

**DISCOVERING PROCESSES AND PATTERNS OF LEARNING IN
COLLABORATIVE LEARNING ENVIRONMENTS USING MULTI-MODAL
DISCOURSE ANALYSIS**

KATE THOMPSON

*Centre for Computer Supported Learning and Cognition,
Faculty of Education and Social Work, University of Sydney, Australia
kate.thompson@sydney.edu.au*

SHANNON KENNEDY-CLARK

*School of Education,
University of Notre Dame, Australia
shannon.kennedy-clark@nd.edu.au*

LINA MARKAUSKAITE

*Centre for Computer Supported Learning and Cognition,
Faculty of Education and Social Work, University of Sydney, Australia
lina.markauskaite@sydney.edu.au*

VILAYTHONG SOUTHAVILAY

*School of Information Technologies,
University of Sydney, Australia
vstoto@it.usyd.edu.au*

Multimodal learning analytics, with a focus on discourse analysis, can be used to discover, and subsequently understand, the processes and patterns of learning in complex learning environments. Our work builds upon and integrates two types of research: (a) process analytic approaches of dynamically captured video and computer-screen activity and (b) learning analytics. By combining previous analyses of a dataset with new analyses of the processes of learning, patterns of successful and unsuccessful collaboration were identified. In this paper, the results of the application of a heuristics miner to utterances coded with the Decision-Function Coding Scheme, are combined with the results of First Order Markov transitions and in-depth linguistic analysis of the discourse to analyse the processes of collaborative problem solving within a scenario-based virtual world. The analysis of dependency graphs extracted from students' event logs revealed problem solving actions enacted by students, as well as the dependency relationships between these actions. The addition of in-depth linguistic analysis explained the micro-level discourse of students, producing the observable patterns. Integration of these findings with those previously reported added to the depth of our understanding about this complex learning environment. We conclude with a discussion about the design of the tasks, the processes of collaboration, and the analytic approach that is presented in this paper.

Keywords: Multimodal learning analytics; CSCL; virtual worlds; design for learning.

1. Introduction

“as more human communication takes place in the networked world for education, commerce, and social activity, an extensive digital trace is being created, a deluge of behavioural data. These data are extremely valuable for modeling human activity and for tailoring responses to the individual—whether for learning or for commerce.” (Borgman et al., 2008, p. 24)

Over the last decades, research at the intersection of cognitive science and artificial intelligence made significant progress in capturing and exploring knowledge acquisition processes and human-computer interaction with learning software (Schulte-Mecklenbeck, Kuehberger, & Ranyard, 2011). In order to investigate knowledge processes, predict outcomes and develop adaptive techniques, a range of research methods to investigate the processes of learning have evolved. These methods include several broad categories: (a) methods for information acquisition, such as eye movement and active information search, (b) methods for tracing thinking processes and information integration, such as think aloud and structured response elicitation, and (c) methods for tracing psychological, neurological and other concomitant of cognitive process, such as reaction time and pupil dilation. In education, these methods range from capturing and analysing software log files, to video observations, screen recordings, and eye tracking traces (Cox, 2007; Derry et al., 2010).

Similarly, research on computer-supported collaborative learning (CSCL) has made significant progress in capturing and analysing student communication and joint decision-making processes in collaborative learning environments. This particularly could be said about a number of studies that have explored how students navigate, communicate and collaborate in asynchronous learning management systems (Dringus & Ellis, 2005; Lin, Hsieh, & Chuang, 2009; Macfadyen & Dawson, 2010). The main enabler of this type of research is the variety of digital traces of students' interaction within and with software, captured in digital media. As the National Science Foundation's Taskforce on Cyberlearning report argues, these learning traces could “aid researchers in developing a more complete and accurate scientific understanding of what makes learning most productive and enjoyable” (Borgman et al., 2008, p. 26). However, this extensive digital trace also requires new research methods and tools to handle the voluminous and complex data generated, and to assist researchers to discover meaningful information. Blikstein (2013) argues that automated, fine-grained analysis could be used to give researchers insight, and could aid in the design of learning tasks, as well as be used to inform assessment. He defines multimodal learning analytics as “a set of techniques that can be used to collect multiple sources of data in high-frequency (video, logs, audio, gestures, biosensors), synchronize and code the data, and examine learning in realistic, ecologically valid, social, mixed-media learning environments” (p. 105).

Although event order and interaction patterns are key aspects characterising productive and non-productive learning in CSCL environments, findings in this research

field are dominated by pre- and post- collaboration measures, such as knowledge and attitudinal tests (Gress, Fior, Hadwin, & Winne, 2010). As Reimann (2009) argues, the theoretical constructs and methods employed in CSCL often neglect to make full use of information relating to the order of events. Kapur (2011) begins to address some differences in the decision-making processes in structured and unstructured environments, and finds that an analysis of temporal measures when carried out as part of a multimodal approach about sequences and transitions enabled researchers to move between macro and micro properties and behaviours of the problem solving process. One of the major challenges for adaptive collaborative learning systems is to develop models of learners' decision-making process that take into account students' interaction with each other as well as with the software environment and that accurately predict student decision-making behaviour in response to these interactions. These collaborative decision-making processes in immersive virtual spaces, such as Multi-User Virtual Environments (MUVES), are particularly important and poorly understood. Among techniques that can be used to discover behaviour patterns of student collaboration and interaction are data mining (Fern, Komireddy, Grigoreanu, & Bumett, 2010; Romero & Ventura, 2006), stochastic modelling and other process research approaches (Poole, Andrew, Dooley, & Holmes, 2000). More recently, research into the processes of learning has addressed decision-making (Kapur, 2011), knowledge construction (Wise & Chiu, 2011), tool use (Thompson & Kelly, 2012), expertise (Thompson, Ashe, Carvalho, Goodyear, Kelly, & Parisio, 2013) and gestures (Evans, Feenstra, Ryon, & McNeill, 2011).

This paper describes our (methodological) experiences capturing and analysing student learning patterns in a MUVE. Traditionally, learning analytics focuses on the use of "big data" (e.g. Macfadyen & Dawson, 2010) for analysis, such as, social network analysis (e.g. Haythornthwaite & de Laat, 2011). Other work has applied these techniques to the analysis of complex learning environments (Blikstein, 2013; Thompson, Ashe, Carvalho, et al., 2013). In our learning setting, students interacted in real-time during relatively short periods. Second, students have interacted not only with each other but also with software tools, and we include this stream of data in our analysis. Our work builds upon and integrates two types of research: process analytic approaches of dynamically captured video and computer screen activity and learning analytics. In this paper, we will address the collaboration (within the dyad and with the tool) with particular emphasis on the analytic techniques employed (multi-modal discourse analysis, incorporating first order Markov Models and heuristics mining). Finally, we will discuss the implications of these findings for the design of learning tasks for complex learning environments.

2. Background

2.1. *Virtual worlds*

There is a growing body of research that supports the use of game-based learning in education due to its potential to enable high quality learning across a range of educational contexts (see, for example, Ketelhut, Clarke, and Nelson, 2010). Over the past two decades, there has been a considerable amount of research undertaken on both the role, and pedagogy, of using computer games, virtual worlds and other similar technologies in education (Bailenson et al., 2008; Barab et al., 2009; Ketelhut et al., 2010; Shaffer & Gee, 2012; Watson, Mong, & Harris, 2011). Research to-date has shown that the value of game-based learning applications in maintaining student motivation and engagement is substantial (Squire, Barnett, Grant, & Higginbottom, 2004; Watson et al., 2011). Specifically, games, such as Quest Atlantis, and MUVES, such as River City, have been shown to increase learner self-efficacy, motivation and engagement (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005; Ketelhut, Dede, Clarke, & Nelson, 2006).

More recently, the focus has shifted from examining how learners feel after participating in a game-based activity to examining what they are actually doing while they are using the game (Keating & Sunakawa, 2010; Kennedy-Clark & Thompson, 2011; Kennedy-Clark, Thompson, & Richards, 2011). Numerous studies have shown that scalability in research on the use of virtual worlds is an issue in educational research and that what works well in one local context may not be appropriate for a wide scale implementation (Clarke, Dede, Ketelhut, & Nelson, 2006). This research has shifted the focus from the end product of the learner experience to the design of the environment.

2.2. *Computer-supported collaborative learning*

“Computer-supported collaborative learning (CSCL) refers to situations in which computer technology plays a significant role in shaping the collaboration” (Goodyear, Jones, & Thompson, 2014, p. 439). After many years of research, the field has moved beyond the naïve assumption that simply providing students with a computer-based tool (such as a virtual world) will result in successful collaboration, and identified the importance of the design of the tasks. CSCL is a broad topic that could include the variety of scales, methods of collaboration, and media (Dillenbourg, 1999). Thus, CSCL researchers inevitably interpret very different (often multimodal) data and, as Strijbos and Fischer (2007) have argued, such research often demands the integration of different analytical techniques. Given the complex nature of CSCL environments, and the importance of design for learning, scaffolding the use of technology, as well as the collaboration, and the task itself, are all considered to be of importance (Kennedy-Clark & Thompson, 2013a).

At present, there are only a small number of studies that examine collaboration specifically in game-based environments. These studies rely on chat data between remote participants to examine human collaborative activity (see, for example, Keating and Sunakawa, 2010; Steinkuehler, 2006). In this paper, the face-to-face conversations

between participants collaborating in dyads were analysed to gain an understanding of their collaborative decision-making. In a conversation there is a sequential development of an interaction where speakers see what happens and what happens next; hence, the question often asked in discourse analysis is why that now? (Nevile & Rendle-Short, 2009) In this study we used Poole and Holmes' Decision-Function Coding Scheme (DFCS) (1995), which has been used successfully in a number of other studies examining the processes of CSCL (see for example, Kapur, 2011; Reimann, 2009; Reimann, Frerejean, & Thompson, 2009; Thompson & Kelly, 2012; Thompson, Kennedy-Clark, Markauskaite, & Southavilay, 2011). Due to the nature of the task, as discussed above, the focus of this work was on participants' identification of navigational or task-based goals.

2.3. Data-driven methods for analysing interaction processes

Examining the processes and patterns of learners' behaviour is an area of research that is expanding in scope and methods. Sequential analytic techniques (such as first order Markov models or lag sequential analysis) in combination with learning analytic techniques (such as social network analysis or parts of speech tagging) mean that we are able to increase our understanding of the complex learning environments created (see Blikstein, 2013; Kapur, 2011; Kennedy-Clark & Thompson, 2013a; Reimann, 2009; Thompson, Ashe, Carvalho, et al., 2013).

One strand of related research has focused on methods to capture, integrate and analyse multiple sources of computer-mediated human interaction data. A typical source has been a recording of the computer screen integrated with "click" data (from log files), audio recording, or visual attention (Romero, Cox, Boulay, Lutz, & Bryant, 2007). Cox (2007) describes several other examples of technology enhanced research in which dynamically recorded computer screen data were used to analyse learning processes, such as to elucidate individual learning differences in logical proof strategies and to identify students' difficulties in learning clinical reasoning (see, also, Dyke, Lund, & Girardot, 2009) for a description of Trace Analysis Tool for Interaction Analysis (Tatiana)). Their approaches and tools primarily focus on how to support human-guided research via synchronisation and visualisation of data. While these tools help to transform and present data in a plethora of ways, the data deduction and pattern discovery process remains in essence a human activity. The advantages of studying the processes of interaction in CSCL research are commonly discussed and the links with learning outcomes often made (e.g. Cox, 2007). What is less common is an analysis and discussion of the processes of interaction with the tool used for learning (Thompson & Reimann, 2010) and it is the integration of these two areas that provides the focus for this paper.

Another strand of research has focused on the automatic analysis of student learning processes and patterns (Blikstein, 2013; Fern et al., 2010; Romero & Ventura, 2006). Much of this research primarily analysed student log files which were automatically generated from their interaction with software and/or each other using various statistical, data and text mining techniques, such as hidden Markov models, process mining, time

series and sequential pattern mining (Jeong, Biswas, Johnson, & Howard, 2010; Pechenizkiy, Tricka, Vasilyeva, van de Aalst, & Bra, 2009; Southavilay, Yacet, & Calvo, 2010). For example, Kay, Maisonneuve, Yacef, and Zaiane (2006) analysed student interaction sequences when they worked collaboratively on software development aiming to detect learning patterns that are indicative of team problems and success. Many other e-learning studies used different data mining algorithms for exploring student learning in e-learning systems, such as analysing student navigational behaviour in a virtual campus environment and identifying gifted students' learning paths (Romero & Ventura, 2006). These computer logs, however, produced more restricted data about student interaction with software (and each other) than could be captured via screen recording.

In other works, a combination of integrating multiple streams of data, and the automation of some data collection and analysis techniques are being trialled. In one, a framework to make sense of multiple streams of data, and apply appropriate analytic techniques to each, in an appropriate context (Thompson, Ashe, Carvalho, et al., 2013). In others, the two are used sequentially, with manual identification of patterns driving the focus of automated process analysis (Kennedy-Clark & Thompson, 2013a; Thompson & Kelly, 2012; Thompson, Ashe, Yeoman, & Parisio, 2013; Thompson, Ashe, Wardek, Yeoman, & Parisio, 2013; Thompson, Kennedy-Clark, Kelly, & Wheeler, 2013). In this paper we used data collected from video screen capture, as well as collaborative discourse collected during interaction with a MUVE. We integrated the results of a new, more automated process analytic technique (heuristics miner) into existing analyses of the data set, in order to demonstrate the impact on our understanding of the complex learning environment.

3. Methods

3.1. *Virtual Singapura*

Virtual Singapura was developed in Singapore as part of a collaborative project between researchers at Singapore Learning Sciences Laboratory (National Institute of Education) and Faculty in Computer Engineering and in Art, Design, and Media at Nanyang Technological University. Virtual Singapura was based on Harvard's River City and was designed to provide students with opportunities to develop scientific inquiry skills via interactions with authentic scientific problems. The scenario is similar in that the residents of the town were affected by three different types of disease: water borne, air borne and insect vector. The story or scenario for the virtual world lends itself to a scientific inquiry domain. Virtual Singapura is set in 19th century Singapore and is based on historical information about several disease epidemics during that period. Students needed to visit several conditions to collect data (Ketelhut et al., 2010). Within the task for learners, the "comparison condition" was the version of Singapore where everyone was sick from malaria, tuberculosis and cholera. There were several "experimental conditions": building new wells, draining the swamp, building new tenements, changing the dry season to the rainy season, changing the night soil coolie's practices. Students

By moving the cursor around the screen the users could select artefacts in the room, such as items on the table and pictures on the wall. If they selected an item, a text-based description would appear on the right-hand side of the screen. If the users right clicked on an agent with a name above the head, they could select from a drop down list of questions and receive scripted answers, such as, directions in the chat window at the bottom of the screen. At the top of the screen in Figure 1 the tool bar can be seen. Participants could use this bar to change perspectives, avatars and to navigate between conditions.

Students could also collect data from wells, bug catchers and air sampling stations. These would, respectively, provide counts for microbes such as cholera and e-coli, mosquitoes and tuberculosis. Students could find the sampling stations either by clicking on the map and teleporting to the location or by walking around the city. If they clicked on the sampling station, the counter would appear on the right-hand side of the screen. Hence, there were multiple routes through the virtual world and students were required to manage their navigation in order to collect sufficient evidence to prove or disprove their hypothesis.

3.2. Data collection

The in-world actions of 12 participants, eight undergraduate students and four postgraduate students, were recorded as they interacted with Virtual Singapore. The data collection was conducted in August 2010. The study took one hour. The participants were directed to complete only the first intervention on the impact on cholera rates on the result of building new wells. In this time, participants were required to participate in the virtual activity, which took approximately 30 to 40 minutes, and to complete post-test surveys. The participants were not provided with a pre-training phase. The instructions were provided in the introduction to the workbook and participants were directed to this information. The participants were assigned to a partner and were then given either a structured or unstructured workbook. The participants' actions were recorded using Camtasia screen capture software. Camtasia recorded the participants' movements in the world (e.g. moving the mouse, changing screens, clicking on objects), a headshot video (webcam) of the pair and their audio communication.

3.3. Multiple-streams of data

Interactions between participants were recorded (via webcam) synchronously with a video screen capture of their in-world actions. Transcriptions of the discourse were coded according to a modified version of the Decision Function Coding System (DFCS) (Poole & Holmes, 1995). DFCS was selected as it allowed for problem definition, orientation and solution development. The resulting coding system has seven main categories and five sub categories.

Screen capture added pertinent information about the context of discourse in terms of the location and activity in Virtual Singapore. The addition of category 7 reflects initial attempts to coordinate both conversational data and data from the video screen shots that were recorded. When coding the data according to the DFCS system, 7 represented

Table 1. Modified coding scheme for collaborative decision-making.

Code	Category	Definition
1	Problem definition	Statements that define or state the causes behind a problem Statements that evaluate problem analysis
2	Orientation	Statements that attempt to orient or guide the group's processes Statements that evaluate or reflect upon the group's progress or processes
Solution development		
3a	Solution analysis	Statements that concern criteria for the decision-making process A direct reference to the solution must be given
3b	Solution suggestions	Suggestion of alternatives
3c	Solution elaboration	Statements that provide detail or elaborate on a previously stated alternative
3d	Solution evaluation	Statements that evaluate alternatives and give reasons, explicit or implicit for the evaluations and therefore include a valuation
3e	Solution Confirmation	Statements that state the decision in its final form or ask for final group confirmation of the decision Statements that concern decisions linked to immediate results
4	Non task	Statements that do not have anything to do with the decision task
5	Agreement	Agreement – response to a question or statement e.g. yeah or yes
6	Disagreement	Disagreement – response to a question or statement e.g. nah or no
7	Implementation	Undertaking agreed upon action

implementation of a goal, in this case the implementation of a goal was often not a verbal utterance, but was reflected in an on screen action, such as teleporting or collecting data. This is particularly important as the additional code, implementation (7), was only able to be added because the implementation of decisions could be observed (Thompson et al., 2011).

3.4. Multi-modal analysis

Including the findings from multiple modes of analysis allows us to form a fuller understanding of the behaviour of learners in complex learning environments such as MUVes. In this paper, we will summarise findings reported in a number of other publications, also derived from this dataset, that include discourse analysis and first order Markov transitions (Kennedy-Clark & Thompson, 2013a; Thompson et al., 2011; Thompson, Kennedy-Clark, et al., 2013). We will then present new analytic work on the process of decision-making in two of the groups using the heuristics miner. As has been widely discussed, the coding and counting approach to understanding the processes in groups alone is not a satisfactory level of analysis (Kapur, 2011; Reimann, 2009). We present the counting and coding data, followed by the presentation and interpretation of

first order Markov transitions, the results of the application of the heuristics miner to two of the groups, and an in-depth linguistic analysis.

In order to prepare the data for analysis, the sequence representing all decision-making process of each dyad was defined as an event log. Each event log was composed of multiple events that represented a sequence of student activities related to a specific purpose, such as, “looking for a doctor”, “discussing with an agent”, or “collecting samples”. Events were considered to be autonomous instances representing the dyad’s decision-making model. Each coded decision-making action – such as, problem definition (code 1), orientation (code 2), solution analysis (code 3a) – was defined as an activity in the event log. In short, each event log included multiple events and each event included multiple activities. Following this definition, each activity can appear multiple times in an event and each event (or different variations of it) can appear several times in the event log (e.g. students may collect samples in multiple places and may talk with multiple agents). While it is possible to expect that certain events may follow a distinct pattern, in this analysis we explored all events together, with the aim of detecting the most important shared features and gaining insight into core patterns in each group’s overall decision-making process.

Process analysis

Process mining techniques aim to discover underlying patterns of various processes by extracting knowledge from observed process data, such as, recorded event logs in organisational management systems, student interactions with each other or software captured in learning software logs (Trčka, Pechenizkiy, & van der Aalst, 2010; van der Aalst, 2011). Process mining techniques involve three broad uses: (a) discovering new patterns that are not known a priori; (b) checking conformance of the observed behaviour to an a priori modelled workflow or behaviour model; (c) extending a priori process models by projecting discovered patterns back on to the initial models and adjusting the processes accordingly (Rozinat, de Jong, Gunther, & van der Aalst, 2007; Weijters & Ribeiro, 2010; Weijters, Van der Aalst, & De Medeiros, 2006). We adopted two approaches to analyse the processes: first order Markov transitions and a heuristics miner.

First-order Markov transitions show the probability that one state will follow the next. Each of the decision-making codes shown in Table 1 is considered to be a state. The dataset is then considered to be a sequence of states, ordered in time. A first-order Markov chain is created by calculating the probabilities for each state transition. Frequencies greater than 10 are recorded on the diagram, as are probabilities greater than 0.25. First-order Markov transitions have been used in other studies examining the patterns and processes of decision-making (Kennedy-Clark & Thompson, 2013a; Thompson & Kelly, 2012).

We also applied a heuristics miner algorithm (Weijters & Ribeiro, 2010; Weijters et al., 2006) to the same dataset, using ProM, an open source process mining framework (ProM, 2011). The heuristics miner was developed for exploratory mining of less structured process data, when the a priori workflow pattern was not known. This mining

algorithm allows the handling of logs with various kinds of “noise,” such as, diversions from common sequences or incomplete traces of process information. Such noise is common in learning data, particularly when data are derived from verbal and observed interactions of student decision-making.

Micro-level discourse analysis

Human coordination is achieved through the processes of social interaction, shared meaning and mutual understanding. Micro level discourse analysis focuses on parts of speech, such as, word types, utterances and affirmations. An analysis of parts of speech can identify speaker(s), hearer(s) /addressee(s) and the “spoken of’s’ roles in the context of an event (Kummerow, 2012). For example, personal pronouns can be analysed as key features of social relationships (Halliday, 1994), and can be linked to other interpersonal components such as the mood element, modality, and the role of clauses in offering and exchanging services and information. We related these micro-level patterns to the macro-level patterns previously identified in the data (Reimann et al., 2009; Thompson & Kelly, 2012).

4. Results

4.1. Coding and counting

The six recordings were transcribed and combined with the corresponding video and screen shot data to provide a detailed account of what the groups were saying and doing while they were using the virtual world. The results presented in this section are a summary of those presented in a number of other publications, all of which are acknowledged. Table 2 shows a summary of basic information about each dyad including the condition, structured or unstructured, to which they were assigned.

Of particular interest in Table 2 is the total number of utterances in each group. Groups 5 and 6 produced more than twice the utterances of the other groups. With the exception of group 6, the unstructured groups had fewer utterances than the structured groups, and a more even distribution between the members. Not all of the groups were successful in arriving at an outcome. Groups 1 and 5 (structured) ran out of time, but were engaged in the activity and remained on task. Groups 3 (structured) and 6 (unstructured) completed the activity and arrived at a conclusion. Groups 2 and 4 (unstructured) did not complete the activity, largely due to difficulties navigating in the world. Multiple measures have been used to analyse this data previously and allowed us to identify patterns of discourse analysis related to the Collaborative Process Analysis Coding Scheme (CPACS) (Kennedy-Clark & Thompson, 2013b) at the group level, as shown in Table 3 (presented in full in Kennedy-Clark and Thompson, 2013a). The categories for these patterns are based on macro-levels of discourse (action, content) as well as micro-level (attitude, tense, modality, pronouns).

Table 2. Summary of dyad information.

Composition				Utterances		Problem Solution
Dyad	Pseudonym	Gender	Education	Total	Split	
Structured Condition						
1	N	Female	Undergraduate	130	63 (49%)	No – ran out of time
	S	Male	Undergraduate		61 (47%)	
	Teacher				6 (4%)	
3	L	Male	Undergraduate	100	52 (52%)	Yes
	S	Female	Undergraduate		45 (45%)	
	Teacher				3 (3%)	
5	X	Male	Postgraduate	217	119 (55%)	No – ran out of time
	Y	Male	Postgraduate		91 (42%)	
	Teacher				7 (3%)	
Unstructured Condition						
2	K	Female	Undergraduate	26	15 (58%)	No
	G	Female	Undergraduate		11 (42%)	
4	T	Female	Undergraduate	69	37 (54%)	No
	U	Female	Undergraduate		28 (41%)	
	Teacher				4 (5%)	
6	A	Male	Postgraduate	240	136 (57%)	Yes
	B	Female	Postgraduate		94 (39%)	
	Teacher				10 (4%)	

Note: Elements of Table 2 have been presented in other publications (Kennedy-Clark & Thompson, 2011; Kennedy-Clark, Thompson, & Richards, 2011; Thompson et al., 2011). The full table was published in Kennedy-Clark and Thompson (2013b).

Table 3 compares the patterns of discourse at the macro and micro levels between two groups. As Kennedy-Clark and Thompson (2013a) report, Dyad 1, classified as unsuccessful, demonstrated patterns of interaction that were identified as successful, and it is clear that given more time, they would have completed the task. Processes existed for discussing the action that they were taking, and the focus of the content of their discourse was on navigation in the virtual world, another reason for their unsuccessful completion of the task. The micro-level patterns indicated that they were engaged with the task, and engaged with both reflection and planning. Dyad 6 demonstrated a more cyclical process of collaboration, in many of the categories presented in Table 3. The context determined their confidence and the pronouns used. This group was able to arrive at a successful solution despite the lack of supportive materials, through the processes of discourse that they derived.

Table 3. Identified group patterns.

	Action	Content	Attitude	Tense	Modality	Pronouns
Dyad 1 (unsuccessful)	Idea generation	Navigation	Learner dependent	Past, present, future	Content dependent	Action dependent
Dyad 6 (successful)	Process – Cyclic	Plan, navigate, collect data	Neutral	Planning and reflecting	Context dependent	Context dependent

Previously reported results have also included discovering patterns of content in relation to decision-making (the macro level of discourse), in combination with student recall of historical data, and content word-arcs (Kennedy-Clark & Thompson, 2013a). In addition, a micro-level analysis of the discourse revealed patterns in learners’ use of tense that indicated a lack of planning as well as a lack of true process in navigating the virtual world in the unsuccessful groups (Thompson, Kennedy-Clark, Wheeler, & Kelly, 2014).

4.2. First-order Markov transitions

First-order Markov transitions were calculated for each group. Figures 2 to 7 shows the process of decision-making for each group. In each case, all codes in the DFCS are included in rectangular boxes. Links between these boxes with arrows, show the order. Two numbers are written on each line – x(y), where x represents the frequency of the link, and y represents the probability that that link occurred. To give a hypothetical example, 10(0.50) on a line with an arrow pointing from problem definition to orientation can be interpreted as follows: the *orientation* code followed *problem definition* on ten occasions, which represented 50% of the codes that followed *problem definition*. Not all links are represented – only those with a frequency greater than or equal to ten, or a probability

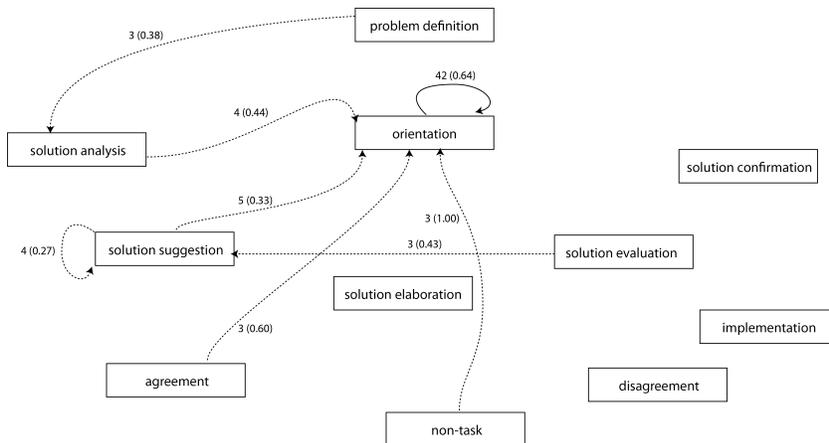


Figure 2. Markov transition diagram for Group 1, structured, unsuccessful (time).

greater than or equal to 0.25 are included in Figure 2. Solid lines represent both (for example, a frequency of 42 and a probability of 0.64), whereas dotted lines represent those cases in which only one criterion is met. Links with a frequency of one or two were excluded.

Figure 2 shows very few links between phases of decision-making, and most included frequencies of less than ten. *Problem definition* was followed by *solution analysis* in three cases. *Orientation* was the main focus for this group. Once they began *orientation*, students tended to continue to do this (64% of the time). The initial stages of solution development – *solution analysis* and *solution suggestion* – were followed by *orientation*. When students *evaluated* a solution, they returned to *solution suggestion*. *Orientation* was also used as an intermediary step after *agreement* and *non-task* discussion.

Group 2 only produced 26 utterances in total. In this group, there is no indication that a process of decision-making was undertaken. If they engaged in *orientation* or *implementation*, they continued to do so. This implies that the group focused on interacting with the virtual world, but was in need of assistance.

After Group 3 *defined the problem*, they engaged in *orientation*. They also returned to *orientation* after *solution suggestion*, *agreement* and *disagreement*. As one of the structured groups, this was most often the support materials that they used to orient themselves. Links *from* orientation are not shown because they were relatively equally distributed between the other codes, which meant that none had a large enough proportion or frequency to report. Figure 4 shows that, again, *orientation* played an important role in this group's process of decision-making.

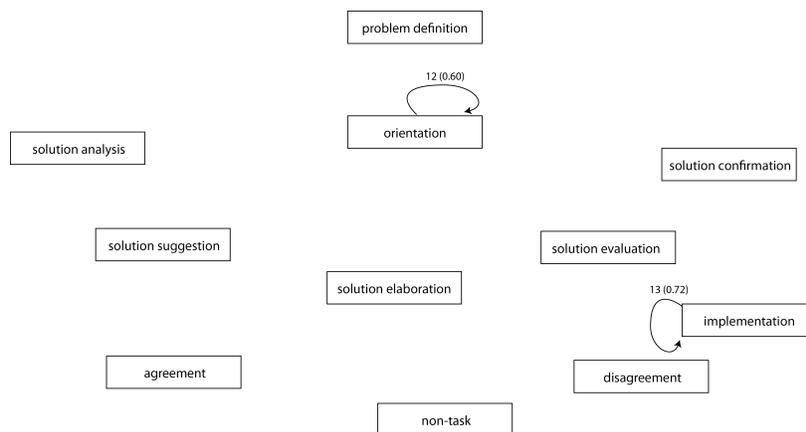


Figure 3. Markov transition diagram for Group 2, unstructured, unsuccessful.

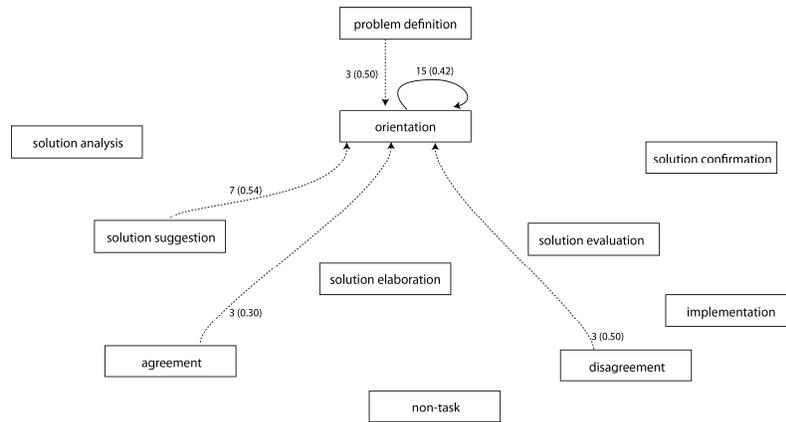


Figure 4. Markov transition diagram for Group 3, structured, successful.

Group 4 (see Figure 5) was the most disengaged group and had the most problems in understanding the purpose of the activity. In this unstructured group *orientation* was again important, and again was not consistently followed by any specific further action, apart from further *orientation*. There was no clear path through solution development. Figure 5 indicates that once the group engaged in *non-task* discourse they tended to continue before returning to *orientation*, as they did for *solution suggestion*.

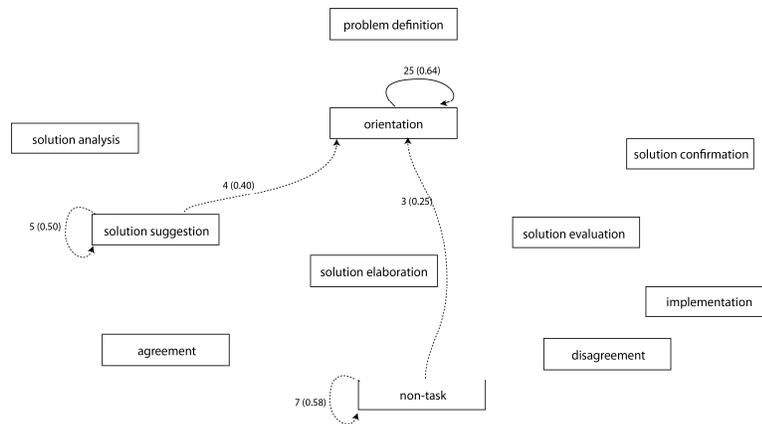


Figure 5. Markov transition diagram for Group 4, unstructured, unsuccessful.

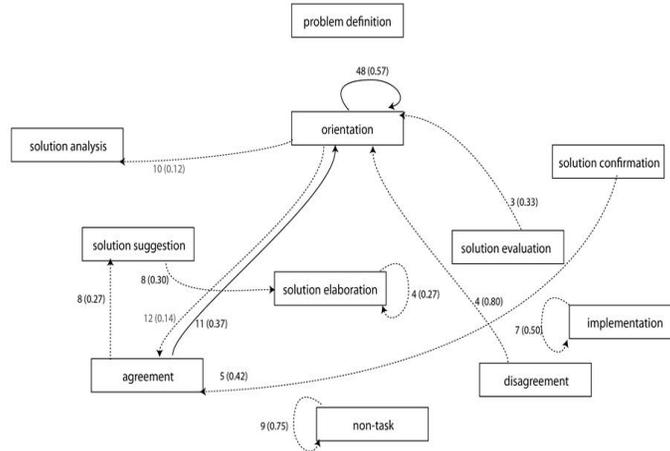


Figure 6. Markov transition diagram for Group 5, structured, unsuccessful (time).

Group 5 did not complete their structured activity, running out of time as they experienced ongoing technical problems with the world. However, when analysing the process data related to the discourse it was apparent that the group was engaged and motivated. Figure 6 shows that, like many of the groups, orientation played a central role throughout the solution development phase. Orientation was followed by itself (57% of the time), solution analysis (12%) or agreement (14%). This may suggest that the materials were able to help students begin the process of solution development. Agreement played an important role for this group, with orientation and solution confirmation both followed by agreement, and then a loop to either solution suggestion or

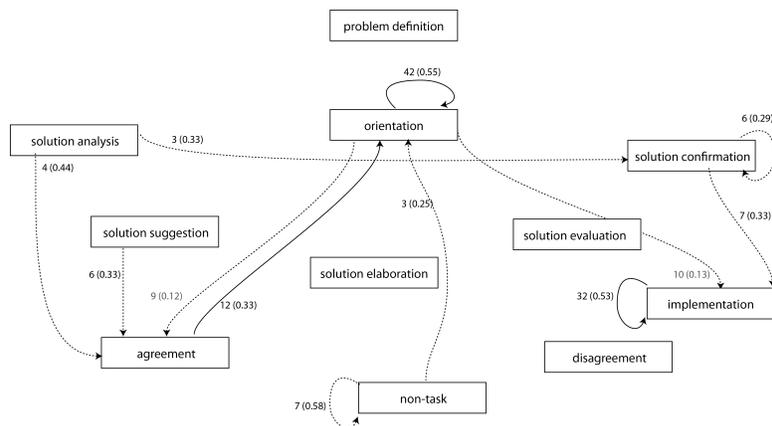


Figure 7. Markov transition diagram for Group 6, unstructured, successful.

back to orientation. When this group engaged in implementing a decision, they tended to remain doing that.

A pattern of collaboration was established early in the activity of Group 6. This group had a different pattern to the other unstructured groups, with central events being orientation, agreement, and implementation, similar to Group 5. Students moved from solution analysis to agreement, and then to orientation. Sometimes solution analysis was followed by solution confirmation. Solution suggestions were also followed by agreement. Solution confirmation was followed by implementation, or further confirmation. Implementation of a decision was informed by solution development, and both members of the dyad reached consensus before acting.

4.3. *Heuristic miner*

Each group's event log was also analysed using a Heuristics Miner algorithm. This mining process is based on dependencies between activities using a frequency metric to represent the certainty of the dependency. In other words, that an activity is causally related to, or dependent upon the activities that precede it. We used an all-activities-connected heuristic, which aims to include all activities in the model and find the links between all activities. Using this algorithm, in each mining step, the ingoing and the outgoing connections with the highest dependency value are identified and included in the model, until all activities are connected. The final heuristic model does not necessarily represent all possible links and dependencies between all activities, but depicts the most strongly dependent actions. Further details about this mining algorithm can be found in the following sources (Rozinat et al., 2007; Weijters & Ribeiro, 2010; Weijters et al., 2006).

Based on the results of the linguistic analysis and Markov transitions, the decision-making patterns of two of the dyads were selected for further exploration using this method: Group 1 and Group 6. One aim of this paper has been to demonstrate the value of multimodal discourse analysis and the application of multiple analytic techniques. In order to gain deeper insights into the decision-making behaviours of both unsuccessful and successful dyads, we specifically selected one unsuccessful dyad from the structured group and one successful dyad from the unstructured group. Group 1 was a structured group, and, despite the provision of guidance, was not successful in completing the task. Group 6 was an unstructured group, and, without instructions, successfully completed the task. Both dyads were similarly interested in the activity and seriously engaged with the task. The heuristic decision-making models are represented in Figures 8(a) and 8(b). The labels and numbers in the boxes indicate the occurrences of different decision-making activities. Two numbers near the path show the relative importance of transition from one activity to another. The upper number indicates heuristic strength (dependency). Its values can range from -1 to 1 with numbers close to 1 indicating a strong dependency of activity that follows on the preceding action. The lower number indicates the frequency of the observed link between two activities in the event logs. Taking into account that the

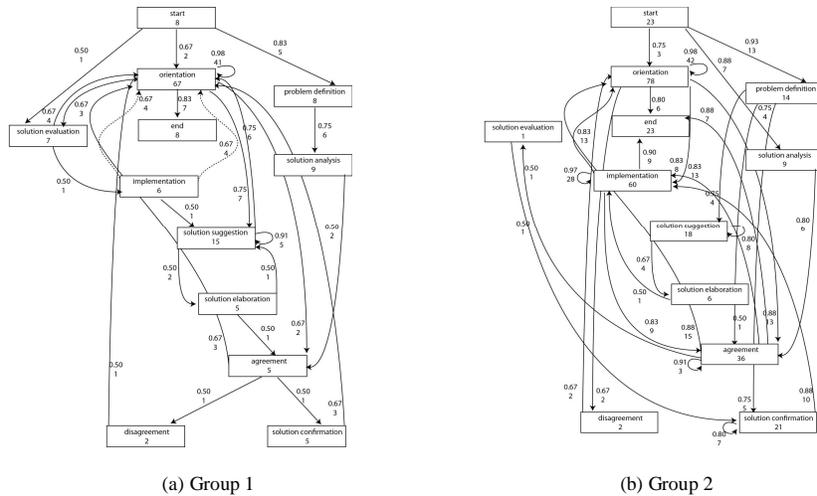


Figure 8. Heuristic decision-making model for Group 1 and 6.

event logs were relatively small, in our interpretations we used both occurrences and dependences between activities.

The heuristic model of Group 1 shows that the dominant and central action was Orientation (67 instances). It was followed by Solution Suggestion but the latter activity was more than four times less frequent (15 instances). Overall, relatively large frequencies of both activities indicate that Group 1 focused on orienting their inquiry process and making new suggestions. Small loops from Orientation to Orientation and from Solution Suggestion to Solution Suggestion show the repetitive nature of these activities. This repetitive behaviour is particularly high for Orientation as one Orientation action was immediately followed by another Orientation in 41 instances out of 67 activities. Further, there were three small two event cycles between Orientation and Solution Suggestion, Orientation and Solution Evaluation, and Orientation and Implementation with dependencies between 0.67 and 0.75. They indicate that two dominant solution development actions and Implementation were also intertwined with Orientation. It is interesting to note that consensus-building actions, such as Agreement and Solution Confirmation, were rare in this group and Orientation behaviours were rarely followed by Agreement (2 out of 5 instances).

Overall, the heuristic model shows that 5 out of 8 events started from problem definition, which was often followed by solution analysis (6 out of 8). The dependency measures between these events are not high (0.75-0.833), but taking into account a relatively short event log and a small number of events, this outcome is indicative of a relatively consistent pattern in the initial stages of Group 1's decision-making. However, only in two cases was this initial Solution Analysis followed by Agreement. A weak

dependency between Solution Analysis and Agreement (0.5) and the lack of other connections confirms that Group 1 did not have a clear decision-making pathway beyond this initial stage, as none of the follow up activities depended on the outcomes of the initial analysis strongly. The rest of the group's actions were focused on Orientation that was intertwined with the development of the solution. While the group sometimes reached Implementation, this activity was relatively rare (6 instances) and was usually followed by Orientation. This heuristic model is indicative of a "process-driven" behaviour. Given the relatively rare occurrence of Agreement, Disagreement and Confirmation in this group (from 2 to 5 instances), it appears that the group's interactions and discussions about the process were not necessary intertwined with the group's attempts to reach a shared agreement.

The heuristic model of Dyad 6 depicted quite different features of decision-making process (Figure 8b). First, three behaviours were dominant in this group: Orientation (78 instances), Implementation (60 instances) and Agreement (36 instances). Orientation and Implementation activities were quite repetitive, which is depicted in the heuristic model by high dependency values and high frequencies near two small feedback loops: from Orientation to Orientation (42 actions out of 78); and from Implementation to Implementation (28 actions out of 60). The central behaviour was Agreement. It was involved in two relatively high dependency feedback loops. The first feedback loop was between Agreement and Orientation (dependency 0.88). It indicates that one group member's attempts to orient or guide decision-making process was often followed by another member's agreement. The second feedback loop was between Agreement and Implementation (dependency 0.83). It depicted that Implementation of agreed actions was often followed by further Agreement on the course during implementation. It was also interesting to note that Group 6 frequently ended events directly after Implementation (9 instances), or after (perhaps further) Orientation (6 instances) and Agreement (7 instances). This indicates that students' subsequent decisions about the transition to a new event were often informed by the results of previous implementation.

The heuristic model of Group 6 depicted that the dyad's decisions (events) often started from Problem Definition (13 events) or Solution Analysis (7 events). This indicated that, in the initial stages of each decision-making event, Group 6 often focused on understanding the problem. These initial problem analysis actions were usually followed by Agreement or Solution Confirmation (dependencies between 0.75 and 0.8) indicating that the group usually reached agreement before starting other inquiry activities. The group then usually proceeded to Orientation or directly started Implementation. While Implementation, Orientation and Agreement activities were repetitive and intertwined, they primarily focused on Implementation rather than further discussion and development of the solution which was agreed beforehand.

Overall, this heuristic model depicted a purposeful "action-based" inquiry pattern. It indicated three important features of their inquiry processes. First, Agreement played an important mediating role in group-decision-making and it followed many other decision-making activities, including initial solution analysis of the problem, Orientation about

groups' process and Implementation. This shared role of Agreement and close association with many decision-making activities might overshadow some (mediated by Agreement) dependencies between other decision-making actions, however, it indicates the collaborative nature of Group 6's inquiry. Second, the group's actions in each event were significantly informed by the initial definition and analysis of the situation, which preceded Orientation and Implementation. Finally, while Orientation was the most frequent action, it was embedded in a consistent pattern of decision-making. Specifically, Orientation activities were primarily linked with Implementation and Agreement, whereas solution development actions were done before and not intertwined with more practical decisions about what to do (Orientation).

The two heuristic models demonstrate three major differences between the groups. Group 1: little engaged in establishing mutual agreement, mixed solution development with orientation actions and focused on orientation rather than implementation activities. In contrast, Group 6: sought to establish mutual agreement about the proposed solutions and implementation; combined orientation with implementation actions; and significantly focused on actions and used outcomes to further inform their inquiry process.

5. Discussions

The analysis of the counting and coding data established that there was a difference in how the structured and unstructured groups approached the inquiry. Dyads that were given structure had a more regulated approach to solving the problem. The workbooks provided guidance on where to gather data in the virtual world, and in what order. It was found that the dyads that made the most use of the workbooks tended to be more successful than the groups that did not use the workbooks. It was shown that dyads 2 and 4 (unstructured and unsuccessful) had the least utterances, were off task, could not move beyond navigation and did not arrive at a problem solution. Dyad 6 (unstructured and successful) was able to complete the problem and establish a regulated pattern of goal identification and goal implementation. Dyad 6 had a superior problem solving approach to the other dyads. It was found that the dyads that oriented their processes and explored alternatives before implementing a goal tended to be more successful progressing through the virtual inquiry than the dyads that did not use such strategies.

The results of the process analysis indicate that the groups that established patterns of collaboration were more successful in the task and were more likely to converge on a goal and arrive at an outcome. It also shows that students in the structured groups used the materials, orienting themselves at most stages of developing a solution. Rather than simply relying on agreement between themselves, they also required agreement with the supports given. This appears to have added considerable time to the process of collaboration. The importance of orientation, agreement, and implementation was evident initially, however, the context in which they occurred was only visible when we considered the order in which they happened. Markov transitions and the heuristics miner algorithm allowed us to do this. They showed that those groups that were unstructured

and unsuccessful had no central activities, and while orientation played a role, there were no supportive materials available for students to use.

In those groups that were unsuccessful and structured, we discovered that the structure played a key role in their decision-making process but that this seems to have resulted in more time being taken as students returned to orientation after most steps in their solution development. It also appears in these groups that students were intent on agreeing with the instructions, rather than with each other.

There were two successful groups, one structured and one not. The structured, successful group (Group 3), while returning to orientation, was able to maintain a path moving through the stages of solution development in a way that was indicative of reflective practice. Agreement played a relatively minor role, however, the presence of disagreement and the importance of solution suggestion may both be indicators of progress through the task, and an active discussion. The successful, unstructured group (Group 6) had a clear path through solution development, which was dependent on agreement within the group, rather than with the materials. Orientation was used as a central point after elements not involved in solution development. Implementation was related back to both these hubs, as well as allowing students to move on to a new problem.

Using the heuristics miner allowed us to further analyse two of the groups to determine whether our observations of the dependencies between elements were significant. In the two groups selected, the importance of the different roles of orientation, agreement and implementation were confirmed. In regards to the design of the inquiry, there were two sources of information that could have helped dyads to navigate through the virtual world: the workbook, and information designed into the world itself. It was apparent that the participants did not attend to the information available within the virtual environment. This identifies potential issues with the game design that would need to be addressed in future work. Participants could also ask the teacher for help, which was done on occasion. Redesigning the virtual world to support orientation activities and establish the rules of the environment, and some key collaboration skills, without necessarily guiding students through the inquiry should be considered in future iterations of this work.

6. Conclusions and Implications for Design

Multimodal discourse analysis enhanced our understanding of the processes of learning in a complex learning environment (one which depended on the use of tools, social interactions, and knowledge acquisition). By collecting multiple data sources, and conducting complementary analyses, we were able to understand the effects of the scaffolding on the different elements of the task. The analysis was time-consuming, however, the ultimate aim of this work is to identify indicators of successful and unsuccessful progress through a task so that the instructor, and the students, have a way of monitoring the processes of learning. One promising result was that the action taken after implementation may be an informative indicator of learners' progress through a task.

The regularity of orientation-type behaviour could be another useful indicator, and even the pattern of simple agreement or disagreement.

The analyses gave us insight into the behaviour of the learners in response to the design of the task – of the virtual world, the collaboration, and the role of the scaffolding. The results, which showed the impact of being “lost” in the virtual world on the collaborative processes, suggest that multiple layers of scaffolding are needed for complex tasks. The scaffolding provided to learners concentrated on knowledge acquisition, however, the collaborative processes as well as navigation around the virtual world proved to be challenging for students. Available time was another factor that should be taken into account when designing tasks in complex learning environments. Enough time needs to be given to accommodate substantial consultation with the materials as well as with other learners.

The future of the intersection of the learning sciences and technology calls for considered design based on our knowledge of the relationships between learners, tasks and tools (Thompson, Ashe, Carvalho et al., 2013). In addition to this careful design, we need ways of informing instructors of learners’ progress through tasks as well as their interactions with the tools, and with each other.

Acknowledgements

Part of this work was funded by the Australian Research Council grant FL100100203.

References

- Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., & Jin, M. (2008). The use of immersive virtual reality in the learning sciences: Digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*, 17(1), 102–141.
- Barab, S. A., Dodge, T., Ingram-Goble, A., Volk, C., Pepler, K., Pettyjohn, P., & Solomou, M. (2009). Pedagogical dramas and transformational play: Narratively-rich games for education. In I. A. Lurjel, N. Zagalo & P. Petta (Eds.), *Interactive storytelling* (pp. 332–335). Heidelberg Springer Berlin.
- Barab, S. A., Thomas, M., Dodge, T., Carteaux, R., & Tuzun, H. (2005). Making learning fun: Quest Atlantis, a game without guns. *Educational Technology, Research and Development*, 53(1), 86–107.
- Blikstein, P. (2013). Multimodal learning analytics. In D. Suthers & K. Verbert (Eds.), *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)* (pp. 102–106). NY: ACM.
- Borgman, C. L., Abelson, H., Dirks, L., Johnson, R., Koedinger, K. R., Linn, M. C., ... Szalay, A. (2008). *Fostering learning in the networked world: The cyber learning opportunity and challenge (A 21st century agenda for the National Science Foundation Report of the NSF Taskforce on Cyberlearning)*. Arlington: NSF Task Force on Cyberlearning.
- Clarke, J., Dede, C., Ketelhut, D. J., & Nelson, B. (2006). A design-based research strategy to promote scalability for educational innovations. *Educational Technology*, 46, 3(May-June), 27–36.

- Cox, R. (2007). Technology-enhanced research: Educational ICT systems as research instruments. *Technology, Pedagogy and Education*, 16(3), 337–356.
- Derry, S. J., Pea, R., Barron, B., Engle, R., Erickson, F., Goldman, R., ... Sherin, B. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology and ethics. *Journal of the Learning Sciences*, 19, 1–51.
- Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 1–19). Oxford: Elsevier.
- Dringus, L. P., & Ellis, T. (2005). Using data mining as a strategy for assessing asynchronous discussion forums. *Computers and Education*, 45(1), 141–160.
- Dyke, G., Lund, K., & Girardot, J.-J. (2009). Tatiana: An environment to support the CSCL analysis process. In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopolou, CSCL 2009, *Computer-supported collaborative learning practices Volume 1* (pp. 58–67). Rhodes, Greece: International Society of the Learning Sciences.
- Evans, M. A., Feenstra, E., Ryon, E., & McNeill, D. (2011). A multimodal approach to coding discourse: Collaboration, distributed cognition, and geometric reasoning. *International Journal of Computer-Supported Collaborative Learning*, 6(2), 253–278.
- Fern, X., Komireddy, C., Grigoreanu, V., & Burnett, M. (2010). Mining problem-solving strategies from HCI data. *ACM Trans. Computer-Human Interaction*, 17(1), 1–22.
- Goodyear, P., Jones, C., & Thompson, K. (2014). Computer-supported collaborative learning: Instructional approaches, group processes and educational designs. In J. M. Spector, M. D. Merrill, J. Elen & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology (4th Ed.)* (pp. 439–451). New York: Springer.
- Gress, C. L. Z., Fior, M., Hadwin, A. F., & Winne, P. H. (2010). Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior*, 26(5), 806–814.
- Halliday, M. A. K. (1994). *Introduction to functional grammar*. London: Arnold.
- Haythornthwaite, C., & de Laat, M. (2011). Social network informed design for learning with educational technology. In A. D. Olofsson & J. O. Lindberg (Eds.), *Informed design of educational technologies in higher education: Enhanced learning and teaching* (pp. 352–374). Hershey, PA: IGI Global.
- Jacobson, M., Miao, C., Kim, B., Shen, Z., & Chavez, M. (2008). Research into learning in an intelligent agent augmented multi-user virtual environment. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (pp. 348–351). United States: (IEEE) Institute of Electrical and Electronics Engineers.
- Jeong, H., Biswas, G., Johnson, J., & Howard, L. (2010). Analysis of productive learning behaviours in a structured inquiry cycle using hidden markov models. In R. S. J. d. Baker, A. Merceron & P. I. Jr. Pavlik (Eds.), *Proceedings of the 3rd International Conference on Educational Data Mining* (pp. 81–89). Pittsburgh, PA: LearnLab/Carnegie Learning.
- Kapur, M. (2011). Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *International Journal of Computer Supported Collaborative Learning*, 6(1), 39–56.
- Kay, J., Maisonneuve, N., Yacef, K., & Zaïane, O. (2006). Mining patterns of events in students' teamwork data. In *Online Proceedings of the ITS (Intelligent Tutoring Systems) 2006 Workshop on Educational Data Mining* (pp. 45–52). Jhongli, Taiwan: EDM.

- Keating, E., & Sunakawa, C. (2010). Participation cues: Coordinating activity and collaboration in complex online gaming worlds. *Language in Society*, 39(3), 331–356.
- Kennedy-Clark, S., & Thompson, K. (2011). What do students learn when collaboratively using a computer game in the study of historical disease epidemics, and why? *Games and Culture*, 6(6), 1–25.
- Kennedy-Clark, S., & Thompson, K. (2013a). Using multimodal discourse analysis to identify patterns of problem solving processes in a computer-supported collaborative environment. In F. Sudweeks & H. Lim, (Eds.), *Innovative methods and technologies for electronic discourse analysis* (pp. 94–118). Hershey, PA: IGI Global.
- Kennedy-Clark, S., & Thompson, K. (2013b). Collaborative Process Analysis Coding Scheme (CPACS): Examining the macro- and micro-level of students' discourse in a virtual world. *International Journal of Virtual and Personal Learning Environments*, 4(2), 19–49.
- Kennedy-Clark, S., Thompson, K., & Richards, D. (2011). Collaborative problem solving processes in a scenario-based multi-user environment. In H. Spada, G. Stahl, N. Miyake & N. Law (Eds.), *CSCL 2011, Computer-supported collaborative learning. Connecting research to policy and practice* (pp. 706–710). Hong Kong: International Society of the Learning Sciences.
- Ketelhut, D. J., Clarke, J., & Nelson, B. (2010). The development of River City, a multi-user virtual environment-based scientific inquiry curriculum: Historical and design evolutions. In M. J. Jacobson & P. Reimann (Eds.), *Designs for learning environments of the future* (pp. 89–110). New York: Springer Science + Business Media.
- Ketelhut, D. J., Dede, C., Clarke, J., & Nelson, B. (2006). A multi-user virtual environment for building higher order inquiry skills in science. *British Journal of Educational Technology*, 41(1), 56–68.
- Kummerow, D. (2012). The person that isn't: On defining the third-person, negatively and positively so. *Australian Journal of Linguistics*, 32(2), 259–290.
- Lin, F.-R., Hsieh, L.-S., & Chuang, F.-T. (2009). Discovering genres of online discussion threads via text mining. *Computers and Education*, 52(2), 481–495.
- Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an “early warning system” for educators: A proof of concept. *Computers and Education*, 54(2), 588–599.
- Nevile, M., & Rendle-Short, J. (2009). A conversation analysis view of communication as jointly accomplished social interaction: An unsuccessful proposal for a social visit. *Australian Journal of Linguistics*, 29(1), 75–89.
- Pechevnikiy, M., Tricka, N., Vasilyeva, E., van der Aalst, W. M. P., & Bra, P. M. E. De (2009). Process mining online assessment data. In T. Barnes, M. Desmarais, C. Romero & S. Venture (Eds.), *Proceedings of Educational Data Mining 2009: 2nd International Conference on Educational Data Mining* (pp. 279–288). Cordoba, Spain: International Working Group on Educational Mining.
- Poole, M. S., & Holmes, M. E. (1995). Decision development in computer-assisted group decision making. *Human Communication Research*, 22(1), 90–127.
- Poole, M. S., Andrew, H. V. d. V. H., Dooley, K., & Holmes, M. E. (2000). *Organizational change and innovation processes: Theory and methods for research*. Oxford: Oxford University Press.
- ProM (2011). *Version 5.2*. Retrieved January 21, 2014, from <https://sourceforge.net/projects/prom/files/ProM/5.2/ProM-5.2-Setup.exe/download>

- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257.
- Reimann, P., Frerejan, J., & Thompson, K. (2009). Using process mining to identify models of group decision making in chat data. In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopoulou (Eds.), *Proceedings of the 9th International Conference CSCL – Volume 1* (pp. 98–107). Rhodes, Greece: International Society of the Learning Sciences.
- Romero, A. C., & Ventura, S. (Eds.). (2006). *Data mining in e-learning*. Southampton: WITpress.
- Romero, P., Cox, R., Boulay, B. d., Lutz, R., & Bryant, S. (2007). A methodology for the capture and analysis of hybrid data: A case study of program debugging. *Behaviour Research Methods*, 39(2), 309–317.
- Rozinat, A., de Jong, I. S. M., Gunther, C. W., & van der Aalst, W. M. P. (2007). *Process mining of test processes: A case study* (BETA Working Paper Series, WP 220). Eindhoven: Eindhoven University of Technology.
- Schulte-Mecklenbeck, M., Kuehberger, A., & Ranyard, R. (Eds.). (2011). *A handbook of process tracing methods for decision research: A critical review and users guide*. New York: Psychology Press.
- Shaffer, D. W., & Gee, J. P. (2012). The right kind of GATE: Computer games and the future of assessment. In M. Mayrath, D. Robinson & J. Clarke-Midura (Eds.), *Technology-based assessments for 21st Century Skills: Theoretical and practical implications from modern research* (pp. 211–228). Charlotte, NC: Information Age Publications.
- Southavilay, V., Yacef, K., & Calvo, R. A. (2010). Analysis of collaborative writing processes using hidden Markov models and semantic heuristics. In W. Fan, W. Hsu, G. I. Webb, B. Liu, C. Zhang, D. Gunopulos & X. Wu (Eds.), *Proceedings of the IEEE International Conference on Data Mining Workshops* (pp. 543–548). CA: IEEE CPS.
- Squire, K. D., Barnett, M., Grant, J. M., & Higginbottom, T. (2004). Electromagnetism Supercharged! Learning physics with digital simulation games. In Y. B. Kafai, W. A. Sandoval & N. Enyedy (Eds.), *Proceedings of the Sixth International Conference of the Learning Sciences* (pp. 513–520). Mahwah, NJ: Lawrence Erlbaum.
- Steinkuehler, C. A. (2006). Massively multiplayer online ideogaming as participation in a discourse. *Mind, Culture, and Activity*, 13(1), 38–52.
- Strijbos, J.-W., & Fischer, F. (2007). Methodological challenges for collaborative learning research. *Learning and Instruction*, 17(4), 389–393.
- Thompson, K., Ashe, D., Carvalho, L., Goodyear, P., Kelly, N., & Parisio, M. L. (2013). Processing and visualizing data in complex learning environments. *American Behavioural Scientist*, 57(10), 1401–1420.
- Thompson, K., Ashe, D., Wardek, D., Yeoman, P., & Parisio, M. (2013). Identification of patterns of tool use and sketching practices in a learning by design project. In N. Rummel, M. Kapur, M. Nathan & S. Puntambekar (Eds.), *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning CSCL2013* (pp. 478–485). Madison, WI: International Society of the Learning Sciences.
- Thompson, K., Ashe, D., Yeoman, P., & Parisio, M. (2013). Phases of design: Following idea development and patterns of collaborative discussion in a learning by design project. In N. Rummel, M. Kapur, M. Nathan & S. Puntambekar (Eds.), *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning CSCL2013* (pp. 494–501). Madison, WI: International Society of the Learning Sciences.

- Thompson, K., & Kelly, N. (2012). Processes of decision-making with adaptive combinations of wiki and chat tools. In J. van Aalst, K. Thompson, M. J. Jacobson & P. Reimann (Eds.), *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012) – Volume 2* (pp.459–466). Sydney, Australia: International Society of the Learning Sciences.
- Thompson, K., Kennedy-Clark, S., Kelly, N., & Wheeler, P. (2013). Using automated and fine-grained analysis of pronoun use as indicators of progress in an online collaborative project. Computer-Supported Collaborative Learning. In N. Rummel, M. Kapur, M. Nathan & S. Puntambekar (Eds.), *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning CSCL2013* (pp. 486–493). Madison, WI: International Society of the Learning Sciences.
- Thompson, K., Kennedy-Clark, S., Markauskaite, L., & Southavilay, V. C. (2011). Capturing and analysing the processes and patterns of learning in collaborative learning environments. In H. Spada, G. Stahl, N. Miyake & N. Law, *Connecting Research to Policy: Proceedings of Computer Supported Collaborative Learning CSCL 2011* (pp. 596–600). Hong Kong: International Society of the Learning Sciences.
- Thompson, K., Kennedy-Clark, S., Wheeler, P., & Kelly, N. (2014). Discovering indicators of successful collaboration using tense: Automated extraction of patterns in discourse. *British Journal of Educational Technology*, 45(3), 461–470.
- Thompson, K., & Reimann, P. (2010). Patterns of use of an agent-based model and a system dynamics model: The application of patterns of use and the impacts on learning outcomes. *Computers and Education*, 54(2), 392–403.
- Trčka, N., Pechenizkiy, M., & van der Aalst, W. (2010). Process mining from educational data. In C. Romero, S. Ventura, M. Pechenizkiy & R. J. D. Baker (Eds.), *Handbook of educational data mining* (pp. 123–142). Boca Raton: CRC Press.
- van der Aalst, W.M.P. (2011). *Process mining: Discovery, conformance and enhancement of business processes*. Berlin: Springer.
- Watson, W. R., Mong, C. J., & Harris, C. A. (2011). A case study of the in-class use of a video game for teaching high school history. *Computers and Education*, 56, 466–474.
- Weijters, A. J. M. M., & Ribeiro, J. T. S. (2010). *Flexible Heuristics Miner (FHM)* (BETA Working Paper Series 334). Eindhoven: Eindhoven University of Technology.
- Weijters, A. J. M. M., Van der Aalst, W. M. P., & De Medeiros, A. K. A. (2006). *Process mining with the heuristics mineralgorithm*. (BETA Working Paper Series 166). Eindhoven: Eindhoven University of Technology.
- Wise, A., & Chiu, M. M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *International Journal of Computer-Supported Collaborative Learning*, 6(3), 445–470.