

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

Using Multimodal Learning Analytics to Model Student Behaviour: A Systematic Analysis of Behavioural Framing

Alejandro Andrade

Indiana University, USA

laandrad@indiana.edu

Ginette Delandshere

Indiana University, USA

Joshua A. Danish

Indiana University, USA

ABSTRACT: One of the challenges many learning scientists face is the laborious task of coding large amounts of video data and consistently identifying social actions, which is time consuming and difficult to accomplish in a systematic and consistent manner. It is easier to catalog observable behaviours (e.g., body motions or gaze) without explicitly attempting to identify their social relevance for the participants. We explore the potential for using a multimodal learning analytic approach to identify whether clusters of observable behaviours can be used to identify and characterize behavioural frames in rich video data of student interviews. We argue that by conducting a systematic analysis of behavioural frames using computerized algorithms we can model student frames as a latent class variable. We explore whether those behavioural frames overlap in productive ways with epistemological frames, thus supporting our efforts to interpret rich video data. We believe that a positive feedback loop between methodological approaches and theory will emerge as we further our understanding of framing by developing analytical models leveraged by multimodal learning analytics.

Keywords: Epistemological framing, multimodal learning analytics, hidden Markov models, interviews

Many learning scientists attempt to reduce large corpuses of video data productively by observing, coding, and identifying clusters of learner actions from which they might be able to draw meaningful inferences about learning and cognition. One of the challenges they face is the laborious task of coding large amounts of video data through repeated viewings. This coding is particularly challenging given the socially situated nature of these actions, which require careful viewing to make inferences about what the activities actually mean for the participants (Danish, Enyedy, & Parnafes, 2015; Jordan & Henderson, 1995). Identifying social actions consistently is time consuming and difficult to accomplish in a systematic manner. It is much easier to catalog observable behaviours (e.g., body motions or gaze) without explicitly attempting to identify their social relevance for the participants. While these observable behaviours are not the only way in which participants communicate, they do provide

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

observable cues, which participants pick up on to help determine the social context as they respond to learning opportunities. Given this blend of substantive and methodological issues, we investigate whether these low-level and easier-to-code behaviours allow us to predict aspects of the learning situation more easily for further exploration.

We explore statistical models and data representations that facilitate the identification of behavioural clusters, which we refer to as behavioural frames, and explore how these frames change during a specific activity to enhance our understanding of the relationship between observed behaviours and learning. Specifically, we explore whether behavioural frames can be identified and related to cognition and learning in rich video data of student interviews using a multimodal learning analytic approach (Blikstein, 2013). Prior examinations of social frames that students attend to (Goffman, 1974) require having multiple researchers code all of the student behaviours and then attempt to interpret student frames through repeated watching of the video data (c.f., Russ, Lee, & Sherin 2012, who focused on epistemological frames). Our goal in the present paper was to extend this approach in a manner that would more consistently map observed behaviours to frames, and potentially find shifts in frames at a finer-grained level while adding the ability to report the statistical validity of the approach.

We ground this exploratory analysis in a semi-structured interview context, which constrains the range of relevant social interactions, thus allowing us to explore the feasibility of this approach. To make sense of the interview context, we attempt to relate our behavioural frames to the “epistemological frames” identified by Russ et al. (2012) in interview contexts as these researchers have productively recognized frames as a method for identifying key learning opportunities. Russ et al. (2012) have shown that the ways that students interpret an activity, such as a cognitive interview, will shape how students respond, explore, and learn within the activity. In a given situation during a learning activity, student expectations are of particular importance for how students engage with knowledge. These expectations, called “epistemological frames” (Scherr & Hammer, 2009), have been shown to influence how students respond to a question when they do not know the answer. For example, do they simply state that they do not know, or do they actively explore possibilities and formulate an answer?

Typically, student epistemologies are measured through self-reporting, but because epistemological stance relates to the contexts in which students are learning, self-reports do not always provide accurate information (Scherr & Hammer, 2009). Therefore, epistemological frame analysis has been developed to analyze this contextual dependency of student expectations about knowledge. This method studies the interaction between students’ framing, behaviour, and the content of their speech. It has been used to describe the relationship between student frames and the way they engage with learning activities from middle school through the college level.

In the remainder of this paper, we explore the potential of multimodal learning analytic approaches for identifying behavioural frames, and then explore whether those frames overlap in productive ways with epistemological frames, thus supporting our efforts to interpret rich video data. We then discuss how

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

the analysis of behavioural frames may apply more broadly to the analysis of interactions in learning contexts and the identification of important yet difficult to identify patterns that reach beyond epistemological frames.

1 BACKGROUND

1.1 Framing and Student Behaviour

Research on framing and its impact on social interaction was initially introduced and studied in the context of sociology (Goffman, 1974), anthropology (Bateson, 1955; Tannen, 1993), media communication (Scheufele, 1999), organizations (Spybey, 1984), and cognitive discourse analysis (van Dijk, 1977). These studies have shown that, in the moment-to-moment interaction, people's social expectations are implicit but visible to other participants (Kendon, 1990). These social expectations are referred to as frames, and as participants interact and respond to their interlocutor's speech and action they continually provide cues to their understanding of the frame. These contextualized cues are frequently produced at a *meta-level* of communication that goes beyond denotative aspects of the message, and therefore remains mostly implicit. For instance, when two people play at sword fighting they might say, "I'm going to hurt you" and yet neither person expects to actually get hurt. Rather, because their actions are framed as play, they can make these statements while assuming that the other person recognizes them as part of the game rather than as a sincere threat. These structures of expectations are revealed in abstract elements of the message, and need be interpreted from a combination of verbal and non-verbal behaviours, such as body positioning, eye gaze, gestures, and other prosodic elements of language (e.g., speech volume and pace).

More recently, the term *epistemological framing* has referred to the contextual, subjective expectation as to what kind of knowledge seems to be worth activating in the moment-by-moment interaction. Studies in epistemological framing have provided evidence of strategies students use during learning activities in which certain actions lead to more sophisticated understanding than others (Berland & Hammer, 2012; Elby & Hammer, 2010; Hutchison & Hammer, 2010). To capture student framing, researchers use video-recorded data as their main source of information (Conlin, Gupta, Scherr, & Hammer, 2008; Russ, Scherr, Hammer, & Mikeska, 2008; Scherr & Hammer, 2009). These studies mainly rely on in-depth qualitative analyses of student interactions in classroom tutorials or interview settings. Building on the tradition of interaction analysis (Jordan & Henderson, 1995), researchers iteratively examine short segments of video, trying to identify different patterns of co-occurring behaviours. The goal of the analyst, as described by Scherr and Hammer (2009), is to find a finite number of clusters of behaviours that reflect the implicit *structures of expectations* that belong to those particular social situations.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

1.2 Framing and Interviews

In psychology as well as in educational research, researchers have made extensive use of interviews to gather information about student knowledge, cognition, and reasoning (Clement, 2000; diSessa, 2007; Ginsburg, 1997). An interview is an unfamiliar setting that may affect student performance; therefore, recent research has foregrounded the necessity to pay attention to the ways in which students interact with the interviewer and engage with the content of the interview. Because interviews provide a delineated, one-on-one interaction between the experimenter and the student, the underlying structures of expectations are thought to be limited to just a few. Young students, however, may enter the interview room without a clear set of expectations of what the situation entails (Russ et al., 2012).

Student expectations depend on previous experience, but may shift throughout the course of the interview in response to various cues (intentional or unintentional) provided by the interviewer. After carefully examining several interviews with middle school students, Russ et al. (2012) argue that students perceived at least three types of epistemological frames: a) the inquiry frame, b) the expert frame, and c) the oral examination frame. In Table 1, we summarize the character of these frames as well as their characteristic behaviours. In the inquiry frame, the behaviours are characterized by long pauses in speech, restarting during explanations, little eye contact, and prolific gesturing, hinting that their responses are a sort of sense-making elaboration. In the expert frame, the student shows prolific gesturing, eye contact, use of colloquial terminology, and a lack of hesitation, indicating that their talk reflects their prior knowledge. The oral examination frame is characterized by eye contact, limited use of gesture, use of scientific vocabulary, and lack of hedging language, which indicates that the student may be interpreting the situation as though they ought to produce “correct,” wordy answers. These frames are mostly implicit, unarticulated expectations carried out by the student from previous experiences or related situations.

Table 1: Framing in clinical interviews (summarized from Russ et al., 2012).

Framing	Description	Behaviours
Inquiry	“When faced with a question they cannot immediately answer, students may choose to frame the clinical interview activity as one in which they should engage in <i>inquiry</i> to construct an explanation. Rather than saying ‘I don’t know’ students may attempt to figure out an appropriate answer in the moment of the interview.” (p. 582)	Long pauses in speech Restarts during explanations Little eye contact Prolific gesturing
Oral Exam	“At other times in the interview, students may instead frame the activity as an <i>oral examination</i> . In general, during oral examinations, students are expected to produce a desired response in a clear and concise fashion. We find evidence of students adopting this approach to the interview both when students know what they perceive to be the desired response and when they do not know it.” (p. 586)	Lack of hedging language Eye contact Limited use of gesture Use of scientific vocabulary

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

Framing	Description	Behaviours
Expert	“In contrast to the <i>inquiry</i> and <i>oral examination</i> frames, students often adopt a framing in which they take their task to be that of discussing their own thinking, on which they are the experts, and that is relatively unproblematic for them.” (p. 587)	Lack of hesitation Eye contact Prolific gesturing Use of colloquial terminology

Russ et al.’s (2012) study offers an in-depth qualitative account of the regularities observed in student non-verbal behaviour in clinical interviews, providing a promising interpretation framework as to why these common behaviours cluster together. According to Russ et al. (2012), these behaviours co-occur because they are instantiations of the student framing, which implicitly reveals their expectations of the specific moment-to-moment interactions between student and interviewer. Although other researchers may reach different conclusions, for instance, by paying attention to different types of behaviours, our goal is not to provide a thorough exploration of frame analysis. For instance, we do not address whether these particular frames are the most useful, and at this point in time we can only speculate about how they develop over time, or what sorts of experiences students had that resulted in these sets of expectations. By providing a systematic analysis of framing, however, we aim to support future research about framing and how it affects student performance during interviews.

In this study (Danish, Saleh, Andrade, & Bryan, accepted) we used semi-structured interviews, instead of open-ended clinical interviews, as post-intervention assessments to elicit and document naturalistic forms of thinking. Several differences exist between these two types of interviews. Clinical interviews allow for the exploration of all facets of student knowledge. The interactions within a clinical interview are loosely structured and exhibit an improvisational character where the interviewer is responsive to the emerging thinking of the interviewee (diSessa, 2007). They typically begin with the interviewer briefly describing the activity and asking the student to complete a task or answer questions about a phenomenon of interest (Clement, 2000). In comparison, semi-structured interviews (Drever, 1995; Whiting, 2008) are intended to solicit very specific elements of knowledge. In this case, the student–interviewer interaction was much more scripted and the interviewer followed a protocol of pre-defined questions. Nonetheless, students are most likely unfamiliar with both types of interviews, and we would expect to find a similar framing in our data.

Before we elaborate on our proposed systematic analysis of behaviours, we describe the relationship between student framing and the nature of their reasoning. Highlighting the link is important because the usefulness of behavioural framing as a construct in the learning sciences relies on its relationship to other important constructs, student reasoning in particular.

1.3 Relationship between Framing and Mechanistic Reasoning

Researchers studying epistemological framing argue that, within a given context, student framing influences the nature of student reasoning (Russ et al., 2012). Within the interview context, some

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

frames, such as the inquiry frame, are assumed not only to make in-depth reasoning visible, but may even promote that kind of reasoning. In contrast, other frames such as the examination frame are presumed to encourage brief answers that quickly summarize recall with little depth. In the context of interviews designed to elicit student understanding of scientific concepts, researchers have found that these different frames are also related to whether or not students are likely to engage in valuable aspects of mechanistic reasoning.

According to Russ et al. (2008), knowledge about causal relationships among the elements of a system is referred to as *mechanistic reasoning*. One of the most revealing forms of this type of reasoning is students being able to articulate how present or future element configurations can be deduced from the relationships among elements in past states. This idea is referred to as “chaining.” In measuring mechanistic reasoning in student utterances, researchers look for instantiations of chaining and other related elements (Russ et al., 2008). For instance, Conlin et al. (2008) sought to find a relationship between student framing and mechanistic reasoning. The authors argue that the frames they studied correlated with certain ways of reasoning. In studying college physics tutorials, they found that student behaviours tend to form four types of clusters that represent distinct frames. One of these frames showed students using prolific gesturing, animated tone and face, sitting up straight, engaging in eye contact, and providing clear utterances. This “animated and prolific talk” frame was associated with good collaboration because the students’ structure of expectations was to openly discuss and comment on each other’s ideas. Conlin et al. (2008) also found that most of the productive reasoning, namely *mechanistic reasoning*, took place during those moments where the group was adopting this frame.

In our project, we also focused on how student framing during the individual interviews co-occurred with instantiations of mechanistic reasoning, as inferred from student accounts (i.e., how students understood how bees collect nectar and communicate relevant information about where the flower is by doing a special dance). In the following section, we describe our analytic technique that engages with video data in a much finer-grained way than previous analyses of epistemological framing have attempted. This technique relates to the use of *multimodal learning analytics* (Blikstein, 2013; Oviatt, Cohen, & Weibel, 2013; Worsley, 2012), which systematically analyze various modalities of student behaviour, such as gesture, gaze, and speech.

1.4 Multimodal Learning Analytics

Because there are always more data than researchers can actually analyze, techniques originally used in data mining are currently of growing interest to researchers in the learning sciences (Martin & Sherin, 2013). Specifically, classification algorithms, machine learning, and other statistical techniques such as clustering, support vector machine models, computerized text analysis and social network analysis, are gaining traction as ways of providing automated detection and convergent support to qualitative analysis of cognition, discourse, and interaction (Berland, Baker, & Blikstein, 2014; Berland, Martin, Benton, Petrick Smith, & Davis, 2013; Sherin, 2013; Worsley & Blikstein, 2014). Recent applications of

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

computational techniques for the study of learning have also gone beyond text- or log-based interaction. These applications include the collection of multimodal information in human activity through new data-capturing methods and sensing technologies usually referred to as *Multimodal Learning Analytics* (MLA; Blikstein, 2013).

Recent studies have explored the relationship between non-verbal behaviours and learning through computational approaches. For instance, Worsley and Blikstein (2014) developed a heuristic approach to analyze multimodal information of student activities in engineering designs. The authors studied engineering practices by looking at how college students manipulated objects in an open-ended, hands-on design task. Ochoa and collaborators (Echeverría, Avendaño, Chiluiza, Vásquez, & Ochoa, 2014; Luzardo, Guamán, Chiluiza, Castells, & Ochoa, 2014; Ochoa et al., 2013) differentiate expert high school students' mathematical and presentation skills using classifiers from multimodal features automatically extracted from video and audio data. Shoukry, Göbel, and Steinmetz (2014) used smartphone camera and tracking sensors to extract and analyze automated information about students' eye gaze, facial features, emotion, touch interactions, and data usage to understand ways of enabling non-invasive naturalistic assessment of learning experience. Many more studies have employed eye tracking, facial emotions, and skin sensors to make inferences about student learning (Gomes, Yassine, Worsley, & Blikstein, 2013; Worsley, 2012).

We argue that by conducting a systematic analysis of behavioural frames using computerized algorithms, we can model student epistemological frames as a latent class variable. These representations could then provide us with additional insights into how frames are related to student learning strategies, attitudes, and performance. In the following sections, we provide a description of the conceptual model and the statistical analysis we used to test it.

1.5 A Conceptual Model of Framing as a Latent Variable

In the model shown in Figure 1, *epistemological frames* (at the top) are represented as latent classes — unobservable properties of the student's interaction. Because the context (e.g., interview and interviewer) and the student's past history of experiences limit the types of possible expectations, the student's framing will likely only take on a few distinct frames (e.g., three frames are displayed because this is what we expect based on Russ et al., 2012). In their realization through time, these epistemological frames — the student's underlying expectations — are expressed in the form of clusters of verbal and non-verbal behaviours (at the bottom) — our present analysis only focuses on the non-verbal behaviours to see what patterns in students' frames might become visible without resorting to their discourse. This is the observable part of the model, where we *measure* the behavioural frames (clusters of behaviours). The behavioural frames are estimated using a statistical model (i.e., a latent variable model) based on a set of observable behaviours (i.e., body positioning, gaze, gesture, hedging, and speech prosodic). Note that the relationship between epistemological frames and clusters of behaviours is assumed to be exclusive, in that a cluster is the expression of one and only one frame (e.g.,

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

the cluster in segment 1 stems only from frame 1, likewise the cluster in the second segment stems from frame 2, and so on). This is an assumption of the statistical model we used.

A theoretical caveat, however, is in order. Framing emerges from the participants’ moment-to-moment meta-communicative actions and some may infer that our model treats student actions as though they take place in isolation, independent of the interviewer or the context. However, we presume that student behaviour demonstrates how they are continually responding to the interviewer and context, and thus it is possible to treat these behavioural frames as approximations for how the student is interacting with the interviewer and the context. If successful, future research may extend our model to represent the interactions between the participants and their context more accurately while also continuing to identify those features of the interaction not consistently captured in such a behavioural model. For instance, the unidirectional arrows from frames to clusters do not imply that behaviours do not affect frames necessarily, but that our current analysis only explores this relationship in one direction. In future, we envision more complex models, including instructor behaviour as a covariate to account for the interviewer–student interaction.

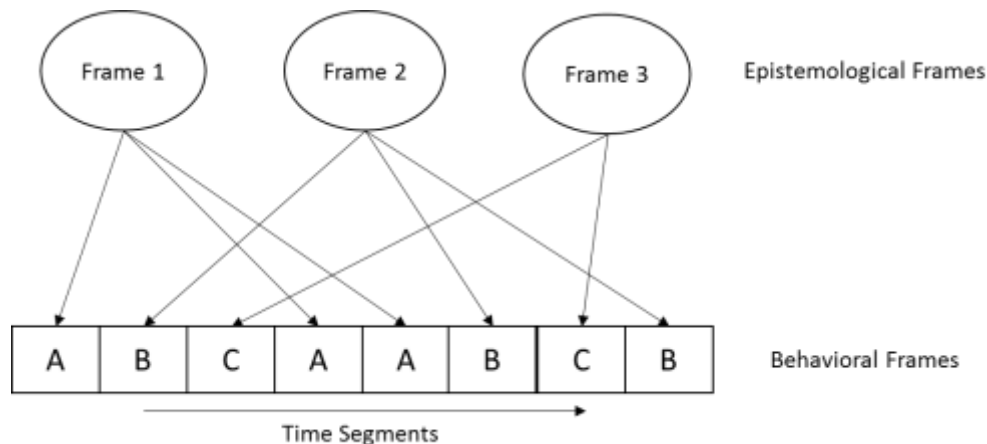


Figure 1: A conceptual model of student epistemological framing.

In order to understand the association between student behavioural frames and reasoning, we explored this relationship in two ways. First, we considered the simple association between each frame and mechanistic reasoning at the aggregate level — i.e., association indices between frequency of frames and mechanistic reasoning scores. Second, we examined frame transitions during the interview. The transition revealed different student typologies or profiles. We hypothesized that the way in which students transitioned from one frame to the other throughout the interview may be related to the depth of their engagement with the content. For instance, some students might start expecting an oral evaluation, but later in the interview begin exploring the space of the problem. Conversely, other students might start engaging with the content, but later feel like providing short answers to the questions.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

In what follows, we introduce our analysis of epistemological framing in semi-structured interviews. Some empirical questions that guided our enquiry were these: What non-verbal behaviours form clusters (behavioural frames)? How do these identified behavioural frames relate to student mechanistic reasoning? And how consistent are these frames with the theoretical definition of epistemological framing and previously identified epistemological frames?

2 METHODS

2.1 Participants, Materials, and Procedure

The analysis that follows is based on a secondary analysis of data collected in the spring of 2012 (Danish et al., accepted). Thirty first- and second-graders (6–7 years old, M=15, F=15) in a Midwestern American elementary school took part in this study and were randomly assigned to either experimental or control conditions. Students in the experimental condition engaged in inquiry about how honeybees collect nectar using a version of the BeeSign software (Danish, 2014) with integrated software scaffolds that help guide the teacher-led inquiry into topics related to how honeybees behave as a system (Danish et al., accepted). All students participated in 30–40-minute instructional activities as a replacement for their science activities for that day. Individual semi-structured interviews were conducted as a post-test to obtain evidence of the students' understanding of complex systems. During the interview, each child answered nine questions about the behaviour of bees and ants gathering food. The questions were asked with the support of pictures (in the case of the bees) and NetLogo (Wilensky, 1999) animations (in the case of the ants). For instance, children saw the picture of a beehive and a flower with nectar and were asked to explain what they thought the bees would do. The average interview was about 12 minutes long.

2.2 Sources of Data and Analysis of Features

While we believe that technology will one day identify behavioural features automatically by means of computer vision, the current analysis was conducted based on researcher-coded behaviours. The coding started by identifying distinct levels of body language, eye gaze, gesture, hedging language, and speech prosodic in students' non-verbal behaviours (see Figure 2). Then, each interview excerpt was divided into a sequence of 10-second intervals. Although the time interval coding has its pros and cons, as it may lack the accuracy of directly recording the onset and offset of events, it is a very convenient way to synchronize the concurrent recording of separate streams of behaviours (Bakeman & Gottman, 1997). Our prior work (Andrade-Lotero, Danish, Moreno, & Perez, 2013) suggests that 10-second intervals are long enough to include sufficient information about the behaviour, yet short enough to provide an adequate number of discrete data points for statistical analyses in this type of short, semi-structured interview format.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>



Body: Forward, Straight, Backward
Gaze: On paper, Interviewer, Away
Gesture: Fidgeting, Gesturing, No gesture
Hedging Language: Yes, No
Speech: Loud, Neutral, Soft

Figure 2: Relevant behaviours and corresponding levels.

To account for the local contingencies in the interview setting, our approach to coding incorporated two important characteristics. First, each behaviour is represented by a categorical variable in which the levels are coded in mutually exclusive and exhaustive ways. For instance, the gaze of a student can be coded as “making eye contact,” “looking at the materials,” or “looking away.” Every category is exclusive of the other two because if the student is looking at the materials she cannot be making eye contact at the same point in time, and all three are exhaustive because there are no other possible directions in which the student can direct their gaze. Note that “looking away” is an all-encompassing category that captures whatever is not classified in the first or second categories. Second, these behaviours, sampled over time, each take the form of a *time-series categorical variable* that follows a multinomial distribution. Therefore, statistical models that deal with multivariate, time-dependent, multinomial distributions can be used to represent the probabilities associated with the occurrence of these behaviours over repeated events.

The resulting data matrix contains 1189 observations and seven variables (subject ID, time, and five behaviours). To optimize the use of these data, we sampled only those 10-second intervals in which the child was speaking. We had two reasons: a) two of the identified features (i.e., hedging language and speech prosody) relate to verbal communication, and b) we wanted to look for reasoning in their talk so this limited us to those frames that could be analyzed in both modalities. The resulting data matrix contains 569 observations.

2.3 Supervised vs. Unsupervised Algorithms

To classify these 10-second intervals, two distinct approaches can be taken. In the machine learning literature (Lantz, 2013; Pedregosa et al., 2011) these approaches are referred to as supervised and unsupervised methods. Supervised approaches are classification algorithms that make use of prior information about the classification variable (for a good review see e.g., Jayaprakash, Moody, Lauría, Regan, & Baron, 2014). An example of a supervised method is predicting whether a hand movement is

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

an iconic gesture or fidgeting, based on a set of annotated still images called the *training set*. Logistic regression, support vector machines, naïve Bayes, neural networks, and decision trees are some examples of supervised algorithms. These algorithms then use characteristic features of the behaviour — i.e., the amplitude of the hand movement, angular distance of the joints, the regularity of the movement — to assign classification probabilities to each case of belonging to either iconic gesturing or fidgeting.

On the other hand, classification algorithms that do not have *a priori* information of what the resulting classes should look like are regarded as unsupervised approaches. For instance, one may want to form clusters of cases that share similar characteristics (e.g., clusters of hand movement behaviours). K-means clustering, agglomerative hierarchical clustering, and model-based clustering are some examples of unsupervised algorithms. We decided on an unsupervised approach for two reasons. First, a statistical reason was the relatively small sample size, which was not sufficient to yield a reliable supervised approach solution. Second, a practical reason was the difficulty of defining the training set from which to infer the relevant classes. A supervised algorithm would have required a training set with the annotated frames and an additional reliability analysis given that the identification of the behavioural transitions can be problematic. In a previous study (Andrade-Lotero, Delandshere, & Danish, 2014), we tried a supervised approach to predict epistemological frames using two raters to identify behavioural transitions for nine students and labelling the segments between transitions using Russ et al. (2012) frames. Although we found that the non-verbal behaviours accounted for about 40% of the frame variability, the training set was too small to provide good interrater reliability (Kappa less than .4).

One additional consideration with our data was the dependency of observations. Because observations that are close in time tend to be similar, they are not time independent. For instance, observations taken at time 2:30 are more likely to be similar to times 2:20 or 2:40 than to 1:10. Hidden Markov Models (Zucchini & MacDonald, 2009) are a class of statistical models that can deal with dependency of observations. We describe our chosen Markov model in greater detail below.

2.4 Unsupervised Analysis Using a Hidden Markov Model

Hidden Markov Models (HMMs), also referred to as Latent Markov Models, are a family of statistical analyses designed to deal with time-series (for application examples of HMMs with discrete or continuous data, univariate and multivariate analysis, with and without covariates, and with and without prior information see, for instance, Visser & Speekenbrink, 2010; Zucchini & MacDonald, 2009). These models have been popular in several disciplines, including sociology, psychology, economics, environmental sciences, genetics, speech recognition, and engineering. HMMs assume that the data are generated by a finite set of mixtures that represent different states, or *latent classes*, responsible for the production of the observable behaviours. Note that in our case this is an unsupervised method because the algorithm does not make use of any *a priori* classification. In HMMs, because the observations are time-dependent, a transition matrix is estimated. The transition matrix is a model of the observed

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

transitions from one state to another. Thus, the transition matrix provides information about the sequence of states.

Because the number of behavioural states is unknown, various models can fit the data. To support a decision with respect to the number of states to extract, some statistical fit indices have been developed. The Bayesian Information Criterion (BIC), used by the *depmixS4* package in *R* (Visser & Speekenbrink, 2010), provides support as to the number of clusters to be extracted — the lower the BIC the better the model fits the data. In addition, an analysis of the uncertainty of the model’s goodness-of-fit can be obtained by analyzing the distribution of the pseudo-residuals.

3 RESULTS

3.1 Behavioural Frame Analysis

There appear to be three distinct latent classes that yield three distinct behavioural frames in our current data set. Six HMMs with 2 to 7 states were fit to the data, and a 3-state model was selected because it fit the data best (2-state BIC = 4887.85, 3-state BIC = 4746.58, 4-state BIC = 4761.59; thus a 3-state has an acceptable pseudo-residual distribution). Table 2 provides the central tendency and dispersion of the distribution of the multimodal behaviours across students. For instance, it is apparent that body straight, eyes on paper, gesturing, no hedging language, and neutral speech were the most prevalent behaviours. The proportion of the behaviours across states is presented in Table 3. Note that the proportion of the behavioural levels add to one in each frame. For instance, the proportions for body backward (.66), forward (.05), and straight (.29) add to one within the hesitant behavioural frame, which is a property of the statistical model we used, and thus larger probabilities carry more weight in defining the latent state. Note, however, that the overall combination of behaviours produces the frames, not just individual behaviours. An examination of the cluster components led us to interpret the first state as a hesitant attitude — the body is backward, soft voice, no eye contact, hedging, and fidgeting. The second state indicates a calm attitude — the body straight, no hedging or gesturing, neutral voice, and eye contact with the interviewer. The third state indicates an active attitude — body forward, gaze on the materials, gesturing, loud voice, and no hedging.

Table 2: Distribution of multimodal behaviours averages (standard deviation).

Body		
Backward	Forward	Straight
26% (30%)	25% (16%)	49% (21%)
Gaze		
Eye contact	No eye contact	On paper
18% (17%)	27% (11%)	55% (12%)

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

Gesture

Fidgeting	Gesturing	No gesture
25% (10%)	41% (16%)	34% (18%)

Hedging

Yes	No
18% (16%)	82% (16%)

Speech

Loud	Neutral	Soft
23% (18%)	51% (14%)	26% (17%)

Table 3: Behavioural frame probabilities.

Observed Behaviour	Behavioural Frame 1 (N=172, 30.22%) Hesitant	Behavioural Frame 2 (N=170, 29.88%) Calm	Behavioural Frame3 (N=227, 39.89%) Active
--------------------	--	--	---

Body

Backward	0.66	0.04	0.03
Forward	0.05	0.13	0.58
Straight	0.29	0.83	0.39

Gaze

Eye contact	0.12	0.40	0.01
Away	0.45	0.37	0.06
On paper	0.44	0.23	0.93

Gesture

Fidgeting	0.43	0.14	0.18
Gesturing	0.18	0.29	0.77
No gesture	0.39	0.57	0.05

Hedging

Yes	0.43	0.06	0.02
No	0.57	0.94	0.98

Speech

Loud	0.03	0.39	0.35
Neutral	0.46	0.53	0.50
Soft	0.51	0.08	0.15

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

We used the algorithm to predict the behavioural frames in each interview and then examined these frames across different students. Although the actual display is idiosyncratic for each student, we noted that the algorithm consistently identifies similar poses regardless of the student. Put differently, although the fidgets identified by the human coders look different, the co-occurring behaviours around the fidgets are consistent across students. Figure 3 shows that students tend to express their hesitation through a combination of the following behaviours: body backwards or straight, gaze on paper or away, fidgeting, hedging language, and soft voice.



Figure 3. Examples of the Hesitant behavioural frame. Body backwards or straight, gaze on paper or away, fidgeting, hedging language, soft voice.

Figure 4 illustrates the Calm state. Students are sitting or standing straight, their gaze is away or on the interviewer, not gesturing, not hedging, and using a neutral or loud voice. Note that two of these students briefly put their hands together as in a short gesture that was neither fidgeting nor referred to the printed materials, but instead seemed to accompany their speech. Due to the brevity of these hand movements, they either were not identified in the 10-second intervals, or coded as “gesturing.”

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>



Figure 4. Examples of the Calm behavioural frame. Body straight, gaze away or on interviewer, not gesturing, not hedging language, neutral or loud voice.

Figure 5 shows when the students were more active. Students were leaning forward or sitting straight, their gaze was on the paper, they were gesturing prolifically, did not use hedging language, and their voice was neutral or soft. The kinds of gestures used varied between simply pointing at a detail in the picture to more iconic or metaphorical gestures.



Figure 5: The Active behavioural frame. Body forwards or straight, gaze on paper, gesturing, not hedging language, voice neutral or soft.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

3.2 Identifying Behavioural Frame Transitions

A main goal of frame analysis is to identify transition instances, we used the 3-cluster HMM algorithm to predict the behavioural frames in an interview as a case study. For this case study, we examined whether these frame transitions corresponded to changes in student expectations as evidenced by verbal exchanges between the interviewer and interviewee. Table 4 shows the data frame that the algorithm used to predict the behavioural frames for student 19. In the following excerpt, we reproduce the interviewer–interviewee exchanges presented in Table 4.

Table 4. Data matrix for student 19. Onset and offset interval times, observed behaviours, and predicted frames are shown.

IN	OUT	BODY	GAZE	GESTURE	HEDGING	SPEECH	FRAME
04:47.1	04:57.0	Straight	Paper	Gesture	No	Neutral	Calm
04:57.0	05:07.0	Backward	Away	Fidgeting	Yes	Soft	Hesitant
05:07.1	05:17.0	Forward	Paper	Gesture	No	Neutral	Active
05:17.1	05:27.0	Straight	Away	Gesture	No	Soft	Active
05:27.0	05:37.0	Forward	Paper	Gesture	No	Neutral	Active

04:47.0 Interviewer: Look at this picture. If you saw the bees flying from the hive, do you think you could tell which hive has the bees that dance and which one doesn't?

Student 19: This one dances (points to Hive A) and this one doesn't (Hive B).



04:57.0 Interviewer: Okay, how do you know?

Student 19: Um, I think that it'll be the opposite, my opposite answer.



05:07.0 Interviewer: What do you mean?

Student 19: Like this one would be the honey one (B) and this would be the bumblebee (A).



(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

05:17.0 Interviewer: Why do you say so?

Student 19: Because in the last picture, it said that this one didn't get as much nectar as this one.



05:27.0 Interviewer: Ok, how do you think the bees will fly though?

Student 19: Um, this one (A) would collect the pollen and then come back. And then get more. This one (B) would collect the pollen, come back and dance and then all the bees would come there.



In this short excerpt, the algorithm predicted two behavioural frame transitions, from calm to hesitant, and from hesitant to active. At time 4:47, the interviewer asks the student to tell which hive has the bees that do the “waggle” dance, and he answers concisely and with no hesitation. At time 4:57, when the interviewer prompts him to elaborate, he hesitantly changes his response. Note his change in body position, fidgeting, and tone of voice. At time 5:07, the interviewer asks the student to explain why he changed his mind, so he points at the beehive that he now thinks has the bees that dance. Again, his body language is different from the previous instance, as his body is slightly leaning forward, he is gesturing profusely, and looks up to think for a little bit and then looks back at the paper. This behavioural frame continues throughout time 5:27. Correspondingly, we interpreted these instances as transitions in the context of the interview and we can then examine the video in greater detail to develop an interpretation of their meaning. In this way, the automated detection of transitions can easily be coupled with interaction analysis.

3.3 Student Profiles

To explore beyond the 10-second behavioural frames and investigate whether longer-term patterns existed during the interview, we searched for typologies or student profiles — hypothesized to be consistent patterns of switching between frames that might predict students’ mechanistic reasoning. A student who is consistently uncomfortable, for example, may not know the content and will perform poorly, whereas a student who becomes quite animated at key moments may be exhibiting greater content knowledge.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

In order to create these profiles, we needed to first identify representative sequences of behavioural states, for which we used an Optimal Matching (OM) algorithm (Gabadinho, Ritschard, Mueller, & Studer, 2011). The OM algorithm computes pairwise dissimilarity values based on the number of transformations required to make two sequences identical — the more transformations required the more dissimilar two sequences are, and therefore the higher the OM value. These pairwise dissimilarity values produce a dissimilarity matrix, from which one can extract representative sequences.

The OM algorithm found three representative sequences, or student profiles. To check whether these three profiles were meaningful, we next compared the state proportions and transition probabilities. The first profile included students who spent most of the interview in the hesitant frame (90% of the time). We labelled this profile as “Stays Hesitant.” The second profile revealed a large proportion of active and hesitant frames (42% and 43%, respectively). It also revealed a high probability of transitioning from active to hesitant (71%). We labelled this profile as “To Hesitant.” A third profile revealed a large proportion of calm and active frames (43% and 47%, respectively). It also revealed a high probability of transitioning from calm to active (88%). We labelled this profile as “To Active” because students moved from a calm disposition to one that showed a more engaged one. Figure 6 shows the empirical distribution, grouped by representative sequence type, of state transitions across all student interviews.

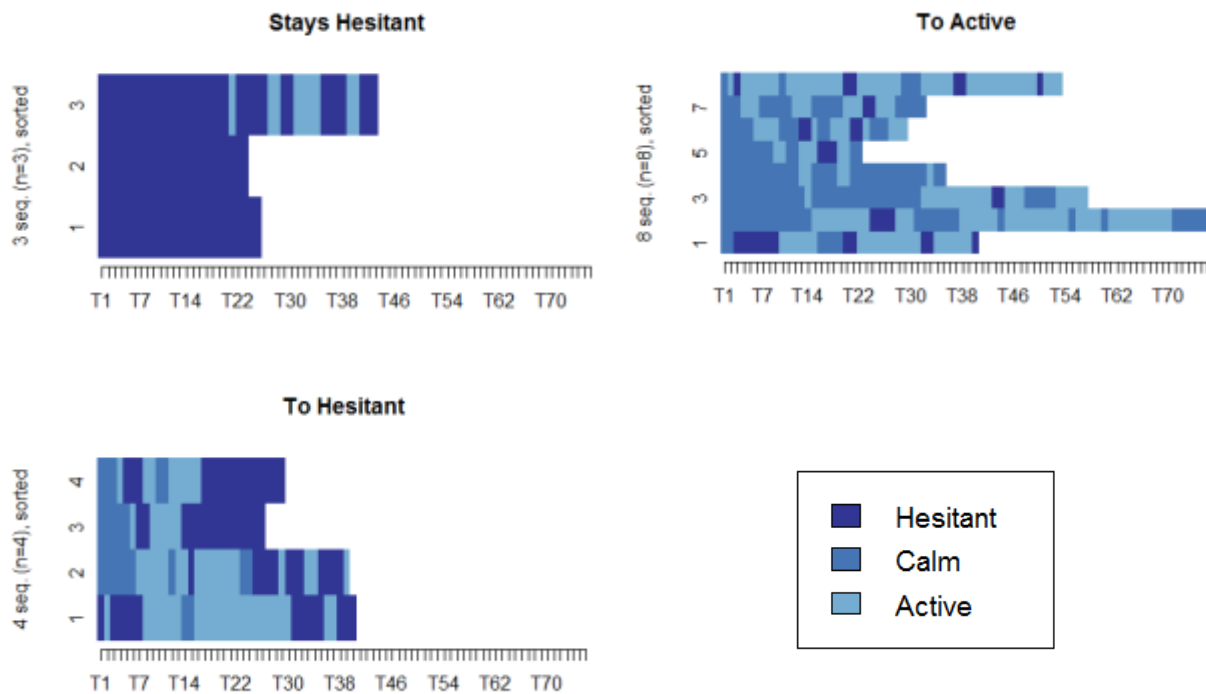


Figure 6: Student profiles as assessed by their interview sequence.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

3.4 Association with Mechanistic Reasoning

To determine whether the behavioural frames and the student profiles were meaningful, the individual interviews were coded with respect to the depth of student reasoning. To appraise the depth of student understanding about the complex mechanisms involved in the way bees collect nectar, we coded student mechanistic reasoning following Russ et al.’s (2008) guidelines. According to these guidelines, the interview transcripts were analyzed for instances of mechanistic reasoning in the form of components, phenomena, and mechanisms (for more information about this analysis see Danish et al., accepted). An analyst coded all utterances and a second analyst coded a randomly selected 30%. After a round of discussion to resolve discrepancies, both analysts recoded all the utterances again. Inter-rater agreement for this second round of coding was 94%. The mechanistic codes for each student were tallied to produce a total score. Because results showed that the experimental group had statistically significantly more instances of mechanistic reasoning than the control group, our results from the classification algorithms were only applied to the experimental group (i.e., fifteen students).

To understand the association between behavioural framing and mechanistic reasoning, we correlated total percent of time spent in each behavioural state and mechanistic reasoning scores. Spearman rank correlation coefficients were computed to evaluate the strength of these associations. Results show that the hesitant state is negatively associated with mechanistic reasoning ($r_s = -.806, p < .001$); the calm state is positively associated with mechanistic reasoning ($r_s = .939, p < .001$); and there seems to be a lack of association between the active state and mechanistic reasoning ($r_s = .327, p = .235$).

To explore this relationship further, we grouped students by profile and classified them into three mechanistic reasoning levels (low, medium, and high; see Table 5). Results show that students in the To-Active profile fell in the high and medium levels of mechanistic reasoning, in the To-Hesitant profile in the medium and low levels, and in the Stays-Hesitant profile in the low level. The association between profile and mechanistic reasoning is statistically significant, according to a Fisher’s Exact Test, $p = .005$. These findings show evidence that the order of transitions from one behavioural frame to the other created distinct long-term patterns in the interviews, which seem to reflect a possible relationship between profiles and mechanistic reasoning.

Table 5: Sequence clusters and mechanistic reasoning levels.

Mechanistic Reasoning	To Active	To Hesitant	Stays Hesitant
High (score > 66%)	5	0	0
Medium (33% < score < 66%)	3	2	0
Low (score < 33%)	0	2	3

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

4 DISCUSSION

Our analysis has demonstrated that it is possible to use multimodal learning analytic techniques to identify clusters of easily observable behaviours in students' interview activities. Furthermore, these behavioural frames are of potential analytic interest because they appear to relate to meaningful measures of student learning. Specifically, we see two important associations: 1) that these behavioural frames relate to performance within the interview context (e.g., the calm behavioural frame appears to co-occur with more demonstrations of competence in the form of mechanistic reasoning); and 2) the transitions between frames during the interview allow us to group students into a small set of profiles correlated with performance.

First, it appears that the behavioural frames are related to students' underlying expectations — i.e., about how they need to respond to the interviewer, and whether they have the expected or appropriate answer. The characteristics of the interactions in the Hesitant frame, as inferred from the HMM analysis, show that usually the student does not make much eye contact, restarts sentences, and provides tentative answers. These behaviours may reflect a student unwillingness to engage in elaborations or perhaps a perception that there is no time or expectation to do so. In contrast, in the Calm frame, the student replies with a quick answer, making eye contact, and gesturing infrequently. Given that the interview is being used as an assessment, and students likely interpret it as such, their calm demeanor may be due to their confidence in their answers. Of course, it is always possible for a participant to be erroneously confident, but this overlap in behavioural frames and evidence of mechanistic reasoning provides an interesting starting point for exploring exactly that sort of question.

Second, with respect to the student profiles, we note that behavioural frames change over time depending on the interviewer–interviewee interaction and the type of questions being asked. Not only was the type of behavioural frame related to mechanistic reasoning, but also the way that these frames transitioned throughout the interview, allowing us to create student profiles. The profile that relates to more instances of mechanistic reasoning transitions from Calm to Active state. A plausible explanation may be related to the type of questions in the interview since the second part of the interview was about ants, a transfer of content beyond the instructional content of bees, and calling for higher instantiations of mechanistic reasoning. The transfer questions may have required students to explore the content, prompting the use of gestures and more attention to the materials, whereas the bee questions prompted more recall. In addition, the student profile transitioning from Active to Hesitant may be because these students were exploring (instead of recalling) the content of bees, and were therefore less successful in making the transfer to the content of ants.

In addition, one more association deserves our attention. Given that our work was inspired, in part, by the research exploring epistemological framing, it is appropriate to ask if a relationship exists between the behavioural frames that we identified and the epistemological frames that others have noted before. Our initial results suggest that there may be a relationship, though it is not yet clear what that

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

relationship is. Tentatively, the behavioural frames we identified appear to be related to the confidence and energy that students feel as they respond within the interview context. Given that students do appear more calm (or active) when engaging in mechanistic reasoning, it would be easy to assume that there is therefore a relationship between these behavioural frames and kinds of epistemological frames that elicit similar forms of reasoning. We imagine that any such relationship would be highly contextual. For example, it may be that our sample included students who were either nervous and fidgety when engaging in inquiry about unfamiliar and challenging concepts, or were quite animated when they did so, and thus conveyed increased energy and confidence. We expect that while these exact patterns will not hold up in other samples or at other times, it is still quite likely that behavioural frames will continue to provide insight into those moments in the data that are worthy of continued attention and re-evaluation. In other words, we cannot yet claim there is a direct relationship between behavioural and epistemological frames, but we believe that behavioural frames can serve as a valuable tool for identifying when students perceive a shift in their current interactional context, thus providing us with a valuable analytic starting point.

5 CONCLUSION

Using coded video data, we have shown that it is possible to identify clusters of behaviours, called behavioural frames, using statistical models and matching algorithms. By providing a statistical model of student framing, our approach has the potential to support the ongoing refinement of theory and methods around behavioural, epistemological, and other social-interactional frames and their relationship to learning. As we further our understanding of how framing and learning are related, we believe that a positive feedback loop between methodological approaches leveraged by multimodal learning analytics and theory would emerge.

To conclude, we discuss some methodological choices that will continue to be at issue in future studies of behavioural frames. In developing a multimodal analytic approach, several decisions have to be made. A supervised approach uses a small subset of data and imposes an *a priori* structure on the rest of the sample, whereas an unsupervised approach takes into account all the data to produce statistical estimates of the behavioural frames. Without a hypothesis about the number and types of behavioural frames, however, this clustering approach is limited by the fact that the decision about the number of clusters is not straightforward and the most “plausible” solution results from model comparisons and statistical model fit indices. The meaning of the behavioural frames is then reconstructed by going back to the video data and interpreting the interaction in each particular context. As we pointed out already, the advantage of a supervised approach is that it is more theoretically grounded, but the danger in this case is to impose a particular structure on the data, making it difficult to see other possible behavioural frames or patterns. On the other hand, an unsupervised approach yields an empirical solution that may lead to speculative meaning reconstruction. Alternating between the different analytical strategies will more likely produce a fruitful dialogue for further theoretical developments in light of empirical findings.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

In addition to the choice of analytical strategies, the relevance of the actions, activities, and behaviours being coded will require careful attention. Surface-level features of actions and behaviours are more easily observed and captured by learning analytics, but what meaning do these hold, and what is the nature of their relationship to learning? A related and equally important consideration is the notion of time and the selection of an appropriate coding interval. Video data capture the dynamic properties of interaction among which time plays an important role in the characterization of behavioural frames. In this study, we chose a 10-second interval to code the behaviours. However, using shorter or longer intervals (e.g., 5 or 15 seconds) could possibly produce different clusters — that is, a different number of behaviour combinations. Additionally, if ignored, the serial dependency of behaviours — i.e., closer-in-time behaviours are expected to be more similar than those coded at longer time intervals — can artificially produce too many transitions in the data stream, and yield questionable inferences about behavioural frames and their relationship to learning. Finally, a quantitative (and even automated) analysis of the content of the student talk (e.g., by using some computational linguistic approach; see for instance Sherin, 2013) could be combined with the analysis of behavioural frames. All these considerations lay the groundwork for future studies and the issues that remain to be addressed to further understand the interplay of talk, behaviour, and learning.

ACKNOWLEDGMENTS

We are grateful to Asmalina Saleh, Jenna MacWilliams, and Branden Bryan for their help throughout various stages of this work.

REFERENCES

- Andrade-Lotero, A., Danish, J. A., Moreno, J., & Perez, L. (2013). Measuring “framing” differences of single-mouse and tangible inputs on patterns of collaborative learning. In N. Rummel, M. Kapur, M. Nathan, S. Puntambekar (Eds.), *Proceedings of the International Conference on Computer-Supported Collaborative Learning (CSCL 2013)*, (Vol.1, pp. 34–41). LuLu: International Society of the Learning Sciences.
- Andrade-Lotero, & Danish, J. A. (2014). Advancing epistemological frame analysis to refine our understanding of inquiry frames in early elementary interviews. In J. Polman, E. Kyza, D.K. O’Neill, I. Tabak, W. R. Penuel, S. Jurow, K. O’Connor, T. Le, L. D’Amico (Eds.), *Proceedings of the International Conference of the Learning Sciences ICLS Workshop on Learning Analytics for Learning and Becoming in Practice (ICLS’14)*, (Vol. 3, pp. 1637–1638). Boulder, CO: International Society of the Learning Sciences.
- Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.
- Bateson, G. (1955). A theory of play and fantasy. *Psychiatric Research Reports*, 2(39), 39–51.
- Berland, L. K., & Hammer, D. (2012). Framing for scientific argumentation. *Journal of Research in Science Teaching*, 49(1), 68–94. <http://dx.doi.org/10.1002/tea.20446>

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

- Berland, M., Baker, R., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning*, 19(1–2), 205–220. <http://dx.doi.org/10.1007/s10758-014-9223-7>
- Berland, M., Martin, T., Benton, T., Petrick Smith, C., & Davis, D. (2013). Using learning analytics to understand the learning pathways of novice programmers. *Journal of the Learning Sciences*, 22(4), 564–599. <http://dx.doi.org/10.1080/10508406.2013.836655>
- Blikstein, P. (2013). Multimodal learning analytics. *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13)*, 102–106. <http://dx.doi.org/10.1145/2460296.2460316>
- Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In R. Lesh & A. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 547–589). Hillsdale, NJ: Lawrence Erlbaum.
- Conlin, L. D., Gupta, A., Scherr, R. E., & Hammer, D. (2008). The dynamics of students' behaviors and reasoning during collaborative physics tutorial sessions. *arXiv preprint arXiv:0803.0323*. <http://dx.doi.org/10.1063/1.2820949>
- Danish, J. A., Enyedy, N., & Parnafes, O. (2015). Working towards an integrated analysis of knowledge in interaction. In A. Disessa, M. Levin, & N. J. S. Brown (Eds.), *Knowledge and interaction: A synthetic agenda for the learning sciences* (pp. 161–181). New York: Routledge.
- Danish, J. A. (2014). Applying an activity theory lens to designing instruction for learning about the structure, behavior, and function of a honeybee system. *Journal of the Learning Sciences*, 23(2), 100–148. <http://dx.doi.org/10.1080/10508406.2013.856793>
- Danish, J. A., Saleh, A., Andrade, A., & Bryan, B. (in press). Observing complex systems thinking in the zone of proximal development. *Instructional Science*.
- diSessa, A. (2007). An interactional analysis of clinical interviewing. *Cognition and Instruction*, 25(4), 523–565. <http://dx.doi.org/10.1080/07370000701632413>
- Drever, E. (1995). *Using semi-structured interviews in small-scale research: A teacher's guide*. Scottish Council for Research in Education.
- Echeverría, V., Avendaño, A., Chiluíza, K., Vásquez, A., & Ochoa, X. (2014). Presentation skills estimation based on video and Kinect data analysis. *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge (MLA '14)*, 53–60. <http://dx.doi.org/10.1145/2666633.2666641>
- Elby, A., & Hammer, D. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp. 409–434). Cambridge, UK: Cambridge University Press.
- Gabadinho, A., Ritschard, G., Mueller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37.
- Ginsburg, H. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge, UK: Cambridge University Press.

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

- Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge, MA: Harvard University Press.
- Gomes, J. S., Yassine, M., Worsley, M., & Blikstein, P. (2013). Analysing engineering expertise of high school students using eye tracking and multimodal learning analytics. *Proceedings of the 6th International Conference on Educational Data Mining (EDM 2013)*, 375–377. Retrieved from http://www.educationaldatamining.org/EDM2013/papers/rn_paper_88.pdf
- Hutchison, P., & Hammer, D. (2010). Attending to student epistemological framing in a science classroom. *Science Education*, 94(3), 506–524. <http://dx.doi.org/10.1002/sce.20373>
- Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics*, 1(1), 6–47. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3249>
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39–103. http://dx.doi.org/10.1207/s15327809jls0401_2
- Kendon, A. (1990). Behavioural foundations for the process of frame-attunement in face-to-face interaction. In A. Kendon (Ed.), *Conducting interaction: Patterns of behaviour in focused encounters* (pp. 239–262). Cambridge, UK: Cambridge University Press.
- Lantz, B. (2013). *Machine learning with R*. Birmingham, UK: Packt Publishing Ltd.
- Luzardo, G., Guamán, B., Chiluíza, K., Castells, J., & Ochoa, X. (2014). Estimation of presentations skills based on slides and audio features. *Proceedings of the 2014 ACM workshop on Multimodal Learning Analytics Workshop and Grand Challenge (MLA '14)*, 37–44. <http://dx.doi.org/10.1145/2666633.2666639>
- Martin, T., & Sherin, B. (2013). Learning analytics and computational techniques for detecting and evaluating patterns in learning: An introduction to the special issue. *Journal of the Learning Sciences*, 22(4), 511–520. <http://dx.doi.org/10.1080/10508406.2013.840466>
- Ochoa, X., Chiluíza, K., Méndez, G., Luzardo, G., Guamán, B., & Castells, J. (2013). Expertise estimation based on simple multimodal features. *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI '13)*, 583–590. <http://dx.doi.org/10.1145/2522848.2533789>
- Oviatt, S., Cohen, A., & Weibel, N. (2013). Multimodal learning analytics: Description of math data corpus for ICMI grand challenge workshop. *Proceedings of the 15th ACM International Conference on Multimodal Interaction (ICMI '13)*, 563–568. <http://dx.doi.org/10.1145/2522848.2533790>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Dubourg, V. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12, 2825–2830.
- Russ, R. S., Lee, V. R., & Sherin, B. L. (2012). Framing in cognitive clinical interviews about intuitive science knowledge: Dynamic student understandings of the discourse interaction. *Science Education*, 96(4), 573–599. <http://dx.doi.org/10.1002/sce.21014>

(2016). Using multimodal learning analytics to model student behavior: A systematic analysis of behavioural framing. *Journal of Learning Analytics*, 3(2), 282–306. <http://dx.doi.org/10.18608/jla.2016.3.2.14>

- 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(3), 499–525. <http://dx.doi.org/10.1002/sc.20264>
- Scherr, R. E., & Hammer, D. (2009). Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics. *Cognition and Instruction*, 27(2), 147–174. <http://dx.doi.org/10.1080/07370000902797379>
- Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of communication*, 49(1), 103–122. <http://dx.doi.org/10.1111/j.1460-2466.1999.tb02784.x>
- Sherin, B. (2013). A computational study of commonsense science: An exploration in the automated analysis of clinical interview data. *Journal of the Learning Sciences*, 22(4), 600–638. <http://dx.doi.org/10.1080/10508406.2013.836654>
- Shoukry, L., Göbel, S., & Steinmetz, R. (2014, September). Towards mobile multimodal learning analytics. Paper presented at the *Learning Analytics for and in Serious Games Workshop*, Graz, Austria. Retrieved from <ftp://www.kom.tu-darmstadt.de/papers/SGS14.pdf>
- Spybey, T. (1984). Frames of meaning: The rationality in organisational cultures. *Acta Sociologica*, 27(4), 311–322. <http://dx.doi.org/10.1177/000169938402700402>
- Tannen, D. (1993). What's in a frame? Surface evidence for underlying expectations. In D. Tannen (Ed.), *Framing in discourse* (Vol. 14, pp. 56). New York: Oxford University Press.
- van Dijk, T. A. (1977). Context and cognition: Knowledge frames and speech act comprehension. *Journal of Pragmatics*, 1(3), 211–231. [http://dx.doi.org/10.1016/0378-2166\(77\)90035-2](http://dx.doi.org/10.1016/0378-2166(77)90035-2)
- Visser, I., & Speekenbrink, M. (2010). depmixS4: An R-package for hidden Markov models. *Journal of Statistical Software*, 36(7), 1–21.
- Whiting, L. S. (2008). Semi-structured interviews: Guidance for novice researchers. *Nursing Standard*, 22(23), 35–40. <http://dx.doi.org/10.7748/ns2008.02.22.23.35.c6420>
- Wilensky, U. (1999). NetLogo: Center for connected learning and computer-based modeling. *Northwestern University*.
- Worsley, M. (2012). Multimodal learning analytics: Enabling the future of learning through multi modal data analysis and interfaces. *Proceedings of the 14th ACM International Conference on Multimodal Interaction (ICMI'12)* 353–356. <http://dx.doi.org/10.1145/2388676.2388755>
- Worsley, M., & Blikstein, P. (2014). Analyzing engineering design through the lens of computation. *Journal of Learning Analytics*, 1(2), 151–186. Retrieved from <https://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3541/0>
- Zucchini, W., & MacDonald, I. L. (2009). *Hidden Markov models for time series: An introduction using R*. Boca Raton, FL: CRC Press.