61)=4.6, p=0.036, medium effect: d=.5), with observers given a dialogue video performing better than those given a monologue video (mean adjusted-gain score for dialogue and monologue: 39.3% vs. 23.6%, respectively, individual group means shown in Figure 3). Thus, in contrast to the analysis above that included implicit and explicit questions and did not find a difference between dialogue and monologue, when the implicit questions were excluded because students were unable to answer them, dialogue was superior to monologue.

Some conjectures for why dialogue fostered larger learning gains

In this section, we contemplate why observing tutorial dialogue might trigger more learning than observing monologue. Since various studies show that interest improves cognitive functioning and performance (Lepper, 1988), one possibility relates to motivation. Anecdotally, our transcripts show that students were more interested in observing tutorial dialogue (e.g., one student expressed this by stating "*single person is way more boring"*). This effect should be particularly true for students in the solo condition, since they did not experience the potentially motivating effect of having a partner to interact with. Students in the solo condition did indeed gain significantly more on the explicit questions when given a dialogue video, as compared to monologue (t(29)=2.7, p=0.01; large effect: d=1.0). In contrast, while dyad observers did gain more when given the dialogue video, as compared to monologue, this difference was not significant, possibly because the effect of collaboration overpowered any *video-type* effects.

Compared to monologue, observing dialogue may also be more motivating because the tutee provides a so-called coping model (Schunk et al., 1987). In particular, observers do not yet have perfect domain knowledge, and so can relate to the tutee struggling or "coping" in the video. Such a model is clearly not present in a monologue since the tutor does not express doubt and does not commit errors. Schunk et al. (1987) did indeed find that observers' self-efficacy, which is related to motivation, improved more after viewing videos in which tutees struggled while solving problems, as compared to videos in which tutees did not.

Related to the last point, in contrast to a monologue, in a dialogue tutees may express misconceptions that are shared by observers. This is important not only because that misconception is subsequently refuted by the tutor, but also because the tutee's perspective is more likely to be aligned to an observer's than a tutor's, and so the misconception is expressed in a manner to which an observer can relate. When we analyzed gains on individual explicit test questions, one question in particular stood out ("*Why does the dye spread from where it was originally added to the water?*"). For this question, the average gain was twice as high for the dialogue observers, compared to ones given monologue (on average, 0.47 vs. 0.26 gain out of a total possible gain of 1). When we checked the instructional videos related to this question, it corresponded to the tutee expressing an incorrect explanation in response to a tutor question:

T: Does anything force them … all the molecules… or some of the molecules to go to the other side?
S: I think… yes…
T: What forces them to go to the other side?
S: The pressure from (shrugs)…

In fact, this was the only instance in the video where the tutee explicitly expressed a misconception (i.e., that pressure makes the molecules to go to the other side). There were other places in the video that confusion occurred, or a simple *"not sure"*, but here the tutee expressed an alternative viewpoint through her misconception. Prior work on scripted dialogue has shown no difference in terms of learning between monologue and dialogue if both contain misconceptions (Muller et al., 2008). In contrast, although our monologue did include misconceptions expressed by the tutor (e.g., "*you might think that…"*), we did find a difference in learning gains between monologue and dialogue. However, our dialogue was naturally occurring, so the embedded misconceptions were expressed from the tutee's perspective. This may explain the difference between our findings and prior work, and highlights the importance of studying the effects of video-type with naturally occurring content, at least until we have a better understanding of the impact of the various factors.

Discussion and future work

Our analysis showed a consistent effect of collaboration, with dyad observers learning more than the solo observers. Although this effect is predicted by other work (e.g., (Chi, 2009)), in our study it was not a given that dyads would in fact be more constructive than the solo observers, since little scaffolding was provided for the overall processes. For instance, a form of scaffolding could correspond to giving students a worksheet to fill in as they watch the videos, as in (Chi et al., 2008). This type of activity encourages collaboration, since students have a specific task to work on. In our study, observers did not have any specific tasks, but did have the simulations, and this may have provided the catalyst for interaction. Students did indeed discuss the simulations: when we analyzed the transcripts for simulation-related utterances, on average students referred to the simulation 16.9 times per session (and 13.1 of those references were clearly to the simulations students could manipulate, as opposed to the ones they could merely observe in the video).

As far as the effect of dialogue versus monologue is concerned, observing dialogue was better than observing monologue for subsequently answering the explicit test questions. This was the first study to show this effect for middle school students. We did not find that dialogue was better than monologue for answering implicit questions, likely because questions such as these require transfer, something that is notoriously difficult to achieve. We provided some interpretations and found some clues for why dialogue was better than monologue, related to motivation and presence of misconceptions. In the future we plan to analyze these factors in more detail, as well as enlist more tutees to compare being tutored against observing tutoring.

Acknowledgements

The work was funded under the Spencer Foundation Grant Program, award number 200800196. The authors thank Johnathan Scheibenpflug and Brittany Gianiorio for helping prepare some of the study materials and the anonymous reviewers for their helpful comments.

References

Caspi, A., Gorsky, P. & Privman, M. (2005). Viewing comprehension: Students' learning preferences and strategies when studying from video. *Instructional Science*, *33*, 31-47.

Chi, M. T. H. (2005). Common sense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, *14(2)*, 161-199.

Chi, M. T. H. (2009). Active-contructive-interactive: a conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73-105.

Chi, M. T. H., Roscoe, R., Slotta, J., Roy, M. & Chase, M. (in press). Misconceived Causal Explanations for Emergent Processes. *Cognitive Science*.

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

Cohen, P. A., Kulik, J. A. & Kulik, C.-L. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, *19*, 237-248.

Craig, S., Driscoll, D. & Gholson, B. (2004). Constructing Knowledge from Dialog in an Intelligent Tutoring System: Interactive Learning, Vicarious Learning, and Pedagogical Agents. *Journal of Educational Multimedia and Hypermedia*, *13(2)*, 163-183.

Craig, S., Sullins, J., Witherspoon, A. & Gloslon, B. (2006). The Deep-Level-Reasoning-Question Effect: The Role of Dialogue and Deep-Level-Reasoning Questions During Vicarious Learning. *Cognition and Instruction*, *24(4)*, 565-591.

Craig, S. D., Chi, M. T. H. & VanLehn, K. (2009). Improving Classroom Learning by Collaboratively Observing Human Tutoring Videos While Problem Solving. *Journal of Educational Psychology*, *101(4)*, 779-789.

Crouch, C. H. & Mazur, E. (2001). Peer Instruction: Ten Years of Experience and Results. *American Journal of Physics 69*, 970-977.

Driscoll, D. M., Craig, S. D., Gholson, B., Hu, X. & Graesser, A. (2003). Vicarious Learning: Effects of Overhearing Dialog and Monologue-like Discourse in a virtual tutoring session. *Journal of Computing Research*, *29(4)*, 431-450.

Fox Tree, J. (1999). Listening in on monologues and dialogues. *Discourse Processes*, *27*, 35-53.

Fox Tree, J. & Mayer, S. A. (2008). Overhearing single and multiple perspectives. *Discourse Processes*, *45*, 160-179.

Johnson, D. W. & Johnson, R. T. (2009). An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. *Educational Researcher*, *28(5)*, 365–379.

Lepper, M. (1988). Motivational Considerations in the study of Instruction. *Cognition and Instruction*, *5(4)*, 289-309.

Meir, E., Perry, J., Stal, D. & Klopfer, E. (2005). How Effective Are Simulated Molecular-level Experiments for Teaching Diffusion and Osmosis? *Cell Biology Education*, *4*, 255-248.

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*, 144-155.

Muller, D. A., Sharma, M. D., Eklund, J. & Reimann, P. (2007). Conceptual change through vicarious learning in an authentic physics setting. *Instructional Science*, *35*, 519-533.

Schunk, D. H., Hanson, R. A. & Cox, P. D. (1987). Peer-Model Attributes and Children's Achievement. *Journal of Educational Psychology*, *79(1)*, 54-61.

Smith, C., Carey, S. & Wiser, M. (1985). On differentiation: A case study of the development of the concept of size, weight, and density. *Cognition*, *21*, 177–237.

Zhang, D., Zhou, L., Briggs, R. O., Nunamaker, J. F. & (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, *43*, 15-27.