Analyzing Patterns of Emerging Understanding and Misunderstanding in Collaborative Science Learning: A Method for Unpacking Critical Turning Points

Alisha Portolese, Lina Markauskaite, Polly K. Lai, and Michael J. Jacobson
Alisha.Portolese@sydney.edu.au, Lina.Markauskaite@sydney.edu.au, Kuan.Lai@sydney.edu.au, Michael.Jacobson@sydney.edu.au
The University of Sydney

Abstract: When students learn science in a computer-supported, collaborative, delayed-instruction environment, how does understanding (and misunderstanding) emerge? Are there patterns in the pivotal moments when emerging understanding turns for the better or worse? While components such as modeling software, delayed instruction methods such as productive failure, and analogical-encoding methods such as contrasting cases have all been shown effective at supporting deep learning in science, little is known about the micro-level mechanisms explaining how and why students might be more or less successful when working in an environment combining all three. This paper details our refinements of an innovative method for unpacking the micro-level mechanisms contributing to turning points in the successes and failures in collaborative understanding when learning science with computer modeling. In unpacking our methodology, we discuss work including Sanderson and Fisher's (1994) exploratory sequential data analysis (ESDA) guidelines and the productive multivocality project (Suthers, 2013) to frame our approach.

Keywords: conceptual change, collaborative learning, science education, turning points, pivotal moments, delayed instruction

Introduction

When learning science, *experiencing* science and engaging in collaborative, authentic, and active learning can be highly beneficial for deep learning (Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001). Providing a space for domain-grounded discussions is particularly beneficial for students to activate prior knowledge, become aware of existing beliefs, realize the extent of their current understanding, and discuss emerging understanding (Greeno, Collins, & Resnick, 1992; Roschelle, 1992; Vosniadou et al., 2001). Additionally, learning environments are strong when they include problems that resemble the complexity inherent in authentic, real-world situations (Jacobson, 2000; Kapur & Kinzer, 2007). Environments with this type of authentic depth foster the rich discussions that support the development of deep learning and transfer.

One method of sculpting a learning environment with these features involves incorporating computer modeling. Models provide the opportunity to manipulate and interpret complex relationships as they happen, and can serve as a visual tool to ground abstract concepts (Nersessian, 2008). Students can manipulate a shared computer model in small groups (collaboration), which can be situated in a delayed instruction learning sequence. For example, in productive failure, students begin by collaboratively exploring possible solutions to a complex, authentic problem they have not yet been taught how to solve. This is followed by a teacher-led consolidation highlighting the critical solution features and comparing and contrasting student and expert solutions (Kapur & Bielaczyc, 2012). Collaborative problem solving with computer modeling software and contrasting cases can be effectively included as part of productive failure's initial idea generation and exploration phase (Jacobson & Markauskaite, 2015, April; Portolese, Markauskaite, Lai, & Jacobson, 2015, April). The effectiveness of contrasting cases is grounded in analogical encoding theory (Gentner, Loewenstein, & Thompson, 2003), which proposes that explicit comparison of multiple cases with different surface features but similar underlying principles can enhance learning the critical features of the core concept.

Despite strong evidence for promoting deep learning and transfer with modeling, delayed instruction, and contrasting cases, the mechanisms underlying how and why these processes work, particularly with all combined, remains relatively unknown. One thread of our research has examined some important questions – when learning in this way, how does collaborative understanding and misunderstanding emerge? And how specifically does understanding take a turn for the better or worse? From a sociocultural perspective, knowledge is co-constructed through interactions (Säljö, 1991). Knowledge building interactions are productive when they allow students to build partial meanings that are gradually refined towards increasingly expert understanding

(Damşa, 2014). Therefore, the micro-level interactional mechanisms fueling emerging knowledge co-construction are a critical yet little understood aspect of deep, collaborative learning. Few studies have explicitly addressed how emerging knowledge relates to interaction over time (Kapur & Kinzer, 2007; Damşa, 2014).

This paper aims to respond by contributing to theory and methodology by providing a detailed, multi-layered, temporally-sensitive approach for analyses designed to unveil the micro-level mechanisms of conceptual change and emerging understanding in complex, rich, computer-supported collaborative learning environments. In preliminary analyses, we explored students' collaborative learning at three parallel grain sizes using an impact coding approach (Portolese et al., 2015, April). With this, we found that superior performance on a far transfer item may have been associated with an idea generation process characterized by producing substantially more ideas, particularly more correct suggestions, predictions, experimental questions, experimental designs, and explanations. We also found that students' collaborative processes seemed to be characterized by small segments of misunderstanding propelling extended correct understanding. However, it remained to be explored what exactly occurred at these critical turning points in understanding; a deeper investigation was required. As we will present in this paper, we expanded our analysis technique to identify and explore turning points in understanding in a rich and meaningful way. We briefly presented our early ideas for how these turning points could be unpacked in a poster at The Computer Supported Collaborative Learning Conference (Portolese, Markauskaite, Lai, & Jacobson, 2015, June). With the support of peer feedback, we present here a refined expansion of this methodology for unpacking turning points in substantially more detail. This detailed explanation of our exploration and conceptualization of turning points is an important contribution to understanding how the mechanisms of emerging understanding and misunderstanding in model-based collaborative learning might be revealed.

In line with the conference's thematic strands, our aim is to unpack the micro-level mechanisms underlying conceptual change and knowledge construction in a science learning, computer-mediated collaborative environment. We believe our main contribution is the detail of an innovative method for unpacking the micro-level mechanisms of turning points in the successes and failures in collaborative understanding when learning science with modeling software. In addition, we present the patterns of emerging understanding and misunderstanding from our data, including insights regarding which aspects might be more and less "productive" to include or withhold scaffolding. While there has been much recent activity regarding such issues (see Kirschner, Sweller, & Clark, 2006; Kapur & Bielaczyc, 2012; Jacobson, Kim, Pathak, & Zhang, 2013) at the larger learning design level, much work is still needed regarding unpacking the mechanisms within powerful learning designs (see Loibl & Rummel, 2014, for a productive move in this direction).

Methods

Context and participants

The two dyads chosen for detailed analyses were selected from a larger study conducted across four Year Nine science classes at a selective girls high school in Australia (see Jacobson & Markauskaite, 2015, April; Portolese et al., 2015, April). Dyads worked collaboratively on inquiry activities that required experimentation with NetLogo (Wilensky, 1999) models. The students from the two selected dyads improved substantially on the target complexity concepts from pretest to posttest within their own groups, in line with the overall results for each group (see for details Jacobson & Markauskaite, 2015, April; Portolese et al., 2015, April).

Details of our coding approach, Phase 1: Impact coding at three-parallel grain sizes

The first phase of our analysis involved exploratory coding at three parallel grain sizes. The approach was developed from our understanding of the data (not a pre-defined model). The incorporation of impact coding to allow for context and time-sensitivity was based on Kapur, Voiklis, and Kinzer's (2008) method; see Portolese et al. (2015, April). As Kapur et al. (2008) argue, this method preserves the temporal sensitivity that is critical in understanding emerging understanding (or misunderstanding). With this, data segments were coded as +1 (moving towards the solution), -1 (moving away from the solution), or 0 (not changing progress). In addition to numeric codes, we also used descriptive labels to add richness to our interpretation.

The micro grain size was the idea level. We defined ideas as a single train of thought (one or multiple speakers) and associated computer actions (e.g. model manipulations). In addition to impact coding, ten categorical labels emerged: the first three about process and the remaining seven about content (see Table 1, and see Portolese et al., 2015, April, for additional details and examples). The content categories that emerged could be understood as cycles of the scientific method (see Figure 1). In summary, each micro segment was assigned a category label and numeric code.

Table 1: Summary of Micro Level idea category descriptions

Idea category	**Description**
Task	Orienting to the task, understanding instructions
Technical	Technical issues often related to the NetLogo modeling software
Representation	What an element of the model represents, and/or how various parts of the model are related to each other (e.g. simulation interface vs. graphical representations)
Suggestion	Novel contributions or solution directions
Experimental Question	Inquiries related to or leading to student experiments
Prediction	Guesses for the outcome of a current or upcoming experiment
Experimental Design	Planning and execution of a modeling experiment
Observation & Data Collection	What the students see in the model (as made overt)
Explanation	How students interpret their observations and experiments
Understanding	Indication of broader comprehension (or miscomprehension) of the target concept

Suggestion ➔ Experimental Question ➔ Prediction ➔ Experimental Design ➔ Observation & Data Collection ➔ Explanation ➔ Understanding ➔ (Suggestion; cycle continues)

Figure 1. The seven content based idea-level categories can be understood as the cycle of the scientific method.

The meso grain size was about change in understanding the epistemic task. Generally, one or multiple idea-level moves indicated this progression of a new phase of understanding. For example, if a student made a prediction and ran an experiment with particular parameters, this could represent their preconception of the concept at hand. Then, during the experiment, students might make observations and explanations, and at this point students might *change* their understanding due to the events they are observing. As students discuss and elaborate ideas, understanding might *change again*. This example broadly outlines a progression of three meso-level segments. In addition to the impact coding, a non-categorical description was provided (see Table 2).

The macro grain size was the experimental level. Students were given guiding questions to softly scaffold their interaction with the models. Working through the questions, students manipulated the model, running simulation experiments. We segmented each macro grain size between when students planned an experiment to when students completed related conclusions. A prototype would include idea segments representing an entire cycle of the process in Figure 1. However, since this was a minimally guided activity with novices, students often did not follow this prototypical method – sometimes they made ad hoc choices with the modeling parameters, or did not make overt observations or discuss their inferences. In these instances, we created segment boundaries based on indicator activities such as manipulating model parameters and refreshing the model, as these actions demonstrate the intention of creating a new experiment. As with the meso level, both a numeric impact code and a non-categorical description was associated with each segment. See Portolese et al. (2015, April) for examples of coding at this level. Importantly, the impact code was assigned based on the position students demonstrated they were in at the end of the segment – as macro segments could contain diverse meso segments within it, understanding could be turbulent and changing throughout the segment.

The three grain sizes were coded somewhat in parallel – the broader context of students' activity regarding their experiments helped make sense of their actions and words at the micro and meso levels subsumed within in (see Figure 2). With the broader context of a macro segment generally understood, the grain sizes were then coded from the micro level segments until reaching the end of a meso segment, and then moving onto the following meso chunk and starting again with the micro segments within it. Labels and descriptions were usually coded before the numeric impact code. Multiple time-aligned factors in our rich data were taken into account, including computer actions, words, and written workbook answers (integrated as they occurred in time). Students' words were considered regarding tone of voice (e.g. sarcasm), implications (utilizing context to infer likely meaning when possible), and focus of attention (e.g. cues from eyes and classroom events). Other relevant events such as interactions with nearby groups or with researchers and teachers were also considered and coded. Overall, a simple yet effective way to help determine how to code a segment at any level was asking ourselves, "At the end of this segment, has the students' understanding changed? If so, in what way?"

Table 2: Meso Level Coding

Impact Level	Definition	Example	Example Description
-1	New or reinforced misunderstand-ing	*"Okay, maybe make it equal" • Computer: wolf-reproduce #4 •"Make this four" • Computer: wolf-gain from food #4 • "But we're gonna keep that number the same, ok? • "Mhmm" • Computer: Go (start) (no response) • "Oh wait, no. Set up" • Computer: Set up, Go (start)*	The students incorrectly thought that manipulating the parameters "wolf gain from food" and "wolf reproduce" so that the values were equal would make the model self sustaining.
0	Do not understand what is happening or no overt change in understanding	*Computer: Go (start) • "Wow okay" • "Gosh" • "Where's the discussion?" • "Wait – where are the wolves?" • "Oh wow, oh gosh" • "Woah" • "Oh wow. That, that's horrible. Stop. How do I stop it?" • "Woah" • "What just happened?" • "I don't know, but the population of sheep is just going up."*	At the beginning of their work, the students do not yet understand what is happening with the model.
+1	New or reinforced understanding	*Computer: Go (stop) • "How do you stop it?" • Computer: Go (start) • "Oh there we go... okay yeah it's still going" • Computer Go (stop) • "ok, um"*	The students learn how to stop the model.

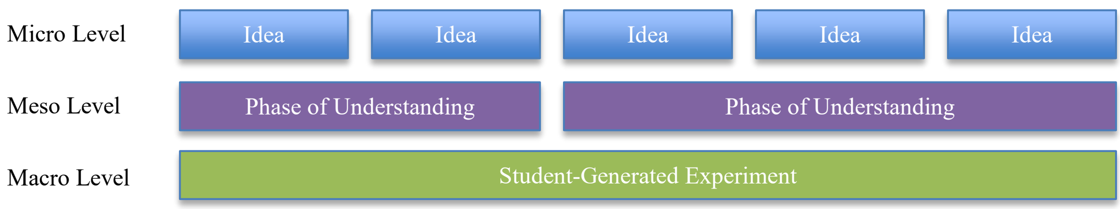

Figure 2. How the three parallel grain sizes fit together.

Details of our coding approach, Phase 2: Turning points analysis

As discussed above, we found that the initial phase of analysis provided useful insights, however, we felt the need to go deeper to truly understand why understanding changed, and what (if any) patterns existed in what we were observing. The impact coding afforded us the possibility to graph a group's cumulative, unfolding understanding and misunderstanding. There were many possible ways this could be done for each group – at each of the grain sizes, and within the micro category using a filter for one or some categories. In our meso level, which we determined focused at the core of what was happening and changing with understanding, we were particularly interested to delve deeper to understand points where understanding seemed to turn. We defined turning points as critical moments in the development of collaborative understanding when understanding changed in an incremental way. The numeric impact coding allowed for an opportunity to provide a clear boundary for identifying turning points; we operationalized turning points as when the impact direction changed and continued for at least two segments. Positive turning points were changes from misunderstanding towards understanding (e.g. -1, +1, +1) and negative turning points were changes from understanding towards misunderstanding (e.g. +1, -1, -1; Portolese et al., 2015, June; see Figure 3).

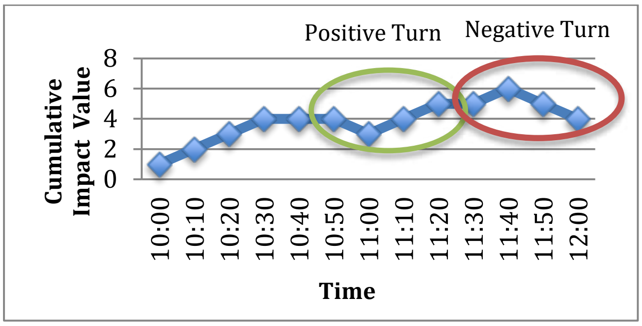

Figure 3. A prototype of how the turning points could be visualized graphically.

In order to understand these turning points, we zoomed into the micro segments in two locations: a) the meso level segment before understanding turned and b) the subsequent meso level segments during the change until two segments in the new direction occurred. These were not always the immediately subsequent segments as sometimes neutral segments spaced in between, such as -1, +1, 0, +1. For example, using the positive turn in Figure 3, the analysis for part (a) would unpack the ideas in the meso segment at time 10:50-11:00, and the analysis for part (b) would be between 11:00-11:20. Table 3 is an example of this breakdown from a turning point in our data. Following this, turning points were interpreted and grouped based on patterns of understanding and misunderstanding which emerged based on a) what happened before and b) what happened during the change in understanding. The analysis aimed to identify the nature of the events that caused the change, to provide us with tangible information on the development of the group processes and understanding.

Table 3. Example from our data of unpacking and analyzing a turning point

Turn Direction	(a) Ideas in Meso Segment Before Turn	(b) Ideas in Meso Segment (i) During Turn		(b) Ideas in Meso Segment (ii) During Turn
Positive	Experimental Design (-1)	Observation (+1)	Technical (+1)	Experimental Design (+1)

Findings and discussion

Patterns of turning points in our data

In total, we identified and analyzed 26 turning points across the two dyads (17 positive, 9 negative). See Tables 4 and 5 for a summary of our turning points analyses. We grouped the turning points based on thematic patterns that emerged. Missing the bigger picture was a common theme, both as a precursor to positive turning points and as the substance of negative turning points. This type of problem can be elusive, because due to the nature of the problem, students likely do not realize there is an issue. Students' understanding turned for the better through additional observations, experimentation and elaborated discussions. Similarly, making incorrect observations was common, both as a precursor to positive turning points and as the substance of negative turning points. Additional correct observations helped students re-ground their developing understanding in correct ideas. We found one instance of students' confusion on a conceptual level as a turning point – engaging in the experimental cycle of ideas rebuilt understanding from the ground up. Similarly, misunderstanding caused by poor experimental designs turned via focused predictions that improved subsequent experimental designs. A less productive pattern was students experiencing technical or representational confusion or errors, which unfortunately in some cases led to deeper misunderstanding – this type of floundering did not appear productive for deep understanding. We would lastly like to highlight a group of negative turning points (last row in Table 5) as examples of instances where basic declarative understanding, even when correct, does not necessarily indicate or lead to deeper conceptual understanding. As seen in our cases, this was even found when the students were largely doing all "right" things regarding following instructions and using a good experimental procedure.

Table 4: Positive turning points turning towards understanding (moving towards solution)

What went wrong: Pre turning point	Turning point frequencies	Key micro level characteristics during turn	How understanding turned & emerged
Conceptual confusion	1	Correct suggestions, experimental design, and observation	Cycles of suggestion, experiments, and observation required to rebuild understanding from the ground up
Technical or representational errors	1	Task ideas (reorientation)	Moved along despite unresolved technical challenges or representational misunderstandings
	2	Correct observation, explanation, and representation ideas	Extended observations and discussions about technical and representational aspects
Incorrect observations	4	Correct observations; Correct suggestions (for some)	New observations corrected misconceived observations
Poor experimental design	4	Correct predictions and experimental design	Predictions typically fueled improved experimental designs

Missing bigger picture	5	Correct observations, explanations, understanding, and experimental design	Additional experimentation and elaborated discussion

Table 5: Negative turning points turning away from understanding (moving away from solution)

What went well: Pre turning point	Turning point frequencies	Key micro level characteristics during turn	How misunderstanding turned & emerged
Understanding based on extended experimentation and discussion	2	Incorrect technical and representation ideas	Misunderstanding based on technical errors
Correct explanation; Solved technical problem; Challenged incorrect idea	3	Incorrect observations	Incorrect observations
Correct task orientation; Focused experiment; Partially correct understanding	4	Incorrect or correct observations, incorrect explanations and misunderstanding	Incorrect elaboration and explanation of understanding; Focusing on the wrong details (missing bigger picture)

Discussing, comparing and evaluating our coding approach

Our analyses of turning points can be related to the pivotal moments in the multivocality project (Suthers, 2013), in particular with Chiu's (2013) statistical discourse analysis of a fractions lesson. Chiu used statistical modeling to map the characteristics of conversation turns (their micro level) within a broader context of the classroom (their meso level). Similar to our method, Chiu evaluated conversation turns with consideration of the context of the previous action, and utilizing a -/+/0 scale as we did. Different to our approach, Chiu considered various dimensions of the micro level, separately considering if each turn: was correct/valid, invited further participation, contained novel content, and was an agreement with the previous turn. Our grain sizes increased in smaller increments, and we kept the students' understanding at the center throughout – our largest macro grain size was still about the students' work (their experiments) and we integrated the classroom context as relevant throughout the grain sizes. We agree with Chiu that student disagreements, even when incorrect, could be productive at stimulating thought, action, and other perspectives.

Also in the multivocality project, Sawyer, Frey, and Brown (2013) found that two main collaborative moves that enable knowledge-building discourse were: (1) collaborative elaboration of ideas and (2) self-monitoring content understanding. On the flip side, they found that groups experienced problems when: (1) the critical features of the problem were not explicitly focused on, (2) students asked closed questions, and (3) a lack of elaboration. Our findings are very much in line with this; we also found that elaboration of ideas was a critical component for the progression of student understanding. The problem of missing critical features we believe is related to the problem of missing the bigger picture, which can be a particularly elusive problem as students might have a false sense of confidence and not realize what they do not know. Similarly, Mameli and Molinari (2011) analyzed the interactive micro-processes and turning points in classroom discourse, and similar to us, found that students had challenges with focusing on the wrong details or making incorrect observations which can lead down a garden path of misunderstanding. They highlight the importance of describing the "order and disorder" in a classroom, which we believe is similar to our consideration of when things turn for the better or for the worse – we believe there is great value in examining when both understanding and misunderstanding emerge. However, their work was limited in that they did not reach a precise definition of turning points – one of the utilities in our methodology is that a turning point could be defined quite precisely.

Sanderson and Fisher (1994) outline "8Cs" as 8 different general transformations that can be done as "primitive smoothing operations" on rich, sequential, human computer interaction data; they include chunking, comments, codes, connections, comparison, constraints, conversion, and computation. As these can be conceptualized as the components to be considered in an analysis that involves analyzing video or observation data in an exploratory way (Dyke, Lund, Suthers, & Teplovs, 2013) such as ours, we have evaluated our coding approach against these 8Cs. Regarding chunking, which refers to how the data is grouped into phases – with great consideration for how our choice of grain size might influence the understanding gleaned, our approach incorporates three parallel and hierarchical grain sizes. Regarding comments, which refers to the informal ideas and notes about the data an analyst might have – in our analysis software (Elan), we preserved multiple tiers for these kinds of notes. We had tiers marked for notes to self, ideas for new coding directions, and notes to discuss at group meetings. It was useful that we built an organized way to record the unorganized ideas that emerge during analysis. Regarding codes, which refer to labels assigned to chunks to reduce variability while preserving meaning

– we found that such codes were useful at the micro idea level, but we feared that with the complexity of the larger meso and macro grain sizes that meaning would be lost if codes were used. However, at the end of our turning point analysis, we were able to meaningfully group together larger-scale patterns. Regarding connections, we have a method that is very strong at providing connections at multiple grain sizes in temporality, and also very strong at connecting related actions (e.g. written responses, computer actions, speakers, classroom activities) in temporality, but perhaps more could be done to connect events in a context that is not organized by temporality or idea category type. Regarding comparison, we had an initial inter-rater reliability of 87% (with 100% agreement following discussion) with a second coder (third author) analyzing approximately 10% of each video. Inter-rater reliability for segment boundaries was not measured, but could be worthwhile in future applications. Within our data in related work (Portolese et al., 2015, April; 2015, June) we compare dyads working in slightly different conditions, and we broadly compare our work to other related collaborative learning and productive failure work. Regarding constraints, which refers to filtering and selecting a part of the data for further analyses – we have done this in the greatest sense with our phase two turning points analyses of what we consider pivotal moments in changes in understanding. We also found it useful when understanding our data to play around with including and excluding various idea categories at the micro level. When doing the initial analyses, we often filtered codes of the same kind to ensure consistency. Regarding conversion, which refers to transforming the data, one of our aims for our next application of this scheme is to experiment with new ways to change and improve the way we represent the data, including the multiple layers and emphasis on patterns in turning points. Finally, regarding computation, our use of numeric impact coding and code categories at the micro level allowed for meaningful numerical summaries. Overall, considering our approach within this framework, our approach has much strength and some specific, tangible areas for us to continue to improve the design and representation of our approach.

Conclusions and implications

Our method of analysis is a useful strategy for unpacking how groups' understanding and/or misunderstanding emerges over time in a deep and rich way. Multiple, integrated grain sizes allow for a deep understanding of critical moments such as turning points by being able to "look up" and "look down" (Russ, Scherr, Hammer, & Mikeska, 2008) a level at the explanatory content and mechanisms. Our results suggest that it is critical that students attempt to elaborate for themselves regarding what they are observing, and perhaps it could be wise that teachers/facilitators check in on their elaborated understanding. There may be utility in explicitly encouraging frequent observation and potentially strengthening students' observation skills. In line with Kapur and Bielaczyc (2012), we saw a benefit when students persevered and generated as many diverse ideas as possible. Students may need to engage in written responses to demonstrate understanding or lack of understanding, and formative feedback on this developing content understanding at a deep level would be productive. When students have misunderstanding, rebuilding understanding from the ground up utilizing the classic experimental procedure (Figure 1) can be useful. It appears that floundering in relation to representative and technical elements is less productive, and more extensive support could be useful in this area. We look forward to continuing our research program by continuing to refine our methodology and strengthen our conclusions as we apply and expand it with larger data sets. Overall, the development of group understanding is an incremental process (Jeong, 2013); understanding the mechanisms of students' developing success and failure in these increments is an important key to understanding how collaborative scientific understanding emerges.

References

Chiu, M. (2013). Social metacognition, micro-creativity, and justifications: Statistical discourse analysis of a mathematics classroom conversation. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 141-159): Springer US.

Damşa, C. (2014). The multi-layered nature of small-group learning: Productive interactions in object-oriented collaboration. *International Journal of Computer-Supported Collaborative Learning, 9*, 247-281.

Dyke, G., Lund, K., Suthers, D. D., & Teplovs, C. (2013). Analytic representations and affordances for productive multivocality. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (639-658): Springer US.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393-408.

Greeno, J., Collins, A., & Resnick, R. (1992). Cognition and learning. In B. Berliner & R. Calfee (Eds.), *Handbook of Educational Psychology* (pp.15-46). New York: Simon & Shuster MacMillan.

Jacobson, M. J. (2000). Problem solving about complex systems: Differences between experts and novices. In B. Fishman & S. O'Connor-Divelbiss (Eds.), *Fourth International Conference of the Learning Sciences* (pp. 14-21). Mahwah, NJ: Erlbaum.

Jacobson, M. J., Kim, B., Pathak, S., & Zhang, B. (2013). To guide or not to guide: Issues in the sequencing of pedagogical structure in computational model-based learning. *Interactive Learning Environments*. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/10494820.2013.792845

Jacobson, M. J., & Markauskaite, L. (2015, April). *Understanding complex systems and climate change: Learning designs with agent-based models, productive failure, and analogical encoding*. Paper presented at the 2015 Annual Meeting of the American Educational Research Association, Chicago, IL.

Jeong, H. (2013). Development of group understanding via the construction of physical and technological artifacts. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 331-351): Springer US.

Kapur, M., & Bielaczyc. K. (2012). Designing for productive failure. *The Journal of the Learning Sciences*, *21*, 45-83.

Kapur, M., & Kinzer, C. K. (2007). Examining the effect of problem type in a synchronous computer-supported collaborative learning (CSCL) environment. *Educational Technology Research and Development, 55*, 439-459.

Kapur, M., Voiklis, J., & Kinzer, C. K. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers and Education, 51*, 54–66.

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*, 75-86.

Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction, 34*, 74-85.

Mameli, C., & Molinari, L. (2013). Interactive micro-processes in classroom discourse: Turning points and emergent meanings. *Research papers in education, 28*(2), 196-211.

Nersessian, N. J. (2008). Mental modeling in conceptual change. In S. Vosniadou (Ed.), *International Handbook of Research on Conceptual Change* (pp. 391-416). New York, NY: Routledge.

Portolese, A., Markauskaite, L., Lai, P. K., & Jacobson, M. J. (2015, April). *Model-based learning with productive failure and analogical encoding: Unpacking learning dynamics with contrasting designs*. Paper presented at the 2015 Annual Meeting of the American Educational Research Association, Chicago, IL.

Portolese, A., Markauskaite, L., Lai, P. K., & Jacobson, M. J. (2015, June). *How collaborative successes and failures become productive: An exploration of emerging understanding and misunderstanding turning points in model-based learning with productive failure.* In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.). Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 2 (pp. 655-656). Gothenburg, Sweden: The International Society of the Learning Sciences.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences, 2*, 235-276.

Russ, R. S., Scherr, R. E., Hammer, D., & Mikeska, J. (2008). Recognizing mechanistic reasoning in student scientific inquiry: A framework for discourse analysis developed from philosophy of science. *Science education, 92*, 499-525.

Säljö, R. (1991). Learning and mediation: Fitting reality into a table. *Learning and Instruction*, *1*, 261-272.

Sanderson, P. M., & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human–Computer Interaction, 9*, 251-317.

Sawyer, K., Frey, R., & Brown, P. (2013). Knowledge building discourse in peer-led team learning (PLTL) groups in first-year general chemistry. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 191-204): Springer US.

Suthers, D. (2013). The productive multivocality project: Origins and objectives. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs, & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions* (pp. 3-19): Springer US.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction, 11*, 381-419.

Wilensky, U. (1999). *NetLogo*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling. Northwestern University (http://ccl.northwestern.edu/netlogo).

Acknowledgments

We thank participating teachers and students, and the anonymous reviewers for their productive suggestions on earlier phases of this work. This research was supported by the Australian Research Council grant #LP100100594.