A Learning Analytics Approach to Monitoring the Quality of Online One-to-One Tutoring

Mutlu Cukurova[1], Madiha Khan-Galaria[2], Eva Millán[3], Rose Luckin[4]

Abstract
One-to-one online tutoring provided by human tutors can improve students' learning outcomes. However, monitoring the quality of such tutoring is a significant challenge. In this paper, we propose a learning analytics approach to monitoring online one-to-one tutoring quality. The approach analyzes teacher behaviours and classifies tutoring sessions into those that are effective and those that are not effective. More specifically, we use sequential behaviour pattern mining to analyze tutoring sessions using the CM-SPAM algorithm and classify tutoring sessions into effective and less effective using the J-48 and JRIP decision tree classifiers. To show the feasibility of the approach, we analyzed data from 2,250 minutes of online one-to-one primary math tutoring sessions with 44 tutors from eight schools. The results showed that the approach can classify tutors' effectiveness with high accuracy (*F* measures of 0.89 and 0.98 were achieved). The results also showed that effective tutors present significantly more frequent hint provision and proactive planning behaviours than their less-effective colleagues in these online one-to-one sessions. Furthermore, effective tutors sequence their monitoring actions with appropriate pauses and initiations of students' self-correction behaviours. We conclude that the proposed approach is feasible for monitoring the quality of online one-to-one primary math tutoring sessions.

Notes for Practice

- The need to monitor the quality of online tutoring has increased significantly in recent years.

- An approach using sequential pattern mining and decision trees is proposed to monitor online one-to-one primary math tutoring.

- Effective online tutors present statistically significantly more frequent appropriate hint provision and proactive planning behaviours in their sessions.

- Effective online tutors sequence monitoring actions with appropriate pausing and initiation of learner self-correction.

Corresponding author [1] Email: m.cukurova@ucl.ac.uk Address: UCL Knowledge Lab, Institute of Education, University College London, 23- 29 Emerald Street, London, WC1N 3QS, United Kingdom. ORCID ID: https://orcid.org/0000-0001-5843-4854
[2] Email: m.khan.16@ucl.ac.uk Address: UCL Knowledge Lab, Institute of Education, University College London, 23-29 Emerald Street, London, WC1N 3QS, United Kingdom. ORCID ID: https://orcid.org/0000-0001-5258-3058
[3] Email: eva@lcc.uma.es Address: Universidad de Málaga Avda. Cervantes, 2. 29071 Málaga, Spain. ORCID ID: https://orcid.org/0000-0001-9178-7600
[4] Email: r.luckin@ucl.ac.uk Address: UCL Knowledge Lab, Institute of Education, University College London, 23-29 Emerald Street, London, WC1N 3QS, United Kingdom. ORCID ID: https://orcid.org/0000-0001-9125-0846

1. Introduction

Online tutoring is an alternative to certain aspects of the traditional face-to-face model of education. However, there is good evidence to demonstrate that merely enabling online interaction opportunities for educators and learners is insufficient to achieve the expected learning outcomes. The quality of the tutoring that is provided is crucial to the effectiveness of online one-to-one tutoring (Donald, Blake, Girault, Datt, & Ramsay, 2009; Torgerson et al., 2016; Rapanta, Botturi, Goodyear, Guàrdia, & Koole, 2020; Hodges, Moore, Lockee, Trust, & Bond, 2020; Carter Jr., Rice, Yang, & Jackson, 2020). However, evaluating the quality of online tutoring sessions is a significant challenge. Many schools assume that the frequency of tutoring interventions is the only success measure required. Issues concerning the quality of that tutoring are often ignored, not least because of a lack of practical and effective ways to monitor tutoring quality. Historically, there has been limited research into

human-human online tutoring, particularly with regards to monitoring the quality of the tutoring (e.g., Johnson & Bratt, 2009; Kopp, Matteucci, & Tomasetto, 2012).

One way to ensure the quality of online tutoring is to employ human evaluators, who can watch a regular, randomly selected set of online tutoring sessions and score them using agreed-upon criteria to differentiate those tutors who are performing well or adequately from those who are not. However, there are various problems with this approach. Firstly, the online tutoring process is complex, meaning that identifying the evaluation criteria in such contexts presents a significant challenge. Secondly, even after such evaluation criteria have been identified correctly, their implementation by the human evaluators creates further challenges (e.g., the validity and reliability of the criteria, including how effectively each "tutor evaluator" can comprehend each criterion, and how much variance there is between different evaluators in their implementation of the same criteria). Thirdly, if we can identify the success criteria of online tutoring sessions accurately, and if we assume that the human evaluators can implement the evaluation criteria objectively and effectively, such an evaluation process still relies heavily on human labour, which is expensive, in short supply, and hard to scale.

In this paper, we propose an approach to monitoring the quality of online human-to-human tutoring sessions. The proposed solution first generates data from online tutoring behaviours with the help of a tagging interface. Then, these data are analyzed with a sequential pattern algorithm (CM-SPAM) to identify emerging pattern frequencies. Finally, the emerging sequential patterns and their frequency values are used to build decision trees to classify effective and less effective tutoring sessions. We investigate two main research questions here:

1. What are the sequential behaviour patterns of effective and less effective online tutors?
2. To what extent can tutors' sequential pattern behaviours be used to accurately classify tutor effectiveness in online one-to-one tutoring settings?

In order to show the value of the approach, we implement it in an online primary math tutoring platform. The online one-to-one tutoring sessions we have sampled follow a structured format, where the session starts with technical checks and warm-up questions, followed by instructional content on lesson objectives and problem-solving on topic questions, and concluding with a lesson plenary/reflection.

2. Literature Review

2.1. Sequential Pattern Mining and Decision Trees

The use of sequential pattern mining and decision tree modelling is not new in the area of learning analytics. However, these approaches are rarely employed together, or in combination with further classification methods. Neither have we found examples of their use to process human-observed behavioural data to explore tutor effectiveness in online one-to-one tutoring settings.

A survey of existing work illustrates, for example, the use of sequences of self- and socially shared regulatory activities from student log and online chat data to identify that more frequent and more diverse regulatory activities occur in data from successful groups than in data from less successful groups (Zheng, Xing, & Zhu, 2019). Similarly, sequential pattern mining, using the cSPADE algorithm applied to log data, has been used successfully when exploring whether differences exist between learners who viewed self-regulated learning prompt videos and those who did not (Wong, Khalil, Baars, de Koning, & Paas, 2019). A framework developed by Ozdagoglu, Oztas, and Cagliyangil (2019) was used to analyze event log data from learning management system platforms, using latent class analysis and sequential pattern mining approaches. Using a similar approach, latent study patterns have been identified in data from learner interactions with course materials in a MOOC (Boroujeni & Dillenbourg, 2019). In addition, there are examples of process mining being combined with sequence analysis to detect students' learning strategies (Matcha et al., 2020). For instance, pattern mining was used to detect the transition between learners' affective states in an open-ended learning environment, and the frequency of these patterns was then correlated with learning outcomes (Andres et al., 2011). Sequence mining has been applied to predict whether students were ready for assessments by detecting patterns of behaviour that were associated with high- and low-performing students (Malekian, Bailey, & Kennedy, 2020) and to identify beneficial and harmful learning habits, as well as to guide instructor interventions when students transition between online platforms (Gitinabard, Heckman, Barnes, & Lynch, 2019). Similarly, a combination of analytics was used by Saint, Gasevic, Matcha, Uzir, and Pardo (2020) to find patterns of self-regulated learning, including simple frequency measures, epistemic network analysis, temporal process mining, and stochastic process mining.

A different approach, and one that is perhaps more closely aligned with the research reported here, can be found in the work of Kopp and colleagues (2012), who looked at the daily support practices of e-tutors and used cluster analysis to identify two profiles of e-tutors. Their study revealed that there is indeed a difference between experienced and inexperienced e-tutors in the way they support online collaboration. E-tutors with experience considered specific cognitive activities to be more important for effective online collaboration, and they seemed to be more familiar with detecting and intervening to avoid dysfunctional social phenomena. However, this study relied solely on data from tutors' self-declared answers to a

questionnaire, not observations of their actual behaviour. Here, we present a learning analytics method for monitoring the effectiveness of online one-to-one tutoring sessions based on the analysis of tutor behaviours during online tutoring sessions.

The first step in the approach we propose here is to identify and describe potentially observable success behaviours exhibited by effective online tutors. We therefore initially present a review of the literature about effective tutoring behaviour. It is important to note that the purpose of this review is not to identify every possible indicator of successful tutoring behaviour; rather, it is to identify a few key features that can be operationalized and detected in online tutoring settings.

2.2. Effective Tutoring Conceptualizations

Numerous studies focus on effective tutor behaviours and their impact on the success of the tutoring process. Graesser, Person, Harter, and the Tutoring Research Group (2000), in their study aiming to understand the tutoring process, observe that tutoring sessions are largely driven by the tutor's agenda, which mainly comprises introducing topics, questions, and problems. Lepper and Woolverton (2002) explain that the tutoring process involves recurrent sequences of phases, including problem selection, problem presentation, problem solution, reflection, and instruction. They use the INSPIRE model to identify tutor behaviours that demonstrate success, namely intelligent, nurturant, Socratic, progressive, indirect, reflective, and encouraging behaviours. Earlier research on one-to-one tutoring, on the other hand, emphasizes the impact of tutors' subject knowledge (Schmidt, Dolmans, Gijselaers, & Des Marchais, 1995) and argues that it is the most significant factor for tutor success and, directly or indirectly, that tutors' subject knowledge shapes most of their tutoring behaviours. Similarly, Ritter, Anderson, Koedinger, and Corbett (2007) argue that in addition to declarative subject knowledge, tutors should have a strong procedural knowledge of the subject domain. In addition to subject knowledge, Anania's (1983) early study showed experimentally that the provision of cues and hints is a key aspect of the success of the tutoring process.

Other studies have focused on tutors' feedback as a variable affecting tutoring success (e.g., Merrill, Reiser, Ranney, & Trafton, 1992; Shute, 2008). For instance, Black and Wiliam (1998) divide tutoring actions into two different functions: directive (tutor provides what needs to be fixed) and facilitative (tutor provides comments and suggestions). As the authors argue, based on the different expected learning outcomes, successful tutoring would require different functions. More recently, Shute (2008) proposed the concept of feedback specificity, which argues that successful tutoring requires specific feedback because uncertainty in feedback leads to a lack of motivation in pupils. The impact of systematic feedback on tutor success has also been shown by Narciss and Huth (2004) in their tutoring feedback conceptual framework, as well as Kopp and colleagues (2012) in their study of differential practices by experienced and inexperienced tutors. Other studies have focused on the concrete aspects of tutors' feedback behaviour, such as the use of language (Schmidt et al., 1995) and the immediacy of tutors' behaviours in tutoring settings (Hastie, Chen, & Kuo, 2007). To emphasize the dialogic nature of the tutoring process, Graesser and colleagues put forward a pervasive student-tutor dialogue pattern of one-to-one online tutoring, consisting of a five-step dialogue frame: Step 1: Tutor asks a question. Step 2: Student answers the question. Step 3: Tutor gives short feedback on the quality of student's answer. Step 4: Tutor and student collaboratively improve the quality of the answer. Step 5: Tutor assesses student's understanding of the answer (Graesser, 1993; Graesser, Person, & Magliano, 1995).

At the metacognitive level, Lin, Schwartz, and Hatano (2005) devised the term *adaptive metacognition* to represent a person's ability to adapt themselves flexibly to their environment in unexpected situations. Their approach does not advocate for tutors to be metacognitively engaged at all times, but rather to reflectively seek out "hidden features" in seemingly "routine situations" (p. 246). The approach has been shown to lead to success in tutoring practice. More recently, Wittwer, Nückles, Landmann, and Renkl (2010) demonstrated that providing information about a pupil's current state of understanding to the online tutors enabled the tutors to more effectively diagnose misconceptions and tailor instructional explanations to their students' needs. Interestingly, the opposite was true when tutors were provided with knowledge deficit information (Sleeman, Kelly, Martinak, Ward, & Moore, 1989), suggesting that tutors are better at following a general sequence of instruction than they are at deconstructing a situation to address specific misconceptions (Wittwer, Nückles, & Renkl, 2008).

A great amount of work also relates to the affective aspects of successful tutoring processes (Plass, Heidig, Hayward, Homer, & Um, 2014; Pardos, Baker, San Pedro, Gowda, & Gowda, 2013; Wang, Li, Geng, & Wang, 2019). Most researchers argue for the impact of positive emotions that benefit students' motivation and learning outcomes (Amiri & Ghonsooly, 2015) and negative emotions that have the opposite effect (cf. Ryan and Deci (2009) for the impact of boredom). However, which emotions are negative and which are positive is not always crystal clear. For instance, although anxiety is traditionally considered as a negative emotion, Hwang, Hsu, Lai, and Hsueh (2017) show that learners with higher levels of anxiety have better learning achievements than those with less anxiety. Simply put, emotions can "trigger, sustain, or reduce academic motivation and related volitional processes" (Pekrun, Goetz, Titz, & Perry, 2002, p. 97). Therefore, successful tutoring also depends on a tutor's ability to manage their pupils' emotions.

Tutoring is a multidimensional, complex process, and its successful delivery depends on a range of factors that could perhaps be broadly categorized as cognitive, metacognitive, and affective (Cukurova, Mavrikis, Luckin, Clark, & Crawford, 2017). However, the separation between these domains is not clear, and various features can be considered under multiple

categories. For example, feedback actions can be considered within any of the three categories, depending on the feedback action's content and implementation. Also, although there is plenty of research on human tutoring in face-to-face environments and intelligent tutoring systems, online human tutoring is still a relatively understudied field. The quality of tutoring is key to the success of an online tutoring session; however, what exact tutor actions are more likely to lead to success is still a question with no clear answers. This is at least in part due to the numerous contextual factors playing a significant role in deciding whether a tutor behaviour is valuable for students' learning or not. For example, when Lepper and Woolverton (2002) argue for the INSPIRE characteristics of successful tutoring, they do not provide the exact details of how such characteristics can be implemented and monitored in real-world practice. Perhaps the chaotic nature of the tutoring process can be used as an excuse for such ambiguous representations of successful tutoring. Just as in many educational practices, the subjective nature of tutor success makes its implementation, evaluation, and training extremely challenging, it also makes its quality-monitoring process very human dependent in many real-world education settings.

3. Methodology

3.1. Potential Signifiers of Tutor Success in Online Tutoring

In this subsection, we present the potential signifiers of effective tutor behaviours that we have identified from the literature reviewed above. These behaviour labels have been studied, either theoretically or experimentally, and argued to help monitor and detect the quality of tutoring in online settings. Table 1 presents the tutor behaviours we have identified and operationalized in building the behaviour-tagging interface developed as part of our quality-monitoring approach. Please note that the purpose of this section is to describe the potential effective tutoring signifiers tested in the particular tutoring context we researched (primary math one-to-one tutoring). It is not to establish a set of criteria that would universally lead to success in all tutoring sessions. Based on the literature reviewed, we identified seven behavioural indicators: appropriate pausing after questions, context-based questions/answers or examples, hint provision, monitoring and clarification actions (or questions), proactive planning actions, initiation of self-correction, and writing on the Virtual Classroom Environment (VCE) interface about the content discussed. We then used these behavioural indicators in a tagging interface built to tag tutor behaviours in real time. Figure 1 shows this interface.
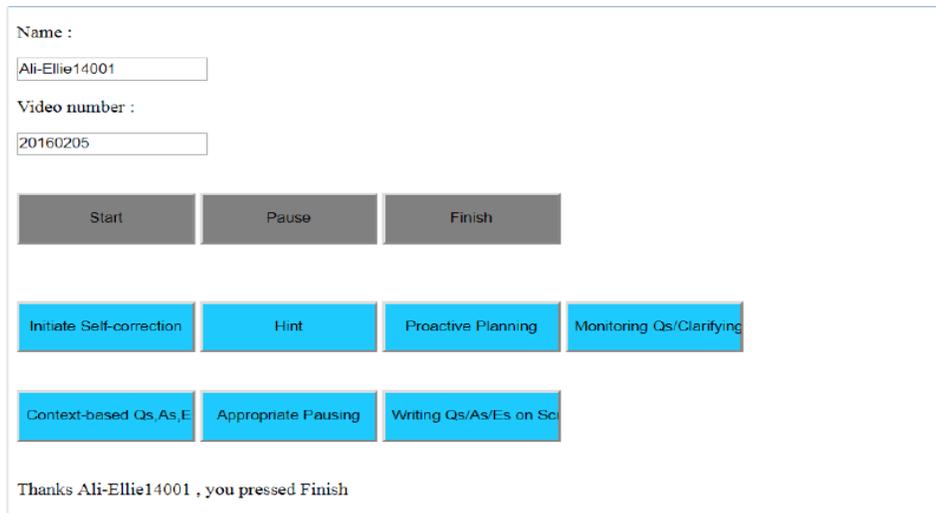

Figure 1: A screenshot of the annotation tool built for tagging tutor behaviour

These behaviours were observed as they emerged while tutors were teaching tutees using the virtual classroom interface presented in Figure 2.
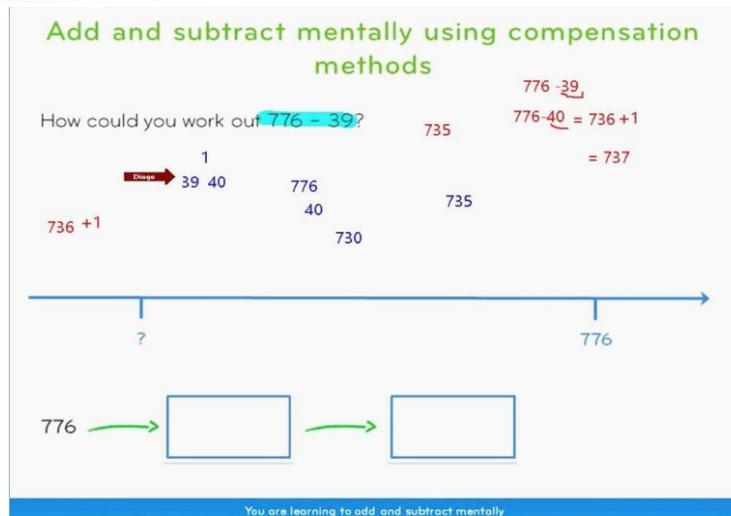
ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

4

Figure 2: A screenshot of the online tutoring video interface

While choosing the potential signifiers for effective tutor behaviours, we particularly focused on experimental evidence. However, we also took theoretical considerations into account because the experimental evidence base for some theoretically effective tutor behaviours was weak and ambiguous. For instance, providing hints rather than direct instruction was often referred to as an effective tutoring behaviour in the literature. However, it was shown by D'Mello and Graesser (2012) that hints are more productive for high-achieving students, whereas low-achieving students should be provided with direct instruction in tutoring sessions. Similarly, it was also argued that successful tutors tend to be involved in delayed-feedback actions more frequently (Butler, Karpicke, & Roediger III, 2007). Indeed, Butler and colleagues (2007) report that students who receive delayed feedback outperform their peers who receive immediate feedback during instruction. On the other hand, in an earlier study, Kulik and Kulik (1988) found that immediate feedback could be more effective in terms of students' knowledge acquisition. However, there is also evidence that although a delay in feedback to a learner's response could enhance their performance in subsequent questions, students still prefer immediate feedback over delayed feedback (Lefevre & Cox, 2017). This has significant implications on the effectiveness of a tutoring session because a tutoring approach that contradicts students' preferences could negatively impact student motivation and, consequently, engagement and learning. There might be various reasons for such differences, and, as argued by Mathan and Koedinger (2002), the appropriate timing of feedback is not something that is fixed, but rather it is related to the nature of the task, the ability of the student, and their individual needs and preferences. Regardless, feedback behaviours such as hint provision and appropriate pausing can indeed indicate effective tutoring. There is plenty of literature that argues for the benefit of waiting for students to reflect on, modify, or elaborate on their answers as well as tutor's questions (Walsh & Sattes, 2016).

Moreover, successful tutoring sessions are argued to involve proactive planning behaviours by tutors (Wittwer et al., 2010). These behaviours are mainly context specific. For instance, tutors in successful tutoring sessions begin each session by ensuring that both they and their tutees are oriented to the same section of the screen and checking the technical issues of the microphone and the tutor tools. They are also proactive in terms of engaging with the tutees—rather than asking them whether they understand the problem or example, they have them read it and then ask what they understand. Such proactive planning is considered to be a strategy to gather more information about the tutees in online tutoring sessions, which in turn has been shown to have a significant impact on the effectiveness of tutoring sessions (Wittwer et al., 2010).

Following the effective feedback theme, the type of questions a tutor poses appears to have an impact on the quality of a tutoring session. In previous work, it was also observed that tutees usually benefit from seeing the questions, answers, or examples clearly written on the screen provided during online tutoring. In their groundbreaking work, Shorrocks-Taylor and Hargreaves (1999) showed that the use of an apparatus to present the questions and answers clearly to tutees facilitated them to better understand what they were expected to do and what was required of them to work out answers to the questions. Similarly, contextualization of the information presented frequently occurs in successful tutoring sessions, particularly within math tutoring. Shield and Dole (2013) argue that mathematical questions based on authentic and real-life contexts lead to students' deep reasoning, therefore leading to more productive tutoring sessions. Similarly, Bell (1993) showed that the presentation of content within a familiar context can allow students to extend this information to other related contexts, therefore leading to better transfer of mathematical knowledge.

Table 1: Tutor Behaviours Identified from the Literature and Coded with the Tagging Interface

Code	Tutor Behaviour	Definition of Behaviour	Reference
a	Appropriate pausing after questions	Tutor pauses after asking a question to allow the learner time to respond.	Walsh and Sattes (2016)
c	Context-based questions, answers, or examples	Tutor asks the learner questions or provides answers and examples that are contextually relevant to guide the learner's understanding.	Shield and Dole (2013)
h	Hint provision	Tutor provides a small piece of information to scaffold the learner as they work through a problem.	D'Mello and Graesser (2012) Mathan and Koedinger (2002)
m	Monitoring and clarification actions or questions	Tutor passively observes a learner or takes clarifying actions.	Lepper and Woolverton (2002)
p	Proactive planning action	Tutor supports the learner in planning an approach to a question and engages with them proactively.	Wittwer et al. (2010)
s	Initiation of self-correction	Tutor prompts the learner to detect and correct their mistakes.	Wylie and Chi (2014)
w	Writing on the virtual classroom interface	Tutor writes on the virtual classroom interface to help explain content to the learner.	Shorrocks-Taylor and Hargreaves (1999)

3.2. Context of the Online Tutoring Platform

For our empirical work, we partnered with an industrial supplier named Third Space Learning (TSL), which delivers math tutoring for primary school children aged 9 to 10 years using an online tutoring platform. Tutoring sessions are organized by subject topic, and learners are assigned to a specific curriculum when they are enrolled in TSL. A tutor is allocated to a learner based on availability, and this tutor will typically deliver most, if not all, of the sessions in the learner's program. Each session lasts about 45 minutes and follows the same format: a warm-up question, lesson objectives, topic questions, and then the lesson plenary/reflection. Tutors guide the learners through the session, and there is a level of flexibility over the structure of the session; for example, tutors can choose to skip certain questions or introduce new ones. In total, online math tutoring sessions were sampled from eight schools.

The online tutoring platform consists of an interactive whiteboard and tools, where the lesson objectives, topic questions, and plenary questions are set out. The learner and the tutor communicate through audio, and they use the pointer and whiteboard tools to draw, type, and erase (Figure 2). If the audio is interrupted, the learner and tutor can also chat through the messaging box displayed on the side of the screen, but there is no video available of either the learner or the tutor. Tutors can reward learners for their work by giving them effort points or presenting them with an emoji or picture.

In the online tutoring environment, learners and tutors log into the system, in which the learner works through a pre-designed online set of questions, with the guidance of their human tutor. All data from tutors and tutees were anonymized and their consent was obtained before the data collection took place.

3.3. Tutor Evaluation

For the ground truth about the effectiveness of online tutoring sessions, we have used a combination of metrics. Tutors are evaluated using four metrics. Firstly, learners rate sessions on several criteria, including enjoyment and clarity. Secondly, learner outcomes for a tutor (i.e. the difference between the pre-diagnostic test score and the post-diagnostic test score) are computed. Thirdly, a panel of three human evaluators (experienced online primary math tutors) evaluate and score a randomly selected sample of sessions for a tutor. A minimum of two sessions are evaluated for each tutor, and the scores from the human evaluators are averaged to provide an overall score. Tutors are evaluated against 28 criteria, including tutor interactions (impacting learning and mindset), subject knowledge, language, familiarity, and safeguarding. The criteria used for human evaluations can be seen at https://bit.ly/3uzhsxF. Finally, the tutors' scores on the subject matter (math) and English tests were used in the evaluation. These four metrics were weighted equally and used to provide an overall evaluation score for each tutor. Based on this aggregated score for tutors, those tutors whose score was higher than the average score for all tutors were categorized as effective, and those with a less than average score were categorized as less effective.

3.4. Participants and Tutor Behaviour Labelling Procedure

In total, 2,250 minutes of online tutoring were recorded from 44 tutors. The research involved 26 male and 18 female tutors, who delivered the tutoring through the online platform to students from eight different schools. Each tutor's practice was evaluated as one incident based on one of their tutoring sessions, which lasted on average 50 minutes. Schools and tutors were from varied backgrounds to improve the representativeness of the sample. The student and tutor were matched randomly depending on the availability of tutors and the demands of the tutees. Using the combined metrics described in Section 3.3, 24 tutors were categorized as effective tutors and 20 tutors were categorized as less effective tutors. Tutor behaviours were tagged

by human observers using the bespoke tagging tool (Figure 1). Human evaluators were experienced primary math tutors who had experience in delivering online tutoring. They were provided with a three-hour workshop during which the tagging tool was introduced and the behaviours were defined and exemplified. Afterwards, they were asked to watch online tutoring sessions and do the tagging in real time. If they had any confusion or questions, they were asked to pause or reverse the recording of a session.

It is important to note that we initially considered more behaviour indicators from the literature reviewed above, but seven indicators was the suitable number for human coders to reliably undertake the tagging task while observing online tutor behaviours in real time. In our pilot studies, we attempted to tag up to ten tutor behaviours. The inter-coder reliability value when human evaluators used seven indicator behaviours to code was significantly higher than when they used ten indicators to code (ordinal k alpha = 0.623 vs. 0.892). We chose the seven indicators that were best understood and easiest to implement by the human evaluators. All tags as well as their timestamps were collected on a local SQL database. Logs involved the video ID, human observer ID, tag name, and timestamp. Then, the log files from the SQL database were exported for data analysis and modelling.

3.5. Pre-processing for Sequential Pattern Mining and Classification Tasks

To be able to observe at which section of a 45-minute-long tutoring session sequential behaviour patterns show differences, we structured each session into time bins based on the structured learning design of the online tutoring sessions described in Section 3.2. A typical session was structured as shown in Figure 3.

| Technical checks (~2 minutes) | Warm up (~5-10 minutes) | Lesson objectives (~10 minutes) | Topic questions (~20-25 minutes) | Lesson plenary/ reflection (~5-10 minutes) |

Figure 3: Structure of an online tutoring session into time bins

We then used the SPMF library to find the sequences of behaviours that appeared more frequently in effective and less effective tutors' online tutoring sessions. SPMF is an open-source data mining library written in Java, specialized in pattern mining (https://www.philippe-fournier-viger.com/spmf/index.php). More specifically, we used the CM-SPAM algorithm, an approach to fast vertical mining of sequential patterns that uses co-occurrence information (Fournier-Viger, Gomariz, Campos, & Thomas, 2014). CM-SPAM discovers all frequent sequential patterns occurring in a sequence database. We used the parameters below in the data analysis. Minimum pattern length was set as two items, and maximum pattern length was set to the maximum possible number of items. We did not specify any particular required item in the mining process and set the maximum gap between items to one. Finally, we used a minsup value of 50%, which meant that a sequential pattern is considered as frequent when it appears in more than half of the total number of sequences.

To build a classification model that would help us automatically categorize tutors into effective and less effective tutor classes based on their behaviour patterns, we first calculated the number of times each sequential pattern emerged in each session and normalized these values according to the total number of patterns in the session. Total pattern lengths varied from 83 to 212 in 44 sessions. Then, using WEKA software, we built several classifiers and tested the performance of each of these models using 10-fold cross-validation.

4. Results

4.1. Frequency of Tutor Behaviours

First, we studied how many of the 44 tutors were tagged with each particular tutor behaviour by the tool in their sessions. The frequency of tutors who present each behaviour at least once in the sessions is presented in Table 2. Appropriate pausing, context-based questioning, monitoring and clarification actions, initiation of self-correction, and writing on the VCE interface about the content discussed were all exhibited at least once by almost all tutors in their sessions. However, there were differences between the appropriate hint provision and proactive planning behaviours of less effective tutors' sessions and those of effective tutors' sessions. Although these behaviours were presented at least once by almost all effective tutors, eight of the less effective tutors never presented an appropriate hint provision behaviour in their sessions. Also, 12 less effective tutors never presented proactive planning behaviours in their sessions. The results indicate that less effective tutors do not present appropriate hint provision and proactive planning behaviours as frequently as effective tutors.

Table 2: Number of Tutors for Whom Each Behaviour Was Observed

Group	Action	Number of Tutors for Whom the Behaviour Was Observed	Group	Action	Number of Tutors for Whom the Behaviour Was Observed	Group	Action	Number of Tutors for Whom the Behaviour Was Observed
All tutors	m	44	Effective tutors	m	24	Less effective tutors	c	20
All tutors	c	44	Effective tutors	a	24	Less effective tutors	m	20
All tutors	a	43	Effective tutors	c	24	Less effective tutors	a	19
All tutors	s	43	Effective tutors	h	24	Less effective tutors	s	19
All tutors	w	41	Effective tutors	s	24	Less effective tutors	w	18
All tutors	h	36	Effective tutors	w	23	Less effective tutors	h	12
All tutors	p	30	Effective tutors	p	22	Less effective tutors	p	8

a = appropriate pausing after questions, c = context-based questions, answers, or examples, h = hint provision, m = monitoring and clarification actions or questions, p = proactive planning action, s = initiation of self-correction, w = writing on the virtual interface.

4.2. Frequency of Tutor Behaviour Patterns in Different Stages of an Online Tutoring Session

The sequential pattern analysis below shows that behaviour pattern differences become particularly apparent during the initial warm-up questions, the instructional section covering the lesson objectives, and the problem-solving on-topic questions stages of a typical tutoring session on the TSL platform. On the other hand, during the technical checks at the beginning and the lesson plenary/reflection phases at the end of the online tutoring sessions, distinctions between effective and less effective tutors' behaviours were not statistically significant.

Table 3 illustrates the chi-square tests undertaken to compare the statistical significance of the differences between the frequencies of sequential patterns of behaviours. The results show that there are statistically significant differences between certain sequential behaviour patterns occurring in the warm-up questions, lesson objectives, and topic questions stages ($p < 0.05$). The purpose of this exploratory investigation was to decide which behaviour patterns are likely to have a significant difference that can be used in decision tree classifications. Therefore, to avoid a possible type II error, no statistical correction was made to the p values.

Looking at Table 3, a couple of interesting results emerge in the differences between effective and less effective tutors' behaviours in online tutoring sessions. First of all, it is interesting to observe that effective tutors sequence their monitoring behaviours with appropriate pausing behaviours. That means monitoring questions or actions are frequently either followed by or preceded by enough time and space for students to think and reflect on their actions. On the other hand, less effective tutors sequence their monitoring actions without allowing appropriate space and time for students' reflections. As can also be seen, less effective tutors might present up to six monitoring actions in a sequence. However, such long monitoring action sequences are non-existent in effective tutors' sessions. Furthermore, particularly during the instructional section covering the lesson objectives, effective tutors sequence their monitoring actions by initiating students' self-correction behaviours. On the contrary, none of the less effective tutors presented this sequence of monitoring actions and initiations of students' self-correction. It is also interesting to observe that there was no statistically significant difference between effective and less-effective tutors in the sequential patterns observed during technical checks and the reflection phase of the tutoring sessions.

4.3. Classification Models for Effective and Less Effective Tutors

The best result was achieved by separating the normalized sequential tutor behaviour patterns into three bins (namely low frequency, medium frequency, and high frequency) and by using the JRIP classifier. The decision tree's classification rule is given below:

> *if* "as" pattern = "low frequency" => Label = Less effective tutor
>> *else* => Label = Effective tutor

Table 3: Sequential Behaviour Patterns in Different Stages of Online Tutoring (* indicates statistical significance $p < 0.05$)

Stage	Pattern	Effective Tutors	Less Effective Tutors	$P(X^2)$
Technical checks	p, c	17	16	21.05
Warm-up	m, a	24	0	0.01*
Warm-up	a, m	24	0	0.01*
Warm-up	m, m	0	18	0.00*
Warm-up	m, m, m	0	10	0.01*
Instruction	m, a	24	12	57.15
Instruction	a, m	24	11	42.94
Instruction	m, a, m	22	0	0.02*
Instruction	m, s	20	0	0.05
Instruction	m, m	0	20	0.00*
Instruction	m, m, m	0	15	0.00*
Problem-solving	m, a	24	0	0.01*
Problem-solving	a, m	24	0	0.01*
Problem-solving	m, m	23	0	0.02*
Problem-solving	m, a, m	23	0	0.02*
Problem-solving	a, m, a	23	0	0.02*
Problem-solving	m, m, m	0	19	0.00*
Problem-solving	m, m, m, m	0	14	0.00*
Problem-solving	m, m, m, m, m	0	13	0.00*
Problem-solving	m, m, m, m, m, m	0	10	0.01*
Reflection	m, s	13	16	5.53

With this rule as tested with a 10-fold cross-validation, 39 out of 44 instances were correctly classified and an *F*-measure of 0.887 was achieved. Table 4 shows the detailed accuracy values and the confusion matrix of the classification model. This result emphasizes the distinctive value of "as" behaviour sequence in online one-to-one primary math tutoring sessions, which follows appropriate pausing behaviours with the initiation of students' self-corrections.

Table 4: Detailed Accuracy Values by Class and Confusion Matrix for the Three-Bins Model

Precision	Recall	*F* Measure	Class	Effective Tutor	Less Effective Tutor	Classified as
0.952	0.833	0.889	Effective tutor	20	4	Effective tutor
0.826	0.950	0.884	Less effective tutor	1	19	Less effective tutor

To build a more detailed classification model and to potentially improve the classification performance, we also tried a model that used the normalized frequency of behaviours according to session duration rather than pattern length. In this case, the best results were achieved by discretizing the variables in four bins (very low, low, high, very high frequency). We used the J48 algorithm for the classification task. The performance was not further improved, but high accuracy of 88.64% was reached. As the decision tree shows (Figure 4), the root of the tree was the pattern of following appropriate pausing behaviours with self-initiation of students' explanations; however, there was further consideration of the "ama" tutor behaviour pattern. This pattern refers to appropriate pausing behaviours being sequenced with monitoring actions and questions.
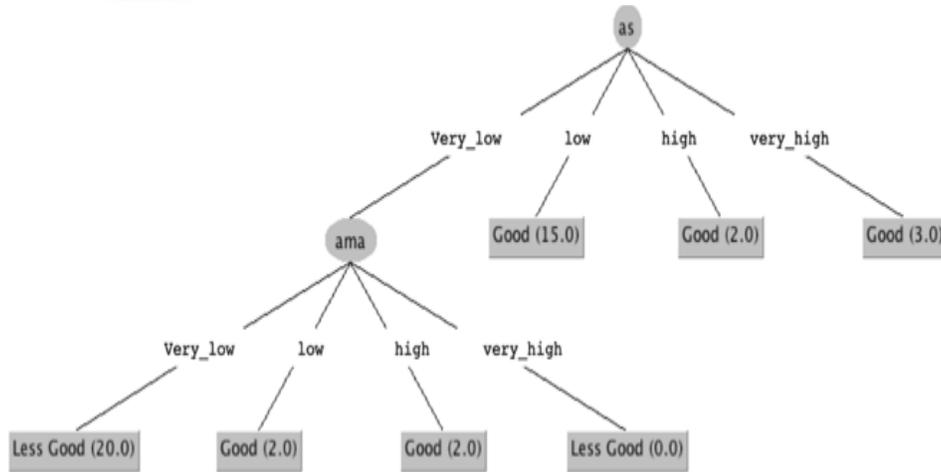
Figure 4: Decision tree for the classification of online tutors based on sequential
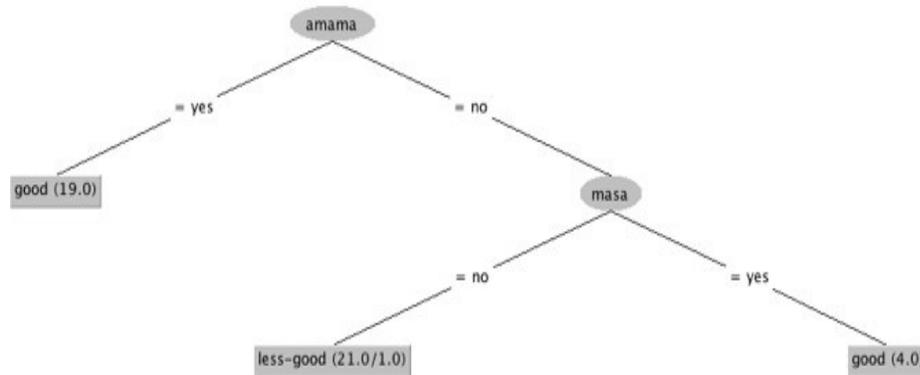behaviour patterns normalized for session time

Figure 5: Decision tree for the classification of online tutors based on sequential behaviour
patterns of a minimum of four behaviours

Table 5: Detailed Accuracy Values by Class and Confusion Matrix for the Long Sequence Model

Precision	Recall	*F* Measure	Class	Effective Tutor	Less Effective Tutor	Classified as
1.000	0.958	0.979	Effective tutor	23	1	Effective tutor
0.952	1.000	0.976	Less effective tutor	0	20	Less effective tutor

4.4. Discussion

In this paper, we present a learning analytics approach to monitoring the quality of online one-to-one primary math tutoring. The proposed solution first uses potential success indicators of tutor behaviours to generate observation data from online tutoring sessions with the help of a bespoke tagging interface. Human evaluators tag seven indicators of effective tutoring behaviours as they occur while they are watching the recordings of online one-to-one tutoring sessions in real time. Then, the data generated are analyzed with the CM-SPAM sequential pattern algorithm to calculate the frequency of sequential tutor behaviours from online tutoring sessions. The frequencies of behaviour sequences are then used to build JRIP and J-48 decision tree classifiers to categorize effective and less effective tutors successfully. The results presented in this paper provide a promising proof of concept of our approach that can detect and classify effective and less effective tutor behaviours in online tutoring sessions with high accuracy.

More specifically, we have investigated two research questions. Firstly, we looked at the behaviour patterns of effective and less effective online tutors. Our results showed that less effective tutors are less likely to provide appropriate hints in their sessions. As discussed in our literature review, there is plenty of previous research showing the value of appropriate hint

provision in face-to-face teaching settings (e.g., Anania, 1983; Merrill et al., 1992; Shute, 2008). Our results confirm these findings and extend them, showing the distinctive value of appropriate hint provision in effective one-to-one online tutoring. D'Mello and Graesser (2012) show that hints are more productive for high-achieving students, whereas low-achieving students should be provided with direct instruction in tutoring sessions. However, in our evaluation, the overall effective tutor behaviours were associated with hint provision rather than direct instruction. It is worth noting that, in this study, we did not categorize students into high- or low-achiever categories, which might have influenced the observed results.

The results also show that effective tutors exhibit proactive planning behaviours significantly more frequently than less effective tutors. These behaviours include, but are not limited to, effective online tutors beginning each slide of the presentation by ensuring that both they and their tutees were oriented to the same section of the screen, checking the technical issues of the microphone and the tutor tools, engaging with the students proactively by asking them to read the problem, and asking them what they have understood from the problem in order to address any possible misunderstandings at an early stage. These results are in alignment with previous research showing the significant impact of proactive planning behaviours on the effectiveness of tutoring sessions (Wittwer et al., 2010).

To provide more detailed insights into effective practice, we looked at the sequences of online tutor behaviours and compared effective tutor behaviours with less effective ones. Indeed, statistically significant differences were observed in the sequences of behaviours presented by effective and less effective online tutors. For instance, effective tutors frequently sequence their monitoring behaviours with appropriate pausing behaviours, whereas less effective tutors sequence their monitoring behaviours with further monitoring actions and questions, rather than pausing and allowing enough time for students to think and reflect. These results are aligned with previous work showing that successful tutors frequently present delayed feedback and monitoring actions to allow space for learners to process information and reflect (Butler et al., 2007; Kulik & Kulik, 1988). Although such appropriate pausing behaviours appear to indicate effective tutor behaviours in online one-to-one tutoring settings, the learner implications of such behaviours are not considered in this research article. As argued by Lefevre and Cox (2017), students usually prefer immediate support and monitoring over delayed actions, which might have significant implications on the effectiveness of an online tutoring session due to learners' unpleasant experiences and their lack of motivation to engage. Also, although we have identified that effective tutors present appropriate pausing behaviours sequenced with their monitoring behaviours, the timing of such behaviours is not fixed and is dependent on the nature of the task, the ability of the learner, and their individual needs and preferences (Mathan & Koedinger, 2002).

The results also show that less effective tutors can present up to six monitoring actions in a sequence; however, long sequences of monitoring actions are non-existent in effective tutors' sessions. In addition to sequencing their monitoring actions with appropriate pausing, effective tutors also sequence their monitoring actions with initiating students' self-correction behaviour. Less effective tutors do not present this sequence of monitoring actions combined with the initiation of students' self-correction. Self-correction requires learners to realize the process of their understanding and recognize the gaps in their knowledge through thinking and reflections, and it can lead to higher learning outcomes in various fields, including computer programming, biology, probability, and reading comprehension (Wylie & Chi, 2014). As argued by Chi, Bassok, Lewis, Reimann, and Glaser (1989), self-correction opportunities also promote the process of knowledge transfer. Besides, prompted self-correction made by online tutors might also induce learners' metacognitive development to a certain extent (Chi, De Leeuw, Chiu, & LaVancher, 1994; Hacker et al., 2015). However, in this research paper, we have not focused on the impact of tutor behaviours on students' metacognitive skills. This particular literature gap provides multiple opportunities for future research. See for instance early work focusing on tutor behaviours that stimulate learner metacognition in online tutoring sessions (Khan-Galaria, Cukurova, & Luckin, 2020).

In our results, we found no statistically significant difference in the sequential patterns of effective and less effective tutors during the initial technical check stage as well as the reflection/plenary phases of online tutoring sessions. On the other hand, behaviour pattern differences emerged during the warm-up questions, the instructional stage where the content of the lesson objectives is covered, and the problem-solving stages of the online tutoring session. These results might be specific to the particular online tutoring context we have investigated. Although our results suggest that evaluations of online tutoring sessions might benefit from focusing on those sections in which significant behaviour pattern differences emerge, we recognize that not all online tutoring sessions would follow such a structured approach, nor would they have the same stages as the sessions we investigated. Therefore, these findings should be investigated further in terms of their potential to be transferred to different online tutoring contexts.

Lastly, in our second research question, we intended to investigate to what extent tutors' sequential pattern behaviours can be used to accurately classify tutor effectiveness in online tutoring settings. The results show that we have achieved 88.64% and 97.73% accuracy values with different classification trees. Due to the significantly small sample size as well as the small number of sequence frequencies investigated here, overfitting in our models is a significant concern. However, the high accuracy values achieved are promising for future research to investigate the potential of the approach presented here at larger

scales. The roots and the leaves of the classification trees were mainly associated with the sequences of tutor behaviours that combine monitoring actions with appropriate pausing and the initiation of self-correction behaviours. Because these sequences were statistically significantly different behaviour patterns between effective and less effective tutors in online tutoring settings, these results are not surprising. The highest accuracy in decision trees was achieved when we used longer sequential behaviour patterns as input. For instance, when we only took into account four or more behaviour sequence patterns, the decision tree classified all less effective tutors correctly and it only mislabelled one effective tutor as less effective by error. These results are similar to previous research in the use of sequential pattern mining approaches in education showing that longer behaviour patterns might be stronger indicators of differences. See, for instance, Kinnebrew, Loretz, and Biswas (2013) for the sequences of online learner behaviours in a digital learning environment, or Martinez-Maldonado, Yacef, Kay, Kharrufa, and Al-Qaraghuli (2010) for students' sequential patterns of collaborative learning activity around an interactive tabletop.

4.5. Limitations

In this section, we discuss some limitations of the work presented in this paper.

As stated in Section 2.2, the original purpose of the work presented here is to identify potential signifiers in the particular tutoring context we are studying (primary math one-to-one tutoring). The methodology proposed here has the potential to be used in similar subjects' online tutoring, for example, for more advanced mathematics or other problem solving–based domains in which learning has a similar methodology (e.g., physics). This requires further research. However, the application of this approach to other domains, such as arts and humanities, may not be very productive. Here, we are only focusing on tutor behaviours and we do not consider the wider context for these behaviours, such as learner actions. Effective tutor behaviours are highly contextualized and hard to interpret in isolation. Although the indicator behaviours we have chosen are purposely easy to interpret, further research that examined tutor-learner interactions in context would enable a more rounded view of tutor practices that can be associated with different clusters of tutors. In our current work, we are also investigating the potential of process mining approaches that have the potential to observe tutor behaviours in a more contextually meaningful manner. Also, there are still challenges with the scalability of the approach proposed here. The approach proposed brings in significant efficiency compared to the traditional quality evaluation of an online tutoring session (evaluation of a session using the metrics in Section 3.3 takes around three hours, whereas the learning analytics approach presented here can be implemented in real time in a 45-minute session). More importantly, decision trees built using this approach can reach a classification decision based on a small number of tagged tutor behaviours. As more reliable models are built with more representative data in the future, classification decisions can be reached from the tagged behaviours of randomly selected short segments of an online tutoring session. The sample size in our study was 44 tutors. For sure, it would be desirable to have a larger set of data, but organizational and human resources constraints make it difficult to obtain such a set. Nevertheless, the same problem is present in previously published related studies that focus on teacher behaviours. For example, the sample size of Prieto, Sharma, Kidzinski, Rodríguez-Triana, and Dillenbourg (2018) was two teachers in 12 classroom sections; Camacho, de la Guía, Olivares, Flores, and Orozco-Barbosa (2020) was three teachers in 18 learning activities; Prieto, Sharma, Dillenbourg, and Jesús (2016) was four teachers in four sessions of 35 to 40 minutes; and Riel, Lawless, and Brown (2018) was 42 teachers. We acknowledge that the findings of this study come from a limited sample size and are required to be further confirmed in future studies. However, our sample size is comparable to recent research in the field.

Finally, the ground truth used in our classification is based on the rankings provided by human evaluators. These rankings are likely to include a level of human bias and inconsistency. We recommend that future work use alternative versions of the ground truth, including relatively more standardized measures such as learners' content acquisition. In addition, this study does not include any particular qualitative evaluation, such as tutor and tutee interviews. Although students' post-session feedback was taken into account in the tutor evaluations, detailed interviews with students could also generate insights into how satisfied they are with different tutoring practices. As such, we are unable to triangulate our findings with tutor and tutee self-reports. Triangulation would enable us to further validate our findings and enable a richer discussion of the inferences that can be drawn from the tutor behaviours observed.

5. Conclusion

This article investigates the nature of effective tutor behaviours in online one-to-one tutoring settings and proposes a learning analytics approach to monitoring the quality of online tutoring. Our results showed statistically significant differences in the sequences of effective versus less effective tutors' online tutoring behaviours. We illustrated that the sequential pattern behaviours of tutors can be used to accurately classify tutor effectiveness in online tutoring settings. The effectiveness of an online one-to-one tutoring session is complex and not a problem that can be addressed with one particular approach. However, the results we presented here give promising opportunities for providing analytics solutions to monitor the quality of online

one-to-one tutoring more efficiently, improving the current practice through the implications of its findings, and also generating various future research opportunities.

Declaration of Conflicting Interest

We declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgements

References

Amiri, M., & Ghonsooly, B. (2015). The relationship between English learning anxiety and the students' achievement on examinations. *Journal of Language Teaching and Research*, *6*(4), 855–865. http://dx.doi.org/10.17507/jltr.0604.20

Anania, J. (1983). The influence of instructional conditions on student learning and achievement. *Evaluation in Education: An International Review Series*, *7*, 3–76. https://doi.org/10.1016/0191-765X(83)90002-2

Andres, J. M. A. L., Ocumpaugh, J., Baker, R. S., Slater, S., Paquette, L., Jiang, Y., … Biswas, G. (2019). Affect sequences and learning in Betty's Brain. In *Proceedings of the Ninth International Conference on Learning Analytics & Knowledge* (LAK 2019), 4–8 March 2019, Tempe, Arizona, USA (pp. 383–390). New York: ACM. https://doi.org/10.1145/3303772.3303807

Bell, A. (1993). Principles for the design of teaching. *Educational Studies in Mathematics*, *24*(1), 5–34. https://doi.org/10.1007/BF01273293

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Boroujeni, M. S., & Dillenbourg, P. (2019). Discovery and temporal analysis of MOOC study patterns. *Journal of Learning Analytics*, *6*(1), 16–33. https://doi.org/10.18608/jla.2019.61.2

Butler, A. C., Karpicke, J. D., & Roediger III, H. L. (2007). The effect of type and timing of feedback on learning from multiple-choice tests. *Journal of Experimental Psychology: Applied*, *13*(4), 273. https://doi.org/10.1037/1076-898X.13.4.273

Camacho, V. L., de la Guía, E., Olivares, T., Flores, M. J., & Orozco-Barbosa, L. (2020). Data capture and multimodal learning analytics focused on engagement with a new wearable IoT approach. *IEEE Transactions on Learning Technologies*, *13*(4), 704–717. https://doi.org/10.1109/TLT.2020.2999787

Carter Jr., R. A., Rice, M., Yang, S., & Jackson, H. A. (2020). Self-regulated learning in online learning environments: strategies for remote learning. *Information and Learning Sciences*, *121*(5/6), 321–329. https://doi.org/10.1108/ILS-04-2020-0114

Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*(2), 145–182. https://doi.org/10.1207/s15516709cog1302_1

Chi, M. T., De Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439–477. https://doi.org/10.1016/0364-0213(94)90016-7

Cukurova, M., Mavrikis, M., Luckin, R., Clark, J., & Crawford, C. (2017). Interaction analysis in online maths human tutoring: The case of third space learning. In *Proceedings of the International Conference on Artificial Intelligence In Education* (AIED 2017), 28 June–1 July 2017, Wuhan, China (pp. 636-643). Cham: Springer. https://doi.org/10.1007/978-3-319-61425-0_80

Cullinane, C., & Montacute, R. (2020). *COVID-19 and Social Mobility Impact Brief #1: School Shutdown*. London, UK: Sutton Trust. Retrieved from https://eric.ed.gov/?id=ED605807

D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, *22*(2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

Donald, C., Blake, A., Girault, I., Datt, A., & Ramsay, E. (2009). Approaches to learning design: Past the head and the hands to the HEART of the matter. *Distance Education*, *30*(2), 179–199. https://doi.org/10.1080/01587910903023181

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

13

Fournier-Viger, P., Gomariz, A., Campos, M., & Thomas, R. (2014). Fast vertical mining of sequential patterns using co-occurrence information. In *Proceedings of the Pacific-Asia Conference on Knowledge Discovery and Data Mining* (PAKDD 2014), 13–16 May 2014, Tainan, Taiwan (pp. 40–52). Cham: Springer. https://doi.org/10.1007/978-3-319-06608-0_4

Gitinabard, N., Heckman, S., Barnes, T., & Lynch, C. F. (2019). What will you do next? A sequence analysis on the student transitions between online platforms in blended courses. arXiv:1905.00928. Retrieved from https://arxiv.org/abs/1905.00928

Graesser, A. C. (1993). *Questioning Mechanisms during Tutoring, Conversation, and Human-Computer Interaction* (Tech. Rep.). Memphis State University Department of Psychology. https://apps.dtic.mil/sti/pdfs/ADA266420.pdf

Graesser, A. C., Person, N., Harter, D., & Tutoring Research Group. (2000). Teaching tactics in autotutor. Modelling human teaching tactics and strategies. *International Journal of Artificial Intelligence in Education*, *11*, 1020–1029. http://users.sussex.ac.uk/~bend/its2000/graesser.pdf

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, *9*(6), 495–522. https://doi.org/10.1002/acp.2350090604

Hacker, D. J., Dole, J. A., Ferguson, M., Adamson, S., Roundy, L., & Scarpulla, L. (2015). The short-term and maintenance effects of self-regulated strategy development in writing for middle school students. *Reading & Writing Quarterly*, *31*(4), 351–372. https://doi.org/10.1080/10573569.2013.869775

Hastie, M., Chen, N.-S., & Kuo, Y.-H. (2007). Instructional design for best practice in the synchronous cyber classroom. *Journal of Educational Technology & Society*, *10*(4), 281–294. Retrieved from https://www.j-ets.net/collection/published-issues/10_4

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*, *27*. Retrieved from https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Hwang, G.-J., Hsu, T.-C., Lai, C.-L., & Hsueh, C.-J. (2017). Interaction of problem-based gaming and learning anxiety in language students' English listening performance and progressive behavioral patterns. *Computers & Education*, *106*, 26–42. https://doi.org/10.1016/j.compedu.2016.11.010

Johnson, G. M., & Bratt, S. E. (2009). Technology education students: E-tutors for school children. *British Journal of Educational Technology, 40*(1), 32-41. https://doi.org/10.1111/j.1467-8535.2007.00805.x

Khan-Galaria, M., Cukurova, M., & Luckin, R. (2020). A framework for exploring the impact of tutor practices on learner self-regulation in online environments. In *Proceedings of the International Conference on Artificial Intelligence in Education* (AIED 2020), 6–10 July 2020, Ifrane, Morocco (pp. 135–139). Cham: Springer. https://doi.org/10.1007/978-3-030-52240-7_25

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*, *5*(1), 190–219. https://doi.org/10.5281/zenodo.3554617

Kopp, B., Matteucci, M. C., & Tomasetto, C. (2012). E-tutorial support for collaborative online learning: An explorative study on experienced and inexperienced e-tutors. *Computers & Education*, *58*(1), 12–20. https://doi.org/10.1016/j.compedu.2011.08.019

Kulik, J. A., & Kulik, C.-L. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, *58*(1), 79–97. https://doi.org/10.3102/00346543058001079

Lefevre, D., & Cox, B. (2017). Delayed instructional feedback may be more effective, but is this contrary to learners' preferences? *British Journal of Educational Technology*, *48*(6), 1357–1367. https://doi.org/10.1111/bjet.12495

Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. Aronson (ed.), *Improving Academic Achievement* (pp. 135–158). Elsevier. https://doi.org/10.1016/B978-012064455-1/50010-5

Lin, X., Schwartz, D. L., & Hatano, G. (2005). Toward teachers' adaptive metacognition. *Educational Psychologist*, *40*(4), 245–255. https://doi.org/10.1207/s15326985ep4004_6

Malekian, D., Bailey, J., & Kennedy, G. (2020). Prediction of students' assessment readiness in online learning environments: The sequence matters. In *Proceedings of the 10th International Conference on Learning Analytics and Knowledge* (LAK 2020), 23–27 March 2020, Frankfurt, Germany (pp. 382–391). New York: ACM. https://doi.org/10.1145/3375462.3375468

Martinez-Maldonado, R., Yacef, K., Kay, J., Kharrufa, A., & Al-Qaraghuli, A. (2011). Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. In *Proceedings of the Fourth International Educational Data Mining Conference* (EDM 2011), 6–8 July 2011, Eindhoven, Netherlands. Retrieved from https://educationaldatamining.org/EDM2011/wp-content/uploads/proc/edm11_proceedings.pdf

ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

14

Matcha, W., Gasevic, D., Jovanovic, J., Pardo, A., Lim, L., Maldonado-Mahauad, J., … Tsai, Y.-S. (2020). Analytics of learning strategies: Role of course design and delivery modality. *Journal of Learning Analytics*, 7(2), 45–71. https://doi.org/10.18608/jla.2020.72.3

Mathan, S. A., & Koedinger, K. R. (2002). An empirical assessment of comprehension fostering features in an intelligent tutoring system. In *Proceedings of the International Conference on Intelligent Tutoring Systems* (ITS 2002), 2–7 June 2002, Biarritz, France, and San Sebastian, Spain (pp. 330–343). Springer. https://doi.org/10.1007/3-540-47987-2_37

Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences*, 2(3), 277–305. https://doi.org/10.1207/s15327809jls0203_2

Narciss, S., & Huth, K. (2004). How to design informative tutoring feedback for multimedia learning. In H. M. Niegemann, D. Leutner, and R. Brünken (eds.), *Instructional Design for Multimedia Learning*, pp. 181–195. Kornwestheim, Germany: Waxmann Buecher. Retrieved from https://studierplatz2000.tu-dresden.de/toolkit/presentations/CD/Literatur/Publikationen/ID_nahu.pdf

Ozdagoglu, G., Oztas, G. Z., & Cagliyangil, M. (2019). An application framework for mining online learning processes through event-logs. *Business Process Management Journal*, 25(5), 860–886. https://doi.org/10.1108/BPMJ-10-2017-0279

Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., & Gowda, S. M. (2013). Affective states and state tests: Investigating how affect throughout the school year predicts end of year learning outcomes. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (LAK 2013), 8–12 April 2013, Leuven, Belgium (pp. 117–124). New York: ACM. https://doi.org/10.1145/2460296.2460320

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91–105. https://doi.org/10.1207/S15326985EP3702_4

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction*, 29, 128–140. https://doi.org/10.1016/j.learninstruc.2013.02.006

Prieto, L. P., Sharma, K., Dillenbourg, P., & Jesús, M. (2016). Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (LAK 2016), 25–29 April 2016, Edinburgh, UK (pp. 148–157). New York: ACM. https://doi.org/10.1145/2883851.2883927

Prieto, L. P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M. J., & Dillenbourg, P. (2018). Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. *Journal of Computer Assisted Learning*, 34(2), 193–203. https://doi.org/10.1111/jcal.12232

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., & Koole, M. (2020). Online university teaching during and after the Covid-19 crisis: Refocusing teacher presence and learning activity. *Postdigital Science and Education*, 2, 923–945. https://doi.org/10.1007/s42438-020-00155-y

Riel, J., Lawless, K. A., & Brown, S. W. (2018). Timing matters: Approaches for measuring and visualizing behaviours of timing and spacing of work in self-paced online teacher professional development courses. *Journal of Learning Analytics*, 5(1), 25–40. https://doi.org/10.18608/jla.2018.51.3

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. (2007). Cognitive tutor: Applied research in mathematics education. *Psychonomic Bulletin & Review*, 14, 249–255. https://doi.org/10.3758/BF03194060

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wenzel & A. Wigfield (Eds.), *Educational Psychology Handbook Series. Handbook of Motivation at School* (pp. 171–195). Routledge/Taylor & Francis Group. https://psycnet.apa.org/record/2009-24219-009

Saint, J., Gasevic, D., Matcha, W., Uzir, N. A., & Pardo, A. (2020). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In *Proceedings of the 10th International Conference on Learning Analytics & Knowledge* (LAK 2020), 23–27 March 2020, Frankfurt, Germany (pp. 402–411). New York: ACM. https://doi.org/10.1145/3375462.3375487

Schmidt, H. G., Dolmans, D., Gijselaers, W. H., & Des Marchais, J. E. (1995). Theory-guided design of a rating scale for course evaluation in problem-based curricula. *Teaching and Learning in Medicine: An International Journal*, 7(2), 82–91. https://doi.org/10.1080/10401339509539719

Shield, M., & Dole, S. (2013). Assessing the potential of mathematics textbooks to promote deep learning. *Educational Studies in Mathematics*, 82, 183–199. https://doi.org/10.1007/s10649-012-9415-9

Shorrocks-Taylor, D., & Hargreaves, M. (1999). Making it clear: A review of language issues in testing with special reference to the national curriculum mathematics tests at key stage 2. *Educational Research*, *41*(2), 123–136. https://doi.org/10.1080/0013188990410201

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sleeman, D., Kelly, A. E., Martinak, R., Ward, R. D., & Moore, J. L. (1989). Studies of diagnosis and remediation with high school algebra students. *Cognitive Science*, *13*(4), 551–568. https://doi.org/10.1016/0364-0213(89)90023-2

Torgerson, C., Ainsworth, H., Buckley, H., Hampden-Thompson, G., Hewitt, C., Humphry, D., … Torgerson, D. (2016). *Affordable Online Maths Tuition: Evaluation Report and Executive Summary*. Education Endowment Foundation. Retrieved from https://eric.ed.gov/?id=ED581116

Walsh, J. A., & Sattes, B. D. (2016). *Quality Questioning: Research-Based Practice to Engage Every Learner*. Corwin Press. ISBN: 1 74101 442 5. https://books.google.co.uk/books?id=lG51DQAAQBAJ&lr=

Wang, Y., Li, T., Geng, C., & Wang, Y. (2019). Evaluating student learning effect based on process mining. In *Proceedings of the Second International Conference on Applied Informatics* (ICAI 2019), 7–9 November 2019, Madrid, Spain (pp. 59–72). Springer. https://doi.org/10.1007/978-3-030-32475-9_5

Wittwer, J., Nückles, M., Landmann, N., & Renkl, A. (2010). Can tutors be supported in giving effective explanations? *Journal of Educational Psychology*, *102*(1), 74. https://doi.org/10.1037/a0016727

Wittwer, J., Nückles, M., & Renkl, A. (2008). Is underestimation less detrimental than overestimation? The impact of experts' beliefs about a layperson's knowledge on learning and question asking. *Instructional Science*, *36*, 27–52. https://doi.org/10.1007/s11251-007-9021-x

Wong, J., Khalil, M., Baars, M., de Koning, B. B., & Paas, F. (2019). Exploring sequences of learner activities in relation to self-regulated learning in a massive open online course. *Computers & Education*, *140*, 103595. https://doi.org/10.1016/j.compedu.2019.103595

Wylie, R., & Chi, M. T. (2014). 17—The self-explanation principle in multimedia learning. In R. E. Mayer (ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 413–432). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.021

Zheng, J., Xing, W., & Zhu, G. (2019). Examining sequential patterns of self- and socially shared regulation of STEM learning in a CSCL environment. *Computers & Education*, *136*, 34–48. https://doi.org/10.1016/j.compedu.2019.03.005