

Using Learner Analytics to Explore the Potential Contribution of Multimodal Formative Assessment to Academic Success in Higher Education

SAGE Open
April-June 2022: 1–12
© The Author(s) 2022
DOI: 10.1177/21582440221089957
journals.sagepub.com/home/sgo

Lucas Kohnke¹ , Dennis Fount² , and Julia Chen³

Abstract

Blended learning pedagogical practices supported by learning management systems have become an important part of higher education curricula. In most cases, these blended curricula are evaluated through multimodal formative assessments. Although assessments can strongly affect student outcomes, research on the topic is limited. In this paper, we adopted a learning analytics approach to explore student engagement with formative assessments and the power of these assessments to predict student outcomes in an English for Academic Purposes courses in a Hong Kong university. The study retrieved the data logs from 7,815 students and used the data to analyze student engagement with the formative assessments. The results suggested that the students put effort into completing the assessments. The degree to which assessments predict learning outcomes depend on students' level of subject knowledge and their understanding of the relevance of the assessments. This study showed that learning analytics provided reliable evidence for understanding students' engagement and identifying at-risk students. Therefore, learning analytics research has the potential to inform pedagogical practice.

Keywords

multimodal formative assessments, English for Academic Purposes, learning analytics, engagement, blended courses

Introduction

The expanding use of learning management systems (LMSs) has made blended learning the new norm in higher education. Blended learning encompasses a range of pedagogical practices that suit today's learners and maximize the potential for academic success (Garrison & Vaughan, 2013; Porter et al., 2016; Ross & Gage, 2006). In English for Academic Purposes (EAP) courses worldwide, technology-based, and face-to-face approaches are increasingly being blended to facilitate second language (L2) acquisition (Terauchi et al., 2019). Online multimodal formative assessments, delivered through an LMS and aligned with the established principles of L2 acquisition (Chapelle & Sauro, 2017; Hockly, 2018), can enable authentic and engaging learning language-learning experiences that foster student autonomy (Park et al., 2016). Accordingly, there has been a call for higher education institutions to use formative assessments in LMSs to help EAP students succeed in language coursework and other studies.

In recent years, both the number of L2 learners pursuing higher education and efforts to integrate technology into EAP programs have increased greatly. However, research on the value of formative assessment has chiefly examined

other courses and disciplines such as quantitative methods (e.g., Choi et al., 2018; Tempelaar et al., 2018, 2020), business (e.g., Tempelaar et al., 2011), history, social care, maths, and engineering (e.g., Hlosta, et al., 2017). There has been little in-depth analysis of the extent to which EAP formative assessments delivered through LMSs such as Blackboard can predict students' future performance and on the potential of learner analytics to identify L2 students who are academically at risk. The emergence of blended learning as the new norm for language acquisition in a higher-education environment in which the focus is on quality outcomes (Gruba et al., 2019; Terauchi et al., 2019) has raised questions regarding individual students' formative assessment behavior online and the best practices for supporting their success in courses

¹The Education University of Hong Kong, Tai Po, Hong Kong

²The University of British Columbia, Vancouver, Canada

³The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Corresponding Author:

Dennis Fount, School of Journalism, Writing and Media, The University of British Columbia, Vancouver, BC V6T 1Z1, Canada.
Email: dennis.fount@gmail.com



Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (<https://creativecommons.org/licenses/by/4.0/>) which permits any use, reproduction and distribution of

the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

and maximizing the potential of blended learning. Past studies have focused on disciplines in which multimodal formative quizzes are used to help students understand concepts that they will be tested on in a final exam. However, in an L2 setting, students may learn concepts via multimodal formative assessments but they will still need to develop skills—skills that they will have to demonstrate in their final L2 assessment. Therefore, the predictive power of multimodal formative assessments in L2 courses remains uncertain. To gain insight into the power of such assessments to predict students' final course grades, we identified the time required to complete a quiz, the number of quiz attempts, and scores on LMSs as factors to assess whether students were at risk for failing L2 courses.

Literature Review

In higher education, blended learning has become more common owing to the rapidly increasing demand for online and distance courses (Brown, 2016; Moradimokhles & Hwang, 2020). Blended learning integrates didactic teaching pedagogy and multimedia technologies to transform the traditional classroom (Graham, 2006; Joosten et al., 2014). The benefits of this approach may include increased efficiency (Rubio et al., 2018), flexibility, and ease of access for learners (Bonks et al., 2005). Blended learning has been found to increase engagement, retention, and student satisfaction (Baumgartner, 2015; Gadbois & Quildon, 2013; Moradimokhles & Hwang, 2020) and to facilitate language learning in general (Kohnke, 2021; Li, 2017). Effectively incorporating blended learning into course delivery and learning resources (Thomas et al., 2017) requires a thorough understanding of the ways in which higher education students learn (Golonka et al., 2014).

Numerous studies have reported positive effects from blended learning pedagogies (Banditvilai, 2016). Students who engage in this form of learning, compared with those in traditional courses (whether taught face-to-face or online), exhibit modestly better performance (Owston & York, 2018). However, there are significant challenges to using blended learning in course delivery. It can be difficult to stimulate interaction, facilitate student learning and allow flexibility in learning (Boelens et al., 2015). An awareness of these challenges was especially important in framing the present study, as they shed light on the possible facilitating effect of technology in L2 language learning, including in the context of independent learning.

EAP courses prepare students to undertake university studies in English (Flowerdew & Peacock, 2001) by cultivating academic language proficiency and communicative competence (Hyland, 2006). There is a growing awareness of the challenges that L2 students face as they learn to read and write, and EAP education can support their efforts to acquire new discourse practices that will help them to succeed in all of their academic endeavors (Hyland 2018; Lughmani & Fount,

2021). With ever more L2 EAP students enrolling in higher education institutions around the world, it is important to determine the effectiveness of multimodal formative assessments delivered through an LMS in facilitating these students' L2 acquisition and predicting their final course grades.

Formative quizzes or assessments (Gardner, 2012) are integral to L2 learning. They are used to identify and monitor students' knowledge and skills (Hinkelman, 2018). While the literature offers multiple definitions of formative assessment (William & Leahy, 2015), a common conceptualization is to understand it as an assessment *for* learning that allows students to understand what they are learning and what they should aim for during their learning journey (Elwood & Kelnowski, 2002). Although students should always be the central actors in formative assessment (Wong, 2017), teachers can also benefit from its assessment of student learning to inform and adapt their teaching (Popham, 2008; Yan et al., 2021). These principles are important as they guide the design of the multimodal formative assessment examined in this study.

Multimodal formative assessments through LMSs constitute the core of blended learning approaches (Staker & Horn, 2012). Recent studies have found that students tend to have a positive attitude toward online quizzes (Cohen & Sasson, 2016; Wallihan et al., 2018), particularly those that incorporate multimodal elements and active learning (Cook & Babon, 2017; Gamage et al., 2019). Particularly helpful for L2 learners are automated quizzes that can be taken repeatedly and offer timely and actionable feedback. These features allow EAP students to self-assess and guide themselves through the process of acquiring an L2 while still participating in an active and supportive learning environment. But automated quizzes can also provide information on student progress—for example, on how a student performs relative to their peers. Blackboard, for instance, features built-in analytical tools that can assist teachers and administrators to understand students' needs even as they help them to understand the subject matter through the completion of targeted activities (Hinkelman, 2018). However, students may do well on tasks such as understanding the concepts taught in EAP courses and the expectations that they are expected to meet and yet fail to do well in course assessments (i.e., not develop the skills needed to employ EAP).

LMSs such as Blackboard support and enhance traditional course delivery in a blended manner through such built-in tools as discussion forums, chats, file sharing, quizzes, and videos (Cole et al., 2019; Coskuncay & Ozkan, 2013). These functions foster independent language learning (Ghazali, 2018). Previous studies of LMSs and language learning have investigated the use of technology to elicit and assess interactions (Dixon & Fuchs, 2015) and its impact on learners' evaluations of their coursework (Rubio, 2015). That said, technology alone cannot facilitate teaching and learning. Previous studies have established that teachers tend to use LMSs for only basic functions such as uploading course syllabi or PowerPoint slides (Dahlstrom et al., 2014; Graham et al., 2013). Higher

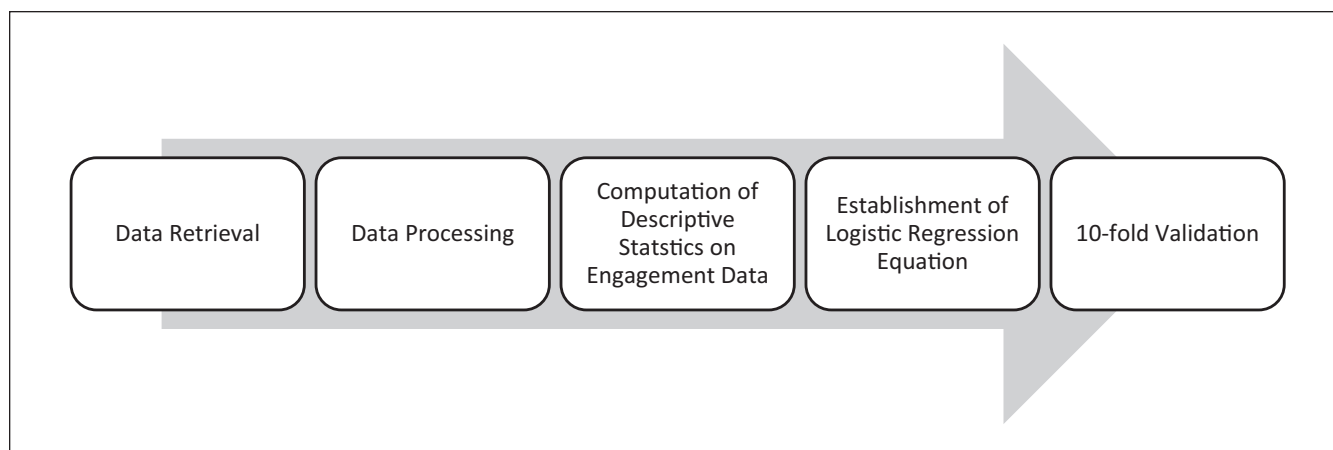


Figure 1. Data analysis procedures.

education institutions have seldom leveraged the full features of LMSs to transform learning and teaching.

A generally accepted definition of learning analytics is “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs” (Ferguson, 2012). The analytics of learners can be of use in understanding the complexities of L2 acquisition (Thomas et al., 2017) because they reveal what students do. They can be used to bridge pedagogy and analytics (Rienties et al., 2018). The major advantage of learning analytics is that engagement is measured via empirical data on students’ actual behaviors; the data are retrieved directly from server logs rather than gathered from self-reported or perception-based questionnaires and interviews. Learning analytics can be used to develop predictive models to measure, understand, and analyze students’ work on LMSs (Atherton et al., 2017). Language learning is a long-term investment, so there is a pressing need for insight into the type of learner resources that students access (e.g., quizzes, discussion boards, or forums), the frequency with which they access them, any unusual behaviors in which they may engage in this regard (Baker et al., 2006; Chen & Fount, 2020; Scheffel et al., 2001) and any positive or negative effects the resources have on them (Conati & Maclaren, 2009; Reffay & Chanier, 2003). By analyzing LMS data, institutions, and language teachers acquire information that can help them to provide personalized and rich learning, predict student success, and support at-risk students (Zhang et al., 2018). LMS data have been used in such ways in a wide range of disciplines in higher education (e.g., Choi et al., 2018; Hlosta et al., 2017; Tempelaar et al., 2018), but little comparable research has been performed in the EAP context.

Recent studies have drawn attention to the lack of standardization of the tools used in learning analytics, which has limited research and made it difficult to compare studies (Hwang et al., 2017; Zhang et al., 2018). Many questions remain regarding the potential of learner analytics to assist

language teachers and administrators in analyzing language acquisition at scale to improve the delivery of course materials and thereby promote academic success. There has been little analysis of the performance of EAP learners in higher education contexts on multimodal quizzes in blended courses using an LMS such as Blackboard over multiple academic semesters. The present quantitative study was accordingly designed to answer the following research questions (RQs):

RQ1: How do students engage in multimodal formative assessments?

RQ2: To what extent do multimodal formative assessments predict EAP students’ performance on final assessments?

Methodology

In this learning analytics study, we aimed to understand students’ engagement with multimodal formative assessments and explore the predictive power of those assessments. We used a five-step procedure to conduct the study (Figure 1). First, we retrieved data from EAP courses across seven cohorts with 7,815 students. This data set included 18 variables, including the final grade, results of 13 quizzes and timestamps of student activities (e.g., recording the time when an activity was completed). We processed the data from an event-based file in which each entry was an event log to a student-based file in which each entry represented a student. We then examined the missing data cases and cleaned the data using a listwise approach. To answer RQ 1, we computed the students’ average percentage of completion and average starting day. To answer RQ 2, we adopted logistic regression to determine the probability of students’ being at risk (grade C or below). The dependent variable was the students’ risk level, which was derived from their final grade (on a ranking scale from 0 to 4.5 with “1” representing 0.0 to 2.5 and denoting “at risk” and “0” representing 2.5 to 4.5 and denoting “not at risk”). All of the quiz results were entered

into the prediction model as the independent variables. The McFadden's p^2 was .17, which is close to .2 (indicating an extremely good fit), and the accuracy of the logistic regression equation was 80%, which was comparable to that achieved in previous studies. This model was validated with a 10-fold cross-validation that confirmed the robustness of the model with a fair agreement in the kappa value.

Participants and Research Context

The data that we present in this study were collected at the largest public university in Hong Kong. Specifically, we collected data from students who were undertaking a compulsory first-year EAP course that focuses primarily on academic writing and presentation skills. Approximately 2,500 to 3,000 individuals participate in the course annually. After processing and cleansing the data, there were 7,815 students in the data set across seven cohorts. The course is delivered over 13 weeks. One multimodal formative learning assessment must be completed by students weekly. Being an assessment for learning, this assessment is designed to show students what they are learning and what their learning goals should be. Therefore, each assessment pertains to an important area of their studies (Table 1).

The quizzes are made available on the Blackboard platform, which is a virtual learning environment. The students are required to complete a total of 50% of the questions from all of the quizzes. Failure to do so results in an overall grade reduction (ranging from 0.5 [e.g., B to C+] to 1 grade [e.g., B to C]). There are no bonus points for students completing more than 50% of the questions. Students may take the quiz multiple times if they wish to do so. The online content for each week consists of a video that focuses on the specific topic of the session, followed by a series of questions (around 10 per quiz), some with multiple-choice answers and others requiring "true" or "false" responses. These formative assessments were introduced four academic years before the data collection for the present study. There were formal quality-assurance mechanisms for teachers, students, and subject leaders in each cohort of students to review the validity and reliability of these formative assessments. No major or minor changes were made to these formative assessments after their introduction. While instrument validation was not conducted quantitatively, the questions for these quizzes should be considered valid and reliable after being used and reviewed for more than six cohorts.

Besides the formative assessments, students must complete three summative assessments in this course: a problem-solution essay, discursive essay, and academic presentation. The university uses a grading scheme that is standard in Hong Kong (Table 2), with grades ranging from 0.0 to 4.5.

Outcome Measures

The outcome measures of these studies (i.e., independent and dependent variables) were chosen based on the

Table 1. List of Multimodal Formative Assessments.

Quiz name and topic
Session 1 Quiz: Academic style
Session 2 Quiz: Referencing
Session 3 Quiz: Reading academic articles
Session 4 Quiz: Integrating sources
Session 5 Quiz: Essay writing (1)
Session 6 Quiz: Essay writing (2)
Session 7 Quiz: Fact vs opinion
Session 8 Quiz: "For" and "Against" essays
Session 9 Quiz: Editing your work
Session 10 Quiz: Academic presentations
Session 11 Quiz: Visual aids
Session 12 Quiz: Effective presentation delivery
Session 13 Quiz: Presentations

Table 2. University Grading Scheme.

Letter grade	Corresponding rating scale	Meaning
A+	4.5	Outstanding
A	4.0	
B+	3.5	
B	3.0	Good
C+	2.5	
C	2.0	Satisfactory
D+	1.5	
F	0.0	Adequate
		Fail

availability of variables on the LMS and the research focus on multimodal formative assessments. While the event logs received from the information technology department contained many variables, we were interested only in the relationship between multimodal formative assessment and assessment results. Therefore, only events indicative of that relationship were processed to constitute the data set for the study. We retrieved and processed the following variables: the final assessment grades (3) and final course grade (1), the results of the final multimodal formative assessment attempt (13), the timestamps of the first attempt at quizzes (13), and the timestamps of the final attempts at quizzes (13). We did not enter the timestamp variables (26) and final assessment grades (3) into the regression equation because the time factor was not the focus of this paper and the assessment results have direct mathematical relationships with the final course grade.

The final assessment results were recoded as a binary variable before being entered into the regression equation. This variable was designed to indicate whether a student was at risk of failing the course. While the literature does not suggest a clear numerical cut-off point for being at risk, it did suggest that the aim of identifying at-risk or struggling students is to provide them with timely assistance (Chingos & Whitehurst, 2011). Therefore, this study defines "at risk" as below 2.0 (i.e.,

Table 3. List of Variables.

Variables (No. of variables)	Variables type (range)
Final course grade (1)	Rating scale (0.0–4.5)
Quiz scores _{first attempt} (13)	Continuous variable (0–100)
Completion rate of multimodal formative assessment (1)	Continuous variable (0–100)
Final assessment grades (3)	
Timestamp _{first attempt} (13)	MS excel serial number – Continuous variable
Timestamp _{last attempt} (13)	MS excel serial number – Continuous variable
Risk (1)	Derived from final course grade
Yes	“1” = 0.0–2.0
No	“0” = ≥ 2.5

below a C grade). Identifying students with grades of C or less can help teachers intervene in a timely fashion.

A summary of the outcome measures is provided in Table 3.

Data Analysis

To answer RQ 1, we used R (version 4.0.0) to compute computed the average of the completion rate of the quizzes and the timestamps for the first and final attempts at quizzes for each student. The timestamps provide a direct indication of engagement by showing when students start and stop doing the quizzes, whereas the completion rate measures engagement indirectly by indicating how much of the quiz students have completed with a given passing rate. Even though the average is a basic descriptive statistics indicator, it can provide “a clear overview of the results at a glance” (Woodrow, 2014, p. 50). Thus, the averages should provide meaningful insights into student engagement.

To answer RQ 2, we adopted direct logistic regression analysis with R (version 4.0.0) to “predict a discrete outcome from a set of variables” (Tabachnick & Fidell, 2013, p. 439). I entered all of the predictors (scores for the 13 quizzes) simultaneously as independent variables and the risk as the dependent variable. The goodness-of-fit was measured with McFadden’s p^2 , with a p^2 between .2 and .4 being considered “indicative of extremely good model fit” (Louviere et al., 2000, p. 54). To validate the model, 10-fold cross-validation was conducted to confirm the model’s robustness, similar to the design for the student outcome prediction study (Zacharis, 2015). As the current study focused on the predictive power of multimodal formative assessments rather than on establishing a model, we decided not to eliminate predictors even if they were not statistically significant predictors as suggested by the Wald test. The overall accuracy rate and Cohen’s kappa, supported by other relevant indicators (e.g., false positives and false negatives), were crucial in evaluating the accuracy of the prediction. The overall accuracy rate was simply the percentage of accurate predictions out of the

Table 4. Overview of Completion Rate.

Range of completion rate (%)	Frequency	Percentage of students
0.00	209	2.67
0.01–25.00	175	2.24
25.00–49.99	312	3.99
50.00–74.99	4,279	54.75
75.00–100.00	2,840	36.34

total number of cases. According to Ramasubramanian and Singh (2017), there is no widely accepted rate for prediction accuracy, which depends on the context. Cohen’s kappa explores “the relationship between observed accuracy and expected accuracy” (p. 462) and falls between 0 and 1.

Results

General Engagement Pattern of Students

The purpose of our first RQ was to investigate how students had engaged with the multimodal formative assessment tasks. Generally speaking, the students engaged extensively with the tasks (see Table 4 for an overview of their scores). While the required completion rate was only 50%, the students’ average completion score was 66.8%. This shows that many students did not stop working on a quiz after achieving the required completion rate. They continued to engage with it despite the lack of incentives in the assessments. It is also encouraging to note that 36.3% of the students achieved scores far beyond the requirement (a completion rate of 75% or higher). These results indicate that the students invested considerable effort in their multimodal assessment tasks and endeavored to perform well.

On average, the students undertook a quiz for the first time 59.8 days into the semester (i.e., the middle of week 8 in a 13-week semester). The average time difference between the students’ first and last attempts at taking a quiz was 1 hour. Thus, they started and ended the assessment tasks on the same day and did not revisit them. On average, 8 of the 13 tasks were completed only once (i.e., the first attempt was the only attempt). The students, on average, completed four tasks more than once and never attempted four tasks.

Predicting Outcomes With Multimodal Formative Assessment

Our aim in posing the second RQ was to develop a prediction model. With it, we hoped that we would be able to predict final student outcomes based on each student’s performance on multimodal formative assessments. In undertaking this exploration, we decided to identify at-risk students: those who obtained grades of C (satisfactory) or below. To do so,

Table 5. Confusion Matrix.

		Actual class	
		Positive (%)	Negative (%)
Predicted class	Positive	7.13 True positive	2.16 False positive
	Negative	16.69 False negative	74.02 True negative

we considered the students' performances on their first attempts at the tasks. The proposed prediction equation was:

$$\begin{aligned}
 \text{Risk (Yes / NO)} &= \text{Quiz 1 Score}_{\text{first}} + \text{Quiz 2 Score}_{\text{first attempt}} \\
 &+ \text{Quiz 3 Score}_{\text{first attempt}} + \text{Quiz 4 Score}_{\text{first attempt}} \\
 &+ \text{Quiz 5 Score}_{\text{first attempt}} + \text{Quiz 6 Score}_{\text{first attempt}} \\
 &+ \text{Quiz 7 Score}_{\text{first attempt}} + \text{Quiz 8 Score}_{\text{first attempt}} \\
 &+ \text{Quiz 9 Score}_{\text{first attempt}} + \text{Quiz 10 Score}_{\text{first attempt}} \\
 &+ \text{Quiz 11 Score}_{\text{first attempt}} + \text{Quiz 12 Score}_{\text{first attempt}} \\
 &+ \text{Quiz 13 Score}_{\text{first attempt}}
 \end{aligned}$$

We have detailed the prediction equation model in Table 5. McFadden's p^2 of the model was .172, which is considered an extremely good model fit after rounding up to .2 (Louviere et al., 2000). Ten-fold cross-validation showed the overall prediction accuracy to be 80.7% (81.15% for the original data set without the validations; Table 5), meaning that four in five students could be identified correctly as being not at risk based on their performance on their first quiz attempt. The scholarship to date has suggested that the rate of accuracy depends on the specific context in which research is conducted; there is no consensus on what constitutes an appropriate accuracy cut-off rate. Our 80.7% accuracy in this study seems acceptable as it is comparable to that of similar studies (e.g., 81.3% for Zacharis, 2015). The kappa value is .34, which is considered a fair agreement between observed and expected accuracy (Ramasubramanian & Singh, 2017, p. 462). However, it should be noted that 16.69% of the misclassified cases were false negatives: those who, despite being at risk, were not identified as such. This is an alarming result and warrants further discussion.

We entered all 13 of the students' assessment scores into the proposed model. Only Quizzes 11, 12, and 13 were not statistically significant predictors (i.e., $p > .05$) as indicated by the results of the Wald test to which we subjected the quiz outcomes (Table 6). As it is hard to interpret the strength

of coefficients in a logistic regression equation, we created some hypothesized scenarios and incorporated them into our prediction model (Table 7). First, we computed the average scores for each of the tasks and entered the results into our model to determine the probability of students' being at risk if they managed to attain the average score for all of the quizzes. We found the probability to be 20.8% (i.e., relatively low). We then entered another 13 hypothesized cases into the model. In each of them, one score (across 13 scores for all quizzes) was 40 points below the average score (out of a possible total score of 100) while all the other scores matched or surpassed the average. We were consequently able to examine the impact of individual quizzes on the at-risk probability ratings. Our investigations determined that Quizzes 1 (Academic Style) and 4 (Integrating Sources) helped students reduce the probability of being at risk in the course. Thus, the students' scores for those two quizzes had the most significant impact on their academic outcomes: students who did not perform well on those quizzes had a higher probability of being at risk. The only quizzes that were not significant predictors were related to giving presentations (i.e., Quiz 11–13), except Quiz 10.

False Negative Investigation

We conducted further analyses to understand why some at-risk students could not be identified. At the outset, it must be acknowledged that the process of predicting student outcomes may not be as straightforward as generating predictions in other scientific disciplines—for example, predicting the 5-year survival rate of cancer patients. Teaching and learning in higher education involve complex processes; their success relies on the delivery skills of teachers, the psychology and demography of students, and even on students' attitudes toward quizzes such as those used in the first-year EAP course. Consequently, a relatively high false-negative (or error) rate should come as no great surprise.

We compared numerous indicators between false-negative and true-positive cases as well as between false-negative and true-negative cases. As expected, the average quiz scores associated with the false-negative cases were higher than those related to the true-positive cases (Table 8). However, they were slightly lower than those of the true-negative cases (that is to say, the students who were not deemed to be at risk) (Table 9). Our further examination of the correlation matrix corroborated this finding (Table 10); the average result for the correlation coefficient between quiz scores and the assessment final scores for true-positive cases was .31 (and as high as .50); for false-negative cases, the result was .003 (and as high as .16). In other words, the students who had false-negative results worked hard but simply did not attain good scores in their summative assessments.

Table 6. Coefficients of Logistics Regression Equation.

	Estimate ^a	SE	z value	p (> z)
(Intercept)	2.37	0.13	18.51	.00**
Session 1 Quiz: Academic style	-0.01	0.00	-6.59	.00**
Session 2 Quiz: Referencing	-0.00	0.00	-2.48	.01*
Session 3 Quiz: Reading academic articles	-0.01	0.00	-3.41	.00**
Session 4 Quiz: Integrating sources	-0.01	0.00	-4.54	.00**
Session 5 Quiz: Essay writing (1)	-0.00	0.00	-3.52	.00**
Session 6 Quiz: Essay writing (2)	-0.01	0.00	-3.95	.00**
Session 7 Quiz: Fact vs opinion	-0.01	0.00	-5.13	.00**
Session 8 Quiz: “For” and “Against” essays	-0.01	0.00	-3.84	.00**
Session 9 Quiz: Editing your work	-0.00	0.00	-3.41	.00**
Session 10 Quiz: Academic presentations	-0.00	0.00	-3.06	.00**
Session 11 Quiz: Visual aids	-0.00	0.00	-0.94	.35
Session 12 Quiz: Effective presentation delivery	-0.00	0.00	-1.70	.09
Session 13 Quiz: Presentations	-0.00	0.00	-0.36	.72

Note. Other information: Null deviance: 8,579.6 on 781 degrees of freedom; residual deviance: 7,106.0 on 7,801 degrees of freedom; AIC: 7,134.

^aPlease note that none of the estimates and standard errors resulted in absolute zero. However, the numbers were too small to be shown in two decimal places.

* $p < .05$. ** $p < .01$.

Table 7. Impact of Decrease in Multimodal Formative Assessment Scores.

Multimodal formative assessments	Probability of being “at risk” (%)	Change of probability (average minus 40) (%)
Baseline probability (Achieving average score in all quizzes)	20.8	—
Session 1 Quiz: Academic style	28.3	+7.5
Session 4 Quiz: Integrating sources	26.0	+5.1
Session 7 Quiz: Fact vs opinion	25.4	+4.6
Session 6 Quiz: Essay writing (2)	24.9	+4.1
Session 3 Quiz: Reading academic articles	24.5	+3.7
Session 8 Quiz: For & against essays	24.5	+3.6
Session 5 Quiz: Essay writing (1)	24.2	+3.3
Session 2 Quiz: Referencing	23.7	+2.9
Session 9 Quiz: Editing your work	23.8	+2.9
Session 10 Quiz: Academic presentations	23.1	+2.3
Session 12 Quiz: Effective presentation delivery	22.4%	+1.6
Session 11 Quiz: Visual aids	21.7	+0.9
Session 13 Quiz: Presentations	21.1	+0.3

Table 8. Comparison Between False Negatives and True Positives (Completion Rates).

	<i>n</i> ^c	<i>M</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i> -Value	Effect size Cohen’s <i>d</i> ^d
False negative ^a	1,304	58.83 (13.30)	48.27 (563)	** $.00$	2.88 (large)
True positive ^b	348	21.24 (12.80)			

^aFalse negative: Students who were at risk but could not be identified as such.

^bTrue positive: Students who were at risk and identified as being at risk.

^cSome students did not complete any quizzes so their completion rates were excluded. Otherwise, the completion rate would have been underestimated.

^dThe cutoff points used were based on Cohen (1988).

** $p < .01$

Table 9. Comparison Between False Negatives and True Negatives (Completion Rates).

	<i>n</i> ^c	<i>M</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i> -Value	Effect size Cohen’s <i>d</i> ^d
False negative ^a	1,304	58.83 (13.30)	-14.34 (1,936)	** $.00$	0.44 (Small)
True negative ^b	5,785	64.67 (13.32)			

^aFalse negative: Students who were at risk but could not be identified as such.

^bTrue negative: Students who were at risk and identified as being at risk.

^cSome students did not complete any quizzes so their completion rates were excluded. Otherwise, the completion rates would have been underestimated.

^dThe cutoff points used were based on Cohen (1988).

** $p < .01$.

Table 10. Correlation Between Assessment Final Scores and Formative Quiz Scores (Summary).

	Average	SD	Minimum	Maximum
False negative ^a	0.003	0.05	-0.14	0.16
True positive ^b	0.311	0.08	0.16	0.50

^aFalse negative: Students who were at risk but could not be identified as such.

^bTrue positive: Students who were at risk and identified as being at risk.

Table 11. Independent Sample *t*-test Between False Negatives and True Positives (Start Day).

	<i>n</i> ^c	<i>M</i> (<i>SD</i>)	<i>t</i> (<i>df</i>)	<i>p</i> -Value	Effect size Cohen's <i>d</i> ^d
False negative ^a	1,281	66.19 (13.30)	8.92 (599)	** .00	1.23 (large)
True positive ^b	339	49.86 (13.32)			

^aFalse negative: Students who were at risk but could not be identified as such.

^bTrue positive: Students who were at risk and identified as being at risk.

^cSome students did not complete any quizzes, so their entries were excluded. Otherwise, the average start day would have been inaccurate.

^dThe cutoff points used were based on Cohen (1988).

***p* < .01.

False-negative and true-negative cases tended to have similar quiz scores, but true-positive cases were markedly different. To understand the extent to which students had undertaken formative tasks, we conducted another round of analysis. This time, we focused on the students' number of attempts at taking the quizzes. On average, true at-risk students failed to complete 9 or more of the 13 tasks, completed 2 tasks only once, and undertook only 1 of the tasks more than once. In contrast, the false-negative cases completed nine tasks once and didn't attempt only two.

The most significant difference between cases was in completion times. On average, the false-negative students started to engage with the quizzes at a later stage (on day 66 of the semester), whereas the true at-risk students started earlier (on day 55). The true at-risk students were, in this way, similar to students who were not at risk. An independent *t*-test confirmed that the differences between the false-negative and true-positive results were statistically significant (*p* < .001; Table 11).

To summarize the findings regarding the false-negative cases, there seems to have been a group of at-risk students who, despite managing to complete the formative quizzes successfully, could not obtain good scores in their summative assessments. This group of at-risk students performed as well as their not-at-risk counterparts in the formative tasks but, notably, obtained these scores later than all of the other students. While there are few direct ways in which

our prediction model could be improved, we may be able to complement it by carefully observing students' performances on their first assessments to spot any strange or concerning results at an early stage as well as carefully noting when students begin to engage with the quizzes.

Discussion and Conclusion

The purpose of this study—conducted in the context of a large-scale EAP course at a higher education institution in Hong Kong—was twofold. Firstly, we aimed to investigate students' levels of engagement with multimodal formative assessments. Secondly, we wished to explore the extent to which students' levels of engagement in such assessments could predict their performance on their final assessments. Our findings have several implications. The results indicated that the students engaged extensively with the quizzes, and the process did not demand much time or effort from them. This is a positive finding. Each quiz focused explicitly on an area of study that the students were required to master in the course and demonstrate in their assignments. Even though doing well on these multiple-choice or true-false quizzes may not help students write an essay well, we found that students seemed to be able to apply the concepts learnt in course assignments, which could explain the predictive power of the quizzes. The quizzes were, as such, set at appropriate levels of language and skill. In this study, we have provided credible evidence that multimodal formative assessments can enable us to engage students and facilitate learning in a language-acquisition context. Identifying academically at-risk students at an early stage through the use of LMS data enhances our potential as educators.

Additionally, in the context of this study, we were able to identify the quizzes that had the strongest impacts on the students' final course outcomes. The quizzes that pertained to core parts of the course or were seen as being useful (namely, Academic Style and Integrating Sources) were more likely to be completed by students. Likewise, the students targeted quizzes that related to specific skills—such as content, organization, and referencing—that they would need to demonstrate to attain certain grades; the students who performed poorly on these types of quizzes were more likely to be academically at risk. By the same logic, the students were less likely to complete the quizzes that they did not deem to be essential or immediately relevant to their studies; they did not prioritize the completion of such quizzes and sometimes did not complete them at all.

All of these findings were highly significant in helping us to identify students in the EAP course who was academically at risk. It should be noted that the students rarely completed the quizzes relating to presentation skills and language (i.e., the use of English). Even though these quizzes pertained to

language skills and aspects of language that was assessed in the final course assignment (Delivering an Academic Presentation) and constituted a major part of the overall course weighting (presentation 40% and language 29%), the students may have completed similar presentation and language assessments in secondary school and therefore didn't consider the topic new.

These results help us to understand how multimodal formative quizzes can facilitate student success and provide the type of content that L2 learners consider valuable in their independent study. The development of reliable prediction models will help teachers and course administrators gain insights into student behavior and learning progress so they can provide actionable, regular feedback in EAP programs.

The results of this study have several notable limitations, especially in relation to the relatively high proportion of false negatives. Certain aspects of academic performance cannot be predicted, such as student motivation to succeed in a compulsory EAP course and absence rates due to timetabling issues (EAP sessions often take place in the early morning or late evening and students are often juggling part-time jobs or heavy course loads). Future researchers should focus on gathering rich information by using both questionnaires and interviews to gather details that complement the LMS data. In doing so, they will be able to gain and share important insights into students' perceptions of the interfaces and designs of multimodal formative quizzes, learners' levels and sources of motivation, and how LMS platforms such as Blackboard can enhance the quality of course delivery. Such knowledge would assist in the development of future blended learning formative assessments as well as early prediction models. In addition, scholars could investigate how students complete quizzes in relation to their timetabled classes and their self-reported levels of motivation for completing the course.

Despite its limitations, this study identified a meaningful correlation between the ways in which—and the extent to which—EAP students completed multimodal formative assessments and their final course grades. The findings of this study confirm that educators will be better placed to identify and monitor students' levels of progress, knowledge, and skill if they encourage them to engage with blended learning through an LMS (Hinkelman, 2018). Moreover, the high levels of engagement with the quizzes exhibited by the students demonstrate that well-designed multimodal quizzes can provide authentic and engaging L2 learning experiences (Park et al., 2016) and foster effective learning climates (Boelens et al., 2015). The implications of these results are significant for the future design of blended courses. Course developers and teachers need to collaborate to design suitable educational activities that promote relevant language skills. Such activities must not only be seen as immediately applicable to students but also of demonstrable benefit to them in their broader higher education

studies. A key challenge for all those who are involved will be to stay up-to-date with technological developments to ensure that the blended learning activities continue to be engaging and attractive to students.

Having access to learner analytics allows teachers and course administrators to affect personalized interventions with each of their students. Moreover, such information enables them to adjust the content or format of their blended learning modules to meet the needs and interests of specific student cohorts (Thomas et al., 2017; Verbert et al., 2013; Vuorikari & Castano Munoz, 2016). As teaching and learning increasingly involve blended pedagogical approaches, combining pedagogical theory with the use of analytics is vital (Rienties et al., 2018). Educators are to understand and meet the needs of EAP students and meaningfully analyze their work to improve their learning experiences. In this study, we have shown that learner analytics can help educators to predict students' academic outcomes. We hope that these findings will inform pedagogical practice and the design of learning activities, leading to greater academic success.

Author Contributions

All authors contributed to the research design and development, discussed the results, and contributed to the final manuscript.

Availability of Data and Materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code Availability

The code used during the current study are available from the corresponding author on reasonable request.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Lucas Kohnke  <https://orcid.org/0000-0001-6717-5719>

Dennis Fount  <https://orcid.org/0000-0002-6769-0582>

References

- Atherton, M., Shah, M., Vazquez, J., Griffiths, Z., Jackson, B., & Burgess, C. (2017). Using learning analytics to assess students engagement and academic outcomes in open access enabling programme. *Open Learning: The Journal of Open and Distance Learning*, 32(2), 119–136. <https://doi.org/10.1080/02680513.2017.1309646>

- Baker, R. S. J., Corbett, A. T., Koedinger, K. R., Evenson, S., Roll, I., Wagner, A. Z., Naim, M., Raspat, J., Baker, D. J., & Beck, J. E. (2006). Adapting to when students game an intelligent tutoring system. In M. Ikeda, K. D. Ashely, & T.-W. Chan (Eds.), *Proceedings of the 8th international conference on intelligent tutoring systems, LNCS* (Vol 4–53, pp. 392–401). Springer.
- Banditvilai, C. (2016). Enhancing students' language skills through blended learning. *Electronic Journal of E-Learning*, 14(3), 220–229. <https://files.eric.ed.gov/fulltext/EJ1107134.pdf>
- Baumgartner, K. (2015). Strategies for hybrid course design in second-year German. *Teaching German*, 48(2), 211–228. <https://doi.org/10.1111/tger.10196>
- Boelens, R., De Weaver, B., & Voet, M. (2015). Four key challenges to the design of blended learning: A systematic literature review. *Educational Research Review*, 22, 1–18. <https://doi.org/10.1016/j.edurev.2017.06.001>
- Bonks, C. J., Kim, K.-J., & Zeng, T. (2005). *Future directions of blended learning in higher education and workplace learning settings*. <https://publicationshare.com/c39-Bonk-Kim-and-Zeng-on-Future-Directions-Bonk-Kim-and-Zeng-updated.pdf>
- Brown, M. G. (2016). Blended instructional practice: A review of the empirical literature on instructors' adoption and use of online tools in face-to-face teaching. *The Internet and Higher Education*, 31, 1–10. <https://doi.org/10.1016/j.ihe-duc.2016.05.001>
- Chapelle, C. A., & Sauro, S. (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell.
- Chen, J., & Fong, D. (2020). A motivational story in Hong Kong: Generating goals for language learners and blended learning designers from a mixed-method learning analytics approach in English for academic purposes. In M. Freiermuth & Z. Nourollah (Eds.), *Technology and the psychology of second language learners and users* (pp. 491–516). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-34212-8_19
- Chingos, M. M., & Whitehurst, G. J. (2011). *Choosing blindly: Instructional materials, teacher effectiveness, and the common core*. The Brookings Institution.
- Choi, S. P. M., Lam, S. S., Cheong Li, K., & Wong, B. T. M. (2018). Learning analytics at low cost: At-risk student prediction with clicker data and systematic proactive interventions. *Journal of Educational Technology & Society*, 21(2), 273–290.
- Cohen, D., & Sasson, I. (2016). Online quizzes in a virtual learning environment as a tool for formative assessment. *Journal of Technology and Science Education*, 6(3), 188–208. <https://doi.org/10.3926/jotse.217>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge Academic.
- Cole, A. W., Lennon, L., & Weber, N. (2019). Student perceptions of online active learning practices and online learning climate predict online engagement. *Interactive Learning Environments*, 29(5), 1–15. <https://doi.org/10.1080/10494820.2019.1619593>
- Conati, C., & Maclaren, H. (2009). Empirically building and evaluating a probabilistic model of user affect. *User Modeling and User-Adapted Interaction*, 19(3), 267–303. <https://doi.org/10.1007/s11257-009-9062-8>
- Cook, B. R., & Babon, A. (2017). Active learning through online quizzes: Better learning and less (busy) work. *Journal of Geography in Higher Education*, 41(1), 24–38. <https://doi.org/10.1080/03098265.2016.1185772>
- Coskuncay, F., & Ozkan, S. (2013). A model for instructors' adoption of learning management systems: Empirical validation in higher education context. *Turkish Journal of Educational Technology*, 12(2), 13–25. <https://files.eric.ed.gov/fulltext/EJ1015409.pdf>
- Dahlstrom, D., Brooks, D. C., & Bichsel, J. (2014). The current ecosystem of learning management systems in higher education: student, faculty, and IT perspectives. *EDUCAUSE*. <https://library.educase.edu/-/media/files/library/2014/9/ers1414-pdf.pdf>
- Dixon, E., & Fuchs, C. (2015). Face to face, online, or MOOC – How the format impacts content, objectives, assignments, and assessments. In E. Dixon & M. Thomas (Eds.), *Researching language learner interactions online: From social media to MOOCs* (pp. 89–105). The Computer Assisted Language Instruction Consortium (CALICO).
- Elwood, J., & Klenowski, V. (2002). Creating communities of shared practice: The challenges of assessment use in learning and teaching. *Assessment & Evaluation in Higher Education*, 27(3), 243–256. <https://doi.org/10.1080/02602930220138606>
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning (IJTEL)*, 4(5/6), 304–317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Flowerdew, J., & Peacock, M. (2001). *Research perspectives on English for academic purposes*. Cambridge University Press.
- Gadbois, M., & Quildon, D. (2013). Using blended learning as an innovative delivery model for an in-house language program. *Canadian Journal of University Continuing Education*, 39(2), 1–9. <https://doi.org/10.21225/D5TS40>
- Gamage, S. H. P. W., Ayres, J. R., Behrend, M. B., & Smith, E. J. (2019). Optimising moodle quizzes for online assessments. *International Journal of Stem Education*, 6, Article 27. <https://doi.org/10.1186/s40594-019-0181-4>
- Gardner, J. (2012). *Assessment and learning*. SAGE. <https://doi.org/10.4135/9781446250808>
- Garrison, D. R., & Vaughan, N. D. (2013). Institutional change and leadership associated with blended learning innovation: Two case studies. *The Internet and Higher Education*, 18, 24–28. <https://doi.org/10.1016/j.ihe-duc.2012.09.001>
- Ghazali, A. F. (2018). Independent language learners in TESOL. *The TESOL Encyclopedia of English language teaching*. Wiley. <https://doi.org/10.1002/9781118784235.eelt0202>
- Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2014). Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27(1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Graham, C. R. (2006). Blended learning systems: Definitions, current trends, and future direction. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: Global perspectives, local designs* (pp. 3–21). Pfeiffer.
- Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *The Internet and Higher Education*, 18, 4–14. <https://doi.org/10.1016/j.ihe-duc.2012.09.003>
- Gruba, P., Hinkelman, D., & Cardenas-Claros, M. (2019). New technologies, blended learning and the “flipped classroom” in

- ELT. In G. Hall (Ed.). *The Routledge handbook of English language teaching* (pp. 135–149). Routledge.
- Hinkelman, D. (2018). *Blended technologies in second language classrooms*. Palgrave Macmillan.
- Hlosta, M., Zdrahal, Z., & Zendulka, J. (2017, March 13–17). *Ouroboros: Early identification of at-risk students without models based on legacy data* [Conference session]. LAK17 – Seventh International Learning Analytics & Knowledge Conference, Vancouver, BC, Canada, pp. 6–15. <http://oro.open.ac.uk/49731/1/paper.pdf>
- Hockly, N. (2018). Blended learning. *ELT Journal*, 72(1), 97–101. <https://doi.org/10.1093/elt/ccx058>
- Hwang, G.-J., Chu, H.-C., & Yin, C. (2017). Objectives, methodologies and research issues of learning analytics. *Interactive Learning Environments*, 25(2), 143–146. <http://www.jstor.org/stable/26388386>
- Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.
- Hyland, K. (2018). Sympathy for the devil? A defence of EAP. *Language Teaching*, 51(3), 383–399. <https://doi.org/10.1017/S0261444818000101>
- Joosten, T. M., Barth, D., Harness, L., & Weber, N. L. (2014). The impact of instructional development and training for blended teaching on course effectiveness. *Blended Learning: Research Perspectives*, 2, 173–189. <https://doi.org/10.30813/jelc.v9i1.1449>
- Kohnke, L. (2021). Professional development and ICT: English language teachers' voices. *Online Learning Journal*, 25(2), 36–53. <https://doi.org/10.24059/olj.v25i2.2228>
- Li, L. (2017). *New technologies and language learning*. Palgrave.
- Louviere, J., Hensher, D., Swait, J., & Adamowicz, W. (2000). Choosing a choice model. In J. J. Louviere, D. A. Hensher, & J. D. Swait (Eds.), *Stated choice methods: Analysis and applications* (pp. 34–82). Cambridge University Press. <https://doi.org/10.1017/CBO9780511753831.003>
- Lughmani, D. S., & Fong, D. (2021). Metacognitive strategy-focused instruction in a writing across the curriculum programme in a Hong Kong university: The impact on writing performance. In B. L. Reynolds & M. F. Teng (Eds.), *Innovative approaches in teaching English writing to Chinese speakers* (pp. 145–178). De Gruyter Mouton. <https://doi.org/10.1515/9781501512643-007>
- Moradimokhles, H., & Hwang, G.-J. (2020). The effect of online vs. blended learning in developing English language skills by nursing student: An experimental study. *Interactive Learning Environments*. Advance online publication. <https://doi.org/10.1080/10494820.2020.1739079>
- Owston, R., & York, D. N. (2018). The nagging question when designing blended courses: Does the proportion of time devoted to online activities matter. *The Internet and Higher Education*, 36, 22–32. <https://doi.org/10.1016/j.iheduc.2017.09.001>
- Park, Y., Yu, J. H., & Jo, I.-H. (2016). Clustering blended learning courses by online behavior data: A case study in a Korean higher education institute. *The Internet and Higher Education*, 29, 1–11. <https://doi.org/10.1016/j.iheduc.2015.11.001>
- Popham, W. J. (2008). *Transformative assessment*. Association for Supervision and Curriculum Development.
- Porter, W. W., Graham, C. R., Bodily, R. G., & Sandberg, D. S. (2016). A qualitative analysis of institutional drivers and barriers to blended learning adoption in higher education. *The Internet and Higher Education*, 28, 17–27. <https://doi.org/10.1016/j.iheduc.2015.08.003>
- Ramasubramanian, K., & Singh, A. (2017). *Machine learning using R*. Springer.
- Reffay, C., & Chanier, T. (2003). How social network analysis can help to measure cohesion in collaborative distance-learning. In B. Wasson, S. Ludvigsen, & U. Hoppe (Eds.), *Designing for change in networked learning environments. Computer-Supported Collaborative Learning* (pp. 343–352). Springer. https://doi.org/10.1007/978-94-017-0195-2_42
- Rienties, B., Lewis, T., McFarlane, R., Nguyen, Q., & Toetenel, L. (2018). Analytics in online and offline language environments: The role of learning design to understand student online engagement. *Computer Assisted Language Learning*, 31(3), 273–293. <https://doi.org/10.1080/09588221.2017.1401548>
- Ross, B., & Gage, K. (2006). Global perspectives on blended learning: Insight from WebCT and our customers in higher education. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: Global perspectives, local designs* (pp. 155–168). Pfeiffer.
- Rubio, F. (2015). The role of interaction in MOOCs and traditional technology-enhanced language courses. In E. Dixon & M. Thomas (Eds.), *Researching language learner interaction online: From social media to MOOCs* (pp. 63–88). CALICO.
- Rubio, F., Thomas, J. M., & Li, Q. (2018). The role of teaching presence and student participation in Spanish blended courses. *Computer Assisted Language Learning*, 31(3), 226–250. <https://doi.org/10.1080/09588221.2017.1372481>
- Scheffel, M., Niemann, K., Pardo, A., Leony, D., Friedrich, M., Schmidt, K., Wolpers, M., & Kloos, C. D. (2001). Usage pattern recognition in student activities. In C. D. Kloos, D. Gillet, R. M. Crespo Garcia, F. Wild, M. Wolpers (Eds.), *Proceedings of EC-TEL'11* (pp. 341–355). Springer-Verlag. https://doi.org/10.1007/978-3-642-23985-4_27
- Staker, H., & Horn, M. B. (2012). *Classifying K-12 Blended Learning*. Innosight Institute.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Pearson.
- Tempelaar, D. T., Nguyen, Q., & Rienties, B. (2020). Learning feedback based on dispositional learning analytics. In M. Virvou, E. Alepis, G. Tsihrintzis, & L. Jain (Eds.), *Machine Learning Paradigms* (pp. 69–89). Springer.
- Tempelaar, D. T., Rienties, B., & Giesbers, B. (2011). Student learning preferences in a blended learning environment: Investigating the relationship between tool use and learning approaches. In P. Van den Bossche, W. H. Gijsselaers, & R. G. Milter (Eds.), *Building learning experiences in a changing world* (pp. 195–212). Springer.
- Tempelaar, D. T., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior*, 78, 408–420. <https://doi.org/10.1016/j.chb.2017.08.010>
- Terauchi, H., Noguchi, J., & Tajino, A. (2019). *Towards a new paradigm for English language teaching: English for specific purposes in Asia and beyond*. Routledge.
- Thomas, M., Reinders, H., & Gelan, A. (2017). Learning analytics in online language learning: Challenges and future directions. In L. Wong & K. Hyland (Eds.), *Faces of English* (pp. 197–212). Routledge.

- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dash-board applications. *American Behavioral Scientist*, 57(10), 1500–1509. <https://doi.org/10.1177/0002764213479363>
- Vuorikari, R., & Castano Munoz, J. (Eds.). (2016). *Research evidence on the use of learning analytics: Implications for education policy*. Publications Office of the European Union.
- Wallihan, R., Smith, K. G., Hormann, M. D., Donthi, R. R., Boland, K., & Mahan, J. D. (2018). Utility of intermittent online quizzes as an early warning for residents at risk of failing the pediatric board certification examination. (Report). *BMC Medical Education*, 18(1), 287. <https://doi.org/10.1186/s12909-018-1366-0>.
- Wiliam, D., & Leahy, S. (2015). *Embedding formative assessment: Practical techniques for K-12 classrooms*. Learning Sciences International.
- Wong, H. M. (2017). Implementing self-assessment in Singapore primary schools: Effects on students' perceptions of self-assessment. *Pedagogies: An International Journal*, 12(4), 391–409. <https://doi.org/10.1080/1554480X.2017.1362348>
- Woodrow, L. (2014). *Writing about quantitative research in applied linguistics*. Palgrave Macmillan.
- Yan, Z., Li, Z., Panadero, E., Yang, M., Yang, L., & Lao, H. (2021). A systematic review on factors influencing teachers' intentions and implementations regarding formative assessment. *Assessment in Education: Principles, Policy & Practice*, 28(3), 228–260. <https://doi.org/10.1080/0969594X.2021.1884042>
- Zacharis, N. Z. (2015). A multivariate approach to predicting student outcomes in web-enabled blended learning courses. *The Internet and Higher Education*, 27, 44–53. <https://doi.org/10.1016/j.iheduc.2015.05.002>
- Zhang, J. H., Zhang, Y. X., Zou, Q., & Huang, S. (2018). What learning analytics tells us: Group behavior analysis and individual learning diagnosis based on long-term and large-scale data. *Journal of Educational Technology & Society*, 21(2), 245–258. <https://www.jstor.org/stable/26388404>