CHAPTER

7

AN ASSESSMENT ANALYTICS FRAMEWORK (AAF) FOR ENHANCING STUDENTS' PROGRESS

Z. Papamitsiou, A.A. Economides

University of Macedonia, Thessaloniki, Greece

1 INTRODUCTION

Students' learning and progress are inferred from the systematic process of assessment (Erwin, 1991; Swan et al., 2006). Performance assessments, "do not offer a direct pipeline into a student's mind. […] assessment is a tool designed to observe students' behavior and produce data that can be used to draw reasonable inferences about what students know" (Pellegrino, 2013, p. 261). Current theoretical, cultural and technological developments influence teaching and learning practices, resulting in an increasingly accepted need for rethinking assessment. The motivation for automating the assessment process originates from the need to alleviate the practical problems introduced by large classes and to harness potential pedagogical benefits.

e-Assessment (or computer-based assessment, CBA; computer assisted assessment, CAA; technology-enhanced assessment, TEA) is the use of information technologies (IT) (eg, desktop computers, mobiles, web-based, etc.) to automate and facilitate assessment and feedback processes (Chatzopoulou and Economides, 2010; Sancho-Vinuesa and Escudero-Viladoms, 2012; Triantafillou et al., 2008), and is usually categorized into formative and summative. While formative assessment provides developmental and prescriptive feedback to learners on their current understanding and skills to assist them in reaching their goals, summative assessment is about establishing whether students have attained the goals set for them, and usually lead to a formal qualification or certification of a skill (Birenbaum, 1996; Economides, 2006, 2009a). Examples of e-assessment methods include portfolio assessment, rubrics, self-assessment, peer assessment (eg, Peat and Franklin, 2002), and, more recently, collaborative and social assessment (Caballé et al., 2011).

The introduction of digital technologies into education opens up new possibilities for tailored, immediate and engaging assessment experiences. Drivers for the adoption of e-assessment include perceived increases in student retention, enhanced quality of feedback, flexibility for distance learning, strategies to cope with large student/candidate numbers, objectivity in marking and more effective use

Formative Assessment, Learning Data Analytics and Gamification. http://dx.doi.org/10.1016/B978-0-12-803637-2.00007-5

of virtual learning environments (Whitelock and Brasher, 2006). Recent studies highlight the importance of continuous e-assessment for learning (Whitelock, 2011).

From a more generalized viewpoint, the use of digital technologies for assessment purposes allows for remote progress tracking and evaluation and is strongly correlated to increased student and teacher (self-)awareness regarding students' learning achievements (Leony et al., 2013; Carnoy et al., 2012). This is one of the reasons why improvement of e-assessment services has been under the microscope of learning analytics research and constitutes one of its main objectives (Chatti et al., 2012; Papamitsiou and Economides, 2014; Tempelaar et al., 2013).

Assessment analytics (AA) is not a new field of enquiry. In fact, it is a subset of the wider area of learning analytics, and by itself it is an emerging research field. Like any other context-aware system, an AA procedure monitors, tracks and records data related to the context, interprets and maps the real current state of these data, organizes them (eg, filter, classify, prioritize), uses them (eg, decide adaptations, recommend, provide feedback, guide the learner) and predicts the future state of these data (Economides, 2009b).

In a sense, AA is about applying analytic methods to multiple types of data in an effort to reveal the intelligence held in e-assessment systems. More specifically, AA attempts to shed light to how students will improve their performance by (a) making practical use of detailed assessment psychometrics data held in e-assessment systems and (b) providing feedback accordingly (MacNeill and Ellis, 2013). It is important to get appropriate adaptive and personalized feedback to both students and teachers (instructors), based on data on the user's (behavioral) model and the learning context (Chatti et al., 2012).

In other words, the main objective of AA is to efficiently and effectively support the improvement of the assessment process. This means that AA's goals target (directly) assisting students, teachers and learning administrators (and indirectly at supporting/informing other stakeholders, like parents or even the Ministry of Education). For students, AA could passively support self-awareness, self-reflection, or self-regulation or actively trigger emotional change, challenge their participation, and motivate further engagement in assessment (and/or learning) activities. In order to explicitly set out the relationship between assessment systems and the sorts of epistemic challenges students might encounter, Knight et al. (2013) discuss the relationships between learning analytics, epistemology, pedagogy and assessment. The authors associated their approach to learning analytics with that of assessment for learning (AfL), which uses continuous assessment with formative feedback to facilitate learning, in contrast to a focus on summative assessment, often through examinations.

As far as it concerns the efficiency of AA for teachers (or similarly for learning administrators), these processes could be used to facilitate the estimation of students' performances, improve the detection and target prevention for students at-risk, enhance the detection of misconceptions and gaps in students' understandings, allow for the identification of students' guessing or cheating behavior, and many more.

The landscape in the domain of AA so far is quite diverse. Indicative research examples in this domain include the selection of the most appropriate next task during adaptive testing (Barla et al., 2010), the recognition of affects and mood during self-assessment (Moridis and Economides, 2009), the determination of students' satisfaction level during mobile formative assessment (Chen and Chen, 2009), the assessment of participatory behavior, knowledge building and performance during discussion processes (Caballé et al., 2011), as well as the construction of sophisticated measures of assessment (Wilson et al., 2011; Worsley and Blikstein, 2013) and many more, which we will discuss in the next sections.

The purpose of this chapter is to provide an AA framework as a reference point suitable to address complex problems in AA systems for which no clear guidelines are available as yet.

The remainder of this chapter is organized as follows: after elaborating on the motivation to develop the proposed AA framework and explicitly setting out the research questions that need to be addressed (see Section 2), we present the methodology we adopted in this study (see Section 3). Next, we describe the framework in general terms and the central concepts it involves (see Sections 4.1 and 4.2). Then, we analyze each dimension of the framework and discuss on the research questions initially posed (see Section 4.3). Finally, we elaborate on other important issues raised during the development of the framework, concerning mostly the security/privacy and ethics during data gathering, authorization and sharing (see Section 5).

2 MOTIVATION

As stated in the introduction, along with the promise and potential embodied in e-assessment systems come many challenges, which are due to the opening up of new possibilities for more personalized, immediate and engaging assessment experiences. Some challenges are foundational and faced by all assessment designers, such as establishing the validity, reliability, precision of measurement, and fairness of the new measures. Other challenges are less trivial and target to the AA experts, such as responsiveness to real-world contexts in real-time, evaluation of complex assignments in large courses (eg, massive open online courses, MOOCs), detection of gaps and misconceptions during assessment and more. Even the broader inclusion of students in developing accurate assessment measures seems to be a challenge for AA designers. This is because previous studies on e-assessment have shown that students find the use of e-assessment more promising, credible, objective, fair, interesting, fun, fast and less difficult or stressful (Conole and Warburton, 2005; Dermo, 2009; McCann, 2010).

A deeper observation of these findings through the microscope of AA partially reveals the intrinsic capabilities of these processes to support self-reflection and self-regulation during assessment procedures. Accordingly, AA should certainly be considered as part of any teaching and learning strategy, due to their potential benefits to better understand the reasons for students' progress or failure during assessment. However, both students and teachers must be well supported in their use. Simply providing the data (in the form of a dashboard, for example) is unlikely to be effective unless students and teachers are offered training in its interpretation and accessible strategies to act upon it.

In all cases, assessment and AA designers require the appropriate tools to conceptualize and tackle design challenges. A thorough search of the literature resulted in various examples of research works examining related issues and evidencing the adoption of AA. The search did not yield any theoretical/conceptual framework for AA. Even the approach suggested by Knight et al. (2013) does not provide a framework for understanding, building and interpreting AA, since it focuses on beliefs about the nature of knowledge, for which analytics grounded in pragmatic, sociocultural theory might be well placed to explore.

In general, a theoretical framework is a visual or written representation that "explains either graphically or in narrative form the main things to be studied—the key factors, concepts or variables—and the presumed relationships among them" (Miles and Huberman, 1994, p. 18). The role of a framework is to provide a theory that will be used to interpret existing AA data and to code them for future use. In that way, a framework would explicitly validate AA results and could be used to move beyond descriptions of "what" to explanations of "why" and "how." Articulating the theoretical framework would permit the transition from simply observing and describing AA to generalizing its various aspects (Jabareen, 2009).

Taking into consideration the diversity of approaches for AA (stated in the introduction) and the need for a framework that explicitly shapes, interprets and validates AA results, in this chapter we propose a theoretical framework. The goal is to develop a conceptual representation to act as a reference point/structure for the discussion of the literature, the methodologies followed and the results from former research studies concerning AA. The framework will also act as a useful guide to understand more deeply, evaluate and design analytics for assessment. In this chapter, we associate related literature to the main concepts of AA and justify the choice of the components of the suggested framework. We identify the key points that need to be closely examined and highlight the critical dimension of AA.

3 METHODOLOGY

After reviewing the existing frameworks for learning analytics (Greller and Drachsler, 2012; Fernández-Gallego et al., 2013; Shum and Crick, 2012), learning and assessment (Economides, 2009b; Haertel et al., 2012), our chosen approach leading to the proposed framework consisted of a sequence of gathering and analysis processes. In particular, we followed an inductive and deductive inquiry methodology for conceptual mapping for sensemaking. This methodology is considered appropriate since the use of inductive analysis is recommended when there are no previous studies dealing with the phenomenon or when knowledge is fragmented. Furthermore, a deductive approach is useful if the aim is to test an earlier theory or to compare categories.

In an inductive approach, once a substantial amount of data has been collected, the next step is the identification of patterns in the data in order to develop a theory that could explain those patterns. This process includes free coding, creating categories and abstraction. The purpose of creating categories is to provide a means of describing the phenomenon, to increase understanding and to generate knowledge. Formulating categories by inductive analysis, leads to a decision, through interpretation, as to which things to put in the same category (Dey, 1993). An inductive approach starts with a set of observations and then moves from data to theory or from the specific to the general.

Moreover, deductive analysis is in general based on earlier work such as theories, models, mind maps and literature reviews and is often used in cases aiming at retesting existing data in a new context (Marshall and Rossman, 1995).

In our approach, we applied a methodology consisting of three discrete steps. More specifically:

1. We initially searched the literature for studies that report results, best practices, central issues, variable construction, measurement techniques, etc., or sparse theoretical approaches regarding AA. Then, we categorized the objectives, methods, measures and results reported in the studies we had collected, into upper classes-concepts of concern. This classification led to an early introduction of the basic general concepts (induction).
2. Creating categories is both an empirical and a conceptual challenge, as categories must be conceptually and empirically grounded (Dey, 1993). Thus, we sought to identify the relationships between these clusters and which more general questions they address, in order to shape the conceptual map. A concept map (Novak, 1981) is a formal, structured diagram showing relationships among a number of unique concepts (concept mapping).
3. Finally, in our study, we designed the framework and tested whether the collected papers fit in that schema (deduction).

4 RESULTS

4.1 INDUCTION PHASE: LITERATURE REVIEW OF ASSESSMENT ANALYTICS APPROACHES

As stated in the introduction, the landscape in the domain of AA so far is often quite dispersed. The main objectives of research examples in this domain include the provision of adaptive feedback during summative assessment, the selection of the most appropriate next task during adaptive testing (Barla et al., 2010; Papamitsiou et al., 2014), student-oriented formative assessment support in real-time (Whitelock et al., 2015) and performance assessment of real-world skills (Knight and Littleton, 2015; Sao Pedro et al., 2013). Other major issues include the recognition of affects and mood during self-assessment, the determination of students' satisfaction level during mobile formative assessment, the assessment of participatory behavior, knowledge building and performance during discussion processes, as well as the construction of sophisticated measures of assessment (eg, Caballé et al., 2011; Chen and Chen, 2009; Moridis and Economides, 2009; Wilson et al., 2011; Worsley and Blikstein, 2013).

More specifically, research on providing appropriate and adaptive feedback during summative assessment targets at adapting the next quiz item to students' abilities during computer-based testing. The combination of different classification methods for selection of the most appropriate next task based on the topic selection using course structure, on item response theory (IRT)—selection of the k-best questions with most appropriate difficulty for the particular user—and on history-based prioritization of questions (eg, not recently asked questions) (Barla et al., 2010) and classification of students' response times according to the correctness of the submitted answers and the amount of time the students remained idle (not submitting an answer) (Papamitsiou et al., 2014) have presented interesting results. Based on students' temporal engagement in summative assessment activities, the different student temporal models of behavior during testing were used aiming at adapting the next quiz item to the students' abilities, as opposed to the more complicated algorithms that make use of psychometrics. The authors also proposed different visualizations of the temporal dimension of students' behavior for increasing awareness during assessment (Papamitsiou and Economides, 2015).

However, the multiple choice questions (MCQs) traditionally used in summative assessment are in general limited in their ability to provide the necessary analyses for guiding real-time scaffolding and remediation for students. Accordingly, student-oriented formative assessment support in real-time has been a major research topic for many authors. To address this issue, approaches to real-time formative assessment have included analyses of student action logs and real-time processing of free-text in open-ended learning environments (OELEs) inside and outside of classroom (Monroy et al., 2013; Sao Pedro et al., 2013; Tempelaar et al., 2013). Furthermore, Worsley and Blikstein (2013) aimed to detect factors and define metrics that could be used primarily as formative assessment tools for sophisticated learning skills acquisition in process-oriented assessment. A combination of speech recognition with knowledge tracing was proposed by the authors as the method for multimodal assessment.

However, interpreting and assessing students' learning behavior is inherently complex; at any point in time, there may be a dozen or more "correct next steps" from which students may choose. The space of possible learning paths—mostly in OELEs—quickly becomes intractable. An AA approach to this issue has been the Model-Driven Assessments, which consist of a model of the cognitive and metacognitive processes important for completing the learning task (Segedy et al., 2013). This approach leverages the cognitive and metacognitive model in interpreting students' actions and behavior patterns (ie, sequences of actions) in terms of the cognitive and metacognitive processes defined by the model. From a different point of view, tracking and examining how students go about solving a problem step-by-step makes transaction-level assessment possible by focusing on diagnosing persistent misconceptions and knowledge gaps using transaction-level data (Davies et al., 2015).

Performance assessments are also central to AA research. Providing performance assessment of real-world skills through real-world behaviors was suggested either as an evidence-centered approach to designing a performance assessment for epistemic cognition (Knight and Littleton, 2015) or as a learner-centered approach to formatively assess the situation of a real-world learner at a given time (Okada and Tada, 2014). In the first case, the study took place in the context of complex multiple document-processing tasks in which students read and synthesize information from multiple documents. In the second study, the objective was to evaluate the performance of individual learners participating in collaboration activities in the real world, through the systematic integration of the spatiotemporal occurrence information of real-world behavior. Finding a way to combine heterogeneous factors—learners' internal situations, their external situations, and their real-world learning field—was identified as a central issue for their research.

Another approach for performance assessment led to the development of a tool that uses six computational intelligence theories according to the web-based learning portfolios of an individual learner, in order to discover useful fuzzy association rules relating to the learning performance assessment and measure students' satisfaction during mobile formative assessment (Chen and Chen, 2009). Learning portfolios provide rich information for reflecting and assessing the performances and achievements of learners, and help learners to engage in meaningful learning accordingly.

In the context of gaining insight into formative assessment procedures from the scope of students' emotional states, recent studies attempted to estimate the students' emotions (eg, boredom, confusion, delight, or frustration) during formative assessment, using sensor data (eg, data from a fixed video camera, a pressure-sensitive mouse, and a pressure-sensitive chair) (D'Mello and Graesser, 2009; Kobayashi et al., 2011; Moridis and Economides, 2009).

Assessment of collaborative and/or teamwork is also considered as a central issue for assessment analytics research (Caballé et al., 2011; Perera et al., 2009). Indicative examples in this direction include a text mining approach to automate teamwork assessment in chat data (Shibani et al., 2015) as well as the use of activity theory (Nardi, 1995) applied to the assessment of computer supported collaborative learning (CSCL) (Xing et al., 2014). In the latter case, the goal was to assess student activities by using cluster analysis to evaluate strengths and weaknesses in individual students' participation in collaborative activities. Moreover, peer assessment in the evaluation of complex assignments in large courses, as in the context of MOOCs (Vozniuk et al., 2014) was also explored. Furthermore, other studies examined the improvement of the collaborative learning process in terms of participation behavior, knowledge building, and performance in the context of learning through discussion (Caballé et al., 2011).

4.2 CONCEPT MAPPING PHASE: CLASSIFICATION OF STUDIES AND MAIN CONCEPTS

The conducted literature review has revealed a number of commonalities and differences in the proposed and explored approaches. More precisely, all of these studies take under consideration the context of the assessment procedure and explore possible ways of providing fruitful and comprehensive feedback to the students being assessed. Moreover, in all studies, the measures adopted and the purpose/scope of the assessment process (either summative or formative) are clearly and explicitly stated. Furthermore, the data gathering and analysis methods/algorithms along with the pedagogical appropriation and benefits, as well as the potential limitations, are explicitly explained. Finally, the implications of the proposed methods and approaches for potential end users (either teachers, learners, institutions, or developers) are discussed.

From the above analysis it becomes apparent that the central concepts involved in an AA framework should include the following: the context, the objectives, the scope, the methods, the instruments, the resources, the people involved, and the limitations and boundaries. Having as reference points both the Model for Learning Analytics suggested by Chatti et al. (2012) and the generic framework for Learning Analytics suggested by Greller and Drachsler (2012), these initial upper-class concepts are mapped to the categories identified from the literature review as illustrated in Fig. 7.1.

Furthermore, these categories can be abstractly assigned to and organized into five more general classes: the who, the how, the why, the what, and the when/where, illustrated in Figs. 7.2 and 7.3.

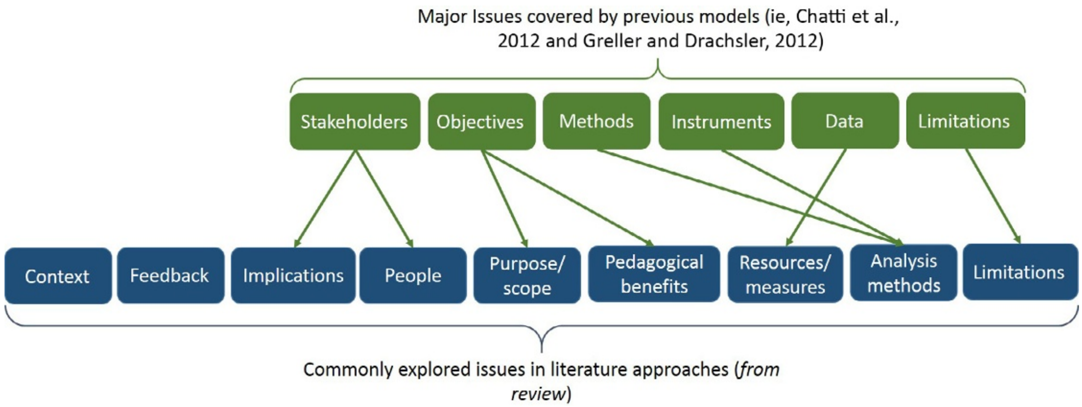

FIGURE 7.1

Mapping of the upper-class concepts to identified categories for assessment analytics.

4.3 THE ASSESSMENT ANALYTICS FRAMEWORK

As stated in the introduction, like any other context-aware system, an AA procedure monitors, tracks and records data related to the context, interprets and maps the real current state of these data, organizes these data (eg, filter, classify, prioritize), uses these data (eg, decide adaptations, recommend, provide feedback, guide the learner) and predicts the future state of these data (Economides, 2009b). Consequently, the suggested assessment analytics framework (AAF) is composed of four "blocks": *input*, *process*, *output* and *feedback* (Fig. 7.4).

The above literature review has revealed a number of commonalities and differences in the proposed and explored approaches. Based on the analysis of these studies, the input to the AA system is contextual information related to (a) *what* should be tracked and assessed (eg, measurements, assessment setting, and infrastructure), (b) *why* is the assessment necessary (eg, objectives, scope, and type of assessment), (c) *who* is the subject and receiver of the assessment (eg, learner-oriented, teacher-oriented) and (d) *when/where* the

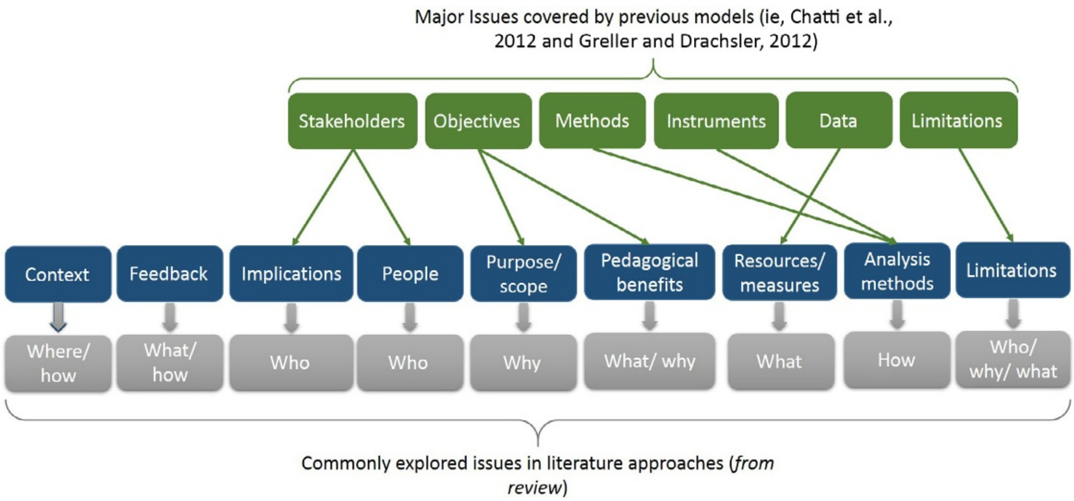

FIGURE 7.2

Abstract classification of identified categories for assessment analytics.

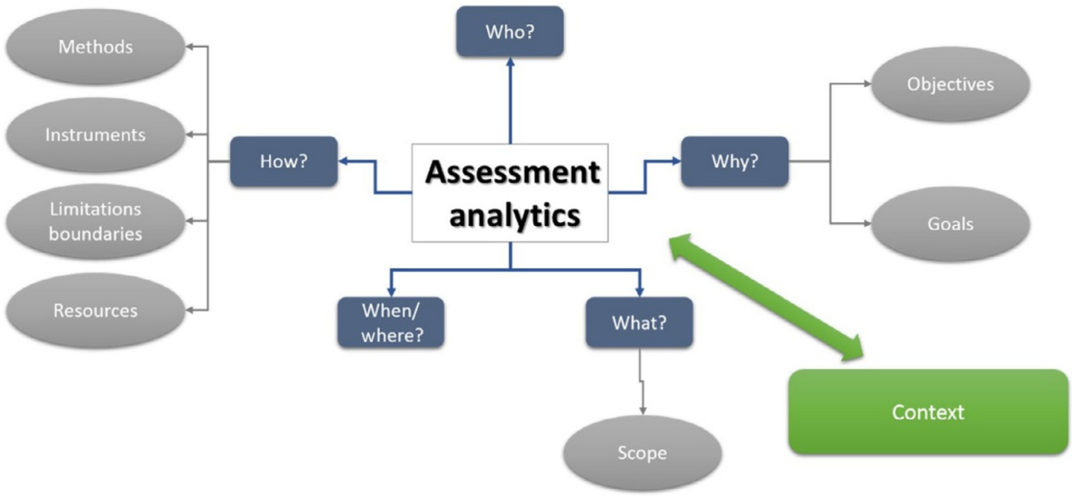

FIGURE 7.3

Upper-class concepts for assessment analytics.

assessment takes place (eg, environmental conditions, time, etc.). The AA *process* itself mostly concerns issues related to *how* it is applied and which parameters are being exploited during the procedure (eg, methods, resources, instruments, limitations and boundaries, pedagogy and instructional design, etc.). The *output* of the AA system is related to the process results and includes (a) *what* should be done next (eg, actions, pedagogical theories, algorithmic changes, educational policy, etc.), (b) *why* it should be done

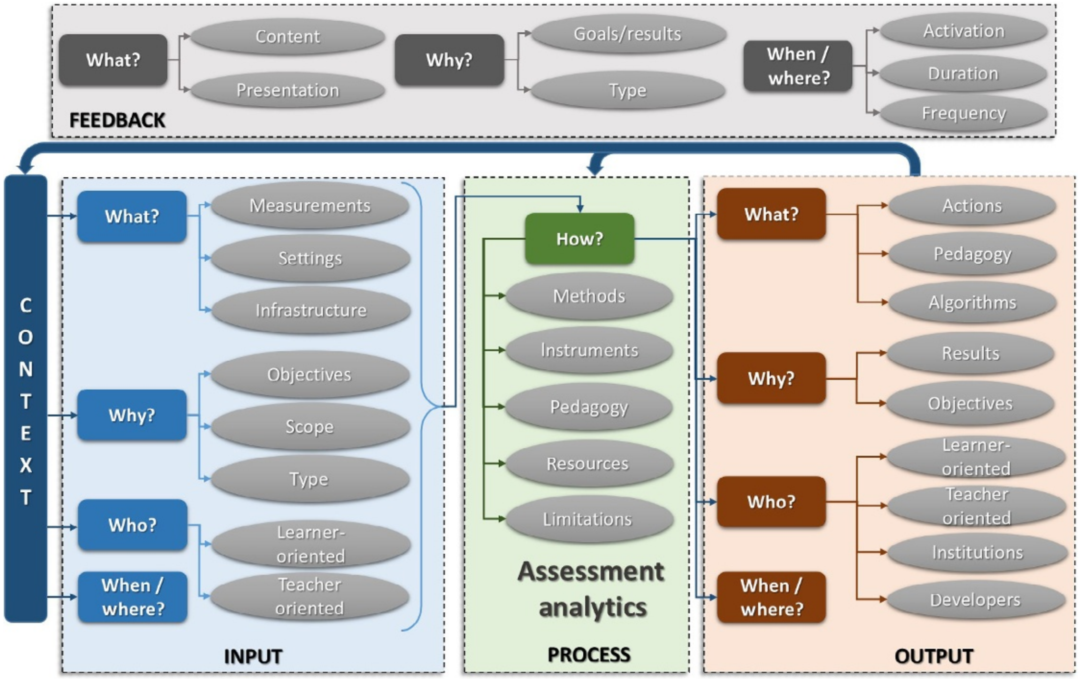

FIGURE 7.4

The proposed assessment analytics framework (AAF)—general schema.

(eg, based on the results and objectives achievement) and (c) *who* is the final receiver of the results (eg, parents, institutions, software developers—beyond students and teachers). Finally, the *feedback* is related to the delivery of the results to the recipient of the assessment result (eg, content, presentation, frequency, type, etc.) in order to change the initial state of the context (Economides, 2005).

The schema also presents the inherent connections and relationships between the major concepts. More specifically, the size and direction of the arrows represent the depth of the analysis from a more abstract to a more concrete concept.

We will discuss each of these blocks individually in the following section and exemplify their instantiations and impact on the AA process and the benefits and opportunities they may determine. We will also elaborate on apparent problem zones and limitations that may hinder any envisaged benefits.

4.3.1 The role of the context

Any information that can be used to characterize the situation of an entity (eg, person, place or object) would be considered as context (Dey and Abowd, 2000). In the field of learning sciences, the belief that context matters leads to the conclusion that research paradigms that simply explore the learning processes as isolated variables within laboratory settings will necessarily lead to an incomplete understanding of their relevance in real-life settings. Context estimation (eg, place of interest, learning topics, unattained knowledge, and tendency of viewpoint) of a real-world learner has been acknowledged as an important parameter during assessment.

In the domain of AA, all studies set out the context of the research experiments, including classroom settings, real-world settings, virtual learning environments (VLEs), learning management systems (LMSs), MOOCs, intelligent tutoring systems (ITS), etc.

The role of context in the AAF is to provide the necessary input information to the AA system (ie, the whole AA system, and not only the analytics engine): who are the learners, who are the teachers, what is the available infrastructure, what will be measured, when and where these measurements will take place, why they are needed and which are their objectives? This initial information will next constitute the "input" block, ie, the input data that will be under processing by the assessment analytics engine.

4.3.2 The "input" block: what, why, who, when, and where is being tracked and assessed?

Every assessment analytics system consists of a number of input parameters that describe it as an entity. It is important to identify from the beginning who will be assessed (eg, children, teenagers, adults, employees, etc.), by whom (eg, self-assessment, the instructor, a computer, etc.), what will be assessed (eg, knowledge, skills, behaviors, etc.), for what purpose (eg, summative, formative, predictive, diagnostic, etc.), and where and when the assessment will take place (eg, in classroom, at home, at an outdoor activity, etc.). It is also important to define from the beginning what will be tracked, which measures will act as the variables that will be analyzed and generate the assessment result. In AA research, these variables vary from study to study and include activity logs, chat and discussion data, temporal data, emotional data, free text, and many more. However, the selection of the most appropriate data type should be aligned with the overall assessment objective and research purpose (fit-for-purpose), and next analyzed with the most appropriate technique (not one-size-fit-all).

4.3.3 The "process" block: how the collected data are analyzed and interpreted?

The heart of every AA system is its processing engine, the mechanisms that are employed in analyzing the input data and producing the exploitable results. The literature review revealed the rich variety of different methods, algorithms and resources involved in the process of analysis. Examples of these components are linguistic analysis, text mining, speech recognition, classification techniques, association rule mining, process mining, machine learning techniques, affect recognition, and many more. Different instruments are adopted for the data processing, including algorithms and modeling methods. Table 7.1 summarizes the cases (application scenarios) according to the learning setting, the data-types and the analysis methods that have been employed.

In addition, the underlying pedagogical usefulness acts as a strong criterion that drives the whole analysis process in order to produce valid and useful results. However, most cases draw the attention to boundaries and limitations related mostly to ethics and security issues on the data manipulation.

4.3.4 The "output" block: what, why, who, when, and where is the outcome of the assessment process?

The result of the assessment procedure is also a central request during the designing of an AA system. According to the analysis of the literature, intelligent feedback, adaptation, personalization, and recommendations, are usually the outcome of the process, along with the diagnosis of misconceptions and knowledge gaps, participation rates, learning portfolios, achievements and summative results, like

Table 7.1 Summary of Application Scenarios With Learning Setting, Data-Types, Analysis Methods

Authors and Year	Learning Setting	Data-Type (Resources)	Data Analysis Method
Barla et al. (2010)	Web-based education—adaptive testing	Question repository—consider test questions as a part of educational material and improve adaptation by considering their difficulty and other attributes	Combination of item response theory with course structure annotations for recommendation purposes (classification)
Caballé et al. (2011)	Computer supported collaborative learning (CSCL) for assessing participation behavior, knowledge building and performance	Asynchronous discussion within virtual learning environment; post tagging, assent, and rating	Statistical analysis
Chen and Chen (2009)	Mobile learning and assessment (mobile devices used as the primary learning mediator)	Web-based learning portfolios—personalized e-learning system and PDA; attendance rate, accumulated score, concentration degree (reverse degree of distraction)	Data mining: statistic correlation analysis, fuzzy clustering analysis, gray relational analysis, K-means clustering, fuzzy association rule mining and fuzzy inference, classification; clustering
Davies et al. (2015)	Computer-based learning (spreadsheets) for detection of systematic student errors as knowledge gaps and misconceptions	Activity-trace logged data; processes students take to arrive at their final solution	Discovery with models; pattern recognition
Leeman-Munk et al. (2014)	Intelligent tutoring systems (ITS) for problem-solving and analyzing student-composed text	Digital science notebook; responses to constructed response questions; grades	Text analytics; text similarity technique; semantic analysis technique
Moridis and Economides (2009)	Online self-assessment tests	Multiple choice questions	Students' mood models—pattern recognition
Okada and Tada (2014)	Real-world learning for context-aware support	Wearable devices; sound devices; audio-visual records; field notes and activity maps; body postures; spatiotemporal behavior information	3D classification; probabilistic modeling; ubiquitous computing techniques
Papamitsiou et al. (2014)	Web-based education—computer-based testing	Question repository—temporal traces	Partial Least Squares—statistical methods
Perera et al. (2009)	CSCL to improve the teaching of the group work skills and facilitation of effective team work by small groups	Open source, professional software development tracking system that supports collaboration; traces of users' actions (create a wiki page or modify it; create a new ticket, or modify an existing, etc.)	Clustering; sequential patterns mining

Continued

Table 7.1 Summary of Application Scenarios With Learning Setting, Data-Types, Analysis Methods—cont'd

Authors and Year	Learning Setting	Data-Type (Resources)	Data Analysis Method
Sao Pedro et al. (2013)	ITS—for inquiry skill assessment	Fine-grained actions were timestamped; interactions with the inquiry, changing simulation variable values and running/pausing/resetting the simulation, and transitioning between inquiry tasks	Traditional data mining, iterative search, and domain expertise
Segedy et al. (2013)	Open-ended learning environments (OELEs) for problem-solving tasks and assess students' cognitive skill levels	Cognitive tutors; causal map edits; quizzes, question evaluations, and explanations; link annotations	Model-driven assessments; model of relevant cognitive and metacognitive processes; classification
Shibani et al. (2015)	Teamwork assessment for measuring teamwork competency	Custom-made chat program	Text mining; classification
Tempelaar et al. (2013)	Learning management systems (LMS) for formative assessment	Generic digital learning environment; LMS; demographics, cultural differences; learning styles; behaviors; emotions	Statistical analysis
Vozniuk et al. (2014)	MOOCS—peer assessment	Social media platform—consensus, agreement, correlations	Statistical analysis
Xing et al. (2014)	CSCL—virtual learning environments (VLE) for participatory learning assessment	Log files about actions, time, duration, space, tasks, objects, chat	Clustering enhanced with activity theory

grades and performance scores. The output of the AA system will become available to the learner, the instructor, the institution, or even the system developer. In all cases, the goal is for this output to be interpretable and comprehensive in order to increase the awareness of its receiver regarding the change that has occurred and been measured by the assessment itself. Of course, this output has to be aligned with (a) the initial objectives of the assessment procedure (eg, automate teamwork assessment, measure students' satisfaction, support in real-time students to successfully complete an assignment, etc.), and (b) the pedagogical goals set by the assessor. It also should provide results that will be accurate, accessible and useful to the end user of the system.

4.3.5 The "feedback" block: what, why, when, and where is delivered to close the loop effectively?

As mentioned before—in the "output" block—the outcome of the assessment process is usually some type of feedback. Accordingly, in an AA system, feedback acts in a twofold way: (a) it constitutes the feedback of the assessment process, ie, it informs the final user (eg, student, teacher, learning administrator, etc.) about the result of the assessment process, and (b) at the same time, it acts like the feedback loop of any iterative system, by keeping the loop continuing and feeding the shift of the former situation of the person under assessment back to the context. Economides (2006) provides a comprehensive description of feedback characteristics, which include the feedback activation reasons, the purposes of feedback and its expected effect, the type (eg, affective, cognitive, positive, negative, etc.), presentation issues of feedback, as well as issues related to the frequency, timing and duration of feedback.

4.4 DEDUCTION PHASE: VALIDATION OF THE AAF

In order to test and validate the proposed AAF, we followed a deduction approach: we aimed at retesting existing data (ie, already published studies) in the new context (ie, the AAF). Thus, we randomly chose two studies from the domain of AA research—not previously used during the induction phase of our methodology—and explored whether they fit in the suggested schema. Table 7.2 illustrates the results of this process, and provides a brief proof-of-concept regarding our proposed theory.

Table 7.2 Validation of the Assessment Analytics Framework—Analysis of Two Studies

Study	Context	Input	Process	Output	Feedback
Leeman-Munk et al. (2014)	Open-ended learning environments—STEM education (science education)	**What:** knowledge—short-text answers to science questions	**How:** text similarity combined with semantic analysis—classification of submitted answers	**Who:** teacher-oriented	**What:** correctness of submitted answer
		Why: real-time formative assessment		**Why:** early warning indicators to teachers to strategize as to how to allocate instructional interventions	**Why:** "train" the system to predict the student's performance
		When/Where: classroom		**What:** grades	**Where/ When:** after every submitted answer
Whitelock et al. (2015)	Academic course assignment – online assessment	**What:** students' essays during writing	**How:** linguistic analysis	**What:** graphical representation related to key linguistic characteristics of the document under development +grades	**What:** multiple types of feedback

Continued

Table 7.2 Validation of the Assessment Analytics Framework—Analysis of Two Studies—cont'd

Study	Context	Input	Process	Output	Feedback
		Why: summative assessment of free-text answers—provide meaningful advice for action		**Why:** self-reflection and support	**Why:** increase students' self-awareness
		Who: the students during writing the essay		**Who:** student-oriented	**Where/When:** during the development of the document—on demand
		Where/When: real-time online during the assignment		**Where/When:** online—during the assignment	

5 DISCUSSION AND CONCLUSIONS

As stated in the motivation section, a deeper observation of the e-assessment research findings through the microscope of AA partially reveals the intrinsic capabilities of these processes to support self-reflection and self-regulation during assessment procedures.

In this chapter we suggested a framework for analyzing and better understanding current research on AA. In order to design and evaluate the framework we followed a deductive and inductive inquiry methodology for concept mapping for sense making. The procedure ended up with a conceptual map consisting of four major clusters, each of which is further analyzed in more specific dimensions. For the validation of the framework we used a number of studies on AA in order to provide a proof-of-concept regarding the theory. In that way, the suggested framework explicitly validates AA results and can be used to move beyond descriptions of "what" to explanations of "why" and "how." By employing a general schema, the proposed framework for AA covers all dimensions of the assessment process and considers them as part of any teaching and learning strategy. The target is to explain the potential benefits of AA to better understand the reasons for students' progress or failure.

However, what is still missing from the existing literature and, consequently, not yet included in details in the proposed framework, are issues related to the security and privacy of tracked and analyzed information during the assessment procedure. For example, students should be able to access their personal data, as well as authorize when and with whom their data are shared. Moreover, students should be able to refuse to make their data available for sharing and to determine who else could have access or exercise control over how their personal data are shared. Further research is required in that direction.

To conclude, we believe that the suggested approach, which is the first of its kind (to the best of our knowledge), has covered central issues for AA. Improvements and extensions are necessary in order to add value and strength to the theory so that it may gain acceptance.

REFERENCES

Barla, M., Bieliková, M., Ezzeddinne, A.B., Kramár, T., Šimko, M., Vozár, O., 2010. On the impact of adaptive test question selection for learning efficiency. Comput. Educ. 55 (2), 846–857.

Birenbaum, M., 1996. Assessment 2000: towards a pluralistic approach to assessment. In: Birenbaum, M., Dochy, F.J.R.C. (Eds.), Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge. Springer, Netherlands, pp. 3–29.

Caballé, S., Daradoumis, T., Xhafa, F., Juan, A., 2011. Providing effective feedback, monitoring and evaluation to on-line collaborative learning discussions. Comput. Hum. Behav. 27 (4), 1372–1381.

Carnoy, M., Jarillo Rabling, B., Castano-Munoz, J., Duart Montoliu, J.M., Sancho-Vinuesa, T., 2012. Who attends and completes virtual universities: the case of the Open University of Catalonia (UOC). High. Educ. 63, 53–82.

Chatti, M.A., Dyckhoff, A.L., Schroeder, U., Thüs, H., 2012. A reference model for learning analytics. Int. J. Technol. Enhanc. Learn. 4 (5/6), 318–331.

Chatzopoulou, D.I., Economides, A.A., 2010. Adaptive assessment of student's knowledge in programming courses. J. Comput. Assist. Learn. 26 (4), 258–269.

Chen, C.-M., Chen, M.-C., 2009. Mobile formative assessment tool based on data mining techniques for supporting web-based learning. Comput. Educ. 52 (1), 256–273.

Conole, G., Warburton, B., 2005. A review of computer assisted assessment. Altern. J. 13 (1), 17–31.

D'Mello, S.K., Graesser, A.C., 2009. Automatic detection of learners' emotions from gross body language. Appl. Artif. Intell. 23 (2), 123–150.

Davies, R., Nyland, R., Chapman, J., Allen, G., 2015. Using transaction-level data to diagnose knowledge gaps and misconceptions. In: Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15). ACM, New York, pp. 113–117.

Dermo, J., 2009. E-Assessment and the student learning experience: a survey of student perceptions of e-assessment. Br. J. Educ. Technol. 40 (2), 203–214.

Dey, I., 1993. Qualitative Data Analysis. A User-Friendly Guide for Social Scientists. Routledge, London.

Dey, A.K., Abowd, G., 2000. Towards a better understanding of context and context-awareness. In: Proceedings of Conference on Human Factors in Computing Systems. The Hague, The Netherlands.

Economides, A.A., 2005. Personalized feedback in CAT. WSEAS Trans. Adv. Eng. Educ. 3 (2), 174–181.

Economides, A.A., 2006. Adaptive feedback characteristics in CAT. Int. J. Instruct. Technol. Dist. Learn. 3(8).

Economides, A.A., 2009a. Conative feedback in computer-based assessment. Comput. Sch. 26 (3), 207–223.

Economides, A.A., 2009b. Adaptive context-aware pervasive and ubiquitous learning. Int. J. Technol. Enhanc. Learn. 1 (3), 169–192.

Erwin, T.D., 1991. Assessing Student Learning and Development: A Guide to the Principles, Goals, and Methods of Determining College Outcomes. Jossey-Bass Inc., San Francisco, CA.

Fernández-Gallego, B., Lama, M., Vidal, J.C., Mucientes, M., 2013. Learning Analytics framework for educational virtual worlds. Procedia Comput. Sci. 25, 443–447.

Greller, W., Drachsler, H., 2012. Translating learning into numbers: a generic framework for learning analytics. Educ. Technol. Soc. 15 (3), 42–57.

Haertel, E., et al., 2012. NAEP: looking ahead—leading assessment into the future. In: A Whitepaper Commissioned by the National Center for Education Statistics. US Department of Education, Washington, DC.

Jabareen, Y., 2009. Building a conceptual framework: philosophy, definitions, and procedure. Int. J. Qual. Methods 8 (4), 49–62.

Knight, S., Littleton, K., 2015. Developing a multiple-document-processing performance assessment for epistemic literacy. In: Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15). ACM, New York, pp. 241–245.

Knight, S., Shum, S.B., Littleton, K., 2013. Epistemology, pedagogy, assessment and learning analytics. In: Suthers, D., Verbert, K., Duval, E., Ochoa, X. (Eds.), Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13). ACM, New York, pp. 75–84.

Kobayashi, A., Muramatsu, S., Kamisaka, D., Watanabe, T., Minamikawa, A., Iwamoto, T., Yokoyama, H., 2011. User movement estimation considering reliability, power saving, and latency using mobile pone. IEICE Trans. Inf. Syst. E94-D (6), 1153–1163.

Leeman-Munk, S.P., Wiebe, E.N., Lester, J.C., 2014. Assessing elementary students' science competency with text analytics. In: Proceedings of the 4th International Conference on Learning Analytics And Knowledge (LAK '14). ACM, New York, pp. 143–147.

Leony, D., Muñoz Merino, P.J., Pardo, A., Kloos, C.D., 2013. Provision of awareness of learners' emotions through visualizations in a computer interaction-based environment. Expert Syst. Appl. 40 (13), 5093–5100.

MacNeill, S., Ellis, C., 2013. Acting on assessment analytics. JISC CETIS Anal. Ser. 2(2).

Marshall, C., Rossman, G.B., 1995. Designing Qualitative Research. Sage Publications, London.

McCann, A.L., 2010. Factors affecting the adoption of an e-assessment system. Assess. Eval. High. Educ. 35 (7), 799–818.

Miles, M.B., Huberman, A.M., 1994. Qualitative Data Analysis: An Expanded Source Book, second ed. Sage, Newbury Park, CA.

Monroy, C., Rangel, V.S., Whitaker, R., 2013. STEMscopes: contextualizing learning analytics in a K-12 science curriculum. In: Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13). ACM, New York, pp. 210–219.

Moridis, C.N., Economides, A.A., 2009. Mood recognition during online self-assessment test. IEEE Trans. Learn. Technol. 2 (1), 50–61.

Nardi, B., 1995. Context and Consciousness: Activity Theory and Human-Computer Interaction. MIT Press, Cambridge, ISBN 0-262-14058-6.

Novak, J.D., 1981. The Use of Concept Mapping and Gowin's "V" Mapping Instructional Strategies in Junior High School Science. Report of the Cornell University 'Learning How to Learn' project, Ithaca, New York.

Okada, M., Tada, M., 2014. Formative assessment method of real-world learning by integrating heterogeneous elements of behavior, knowledge, and the environment. In: Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14). ACM, New York, pp. 1–10.

Papamitsiou, Z., Economides, A.A., 2014. Learning analytics and educational data mining in practice: a systematic literature review of empirical evidence. Educ. Technol. Soc. 17 (4), 49–64.

Papamitsiou, Z., Economides, A.A., 2015. Temporal learning analytics visualizations for increasing awareness during assessment. Int. J. Educ. Technol. Higher Educ. 12 (3), 129–147.

Papamitsiou, Z., Terzis, V., Economides, A.A., 2014. Temporal learning analytics for computer based testing. In: Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14). ACM, New York, pp. 31–35.

Peat, M., Franklin, S., 2002. Supporting student learning: the use of computer based formative assessment modules. Br. J. Educ. Technol. 33 (5), 515–523.

Pellegrino, J.W., 2013. Measuring what matters in a digital age: technology and the design of assessments for multisource comprehension. In: Sampson, D.G., et al., (Ed.), Ubiquitous and Mobile Learning in the Digital Age. Springer Science & Business Media, New York, pp. 259–286.

Perera, D., Kay, J., Koprinska, I., Yacef, K., Zaiane, O.R., 2009. Clustering and sequential pattern mining of online collaborative learning data. IEEE Trans. Knowl. Data Eng. 21 (6), 759–772.

Sancho-Vinuesa, T., Escudero-Viladoms, N., 2012. A proposal for formative assessment with automatic feedback on an online mathematics subject. Revista de Universidad y Sociedad del Conocimiento (RUSC) 9 (2), 240–260.

Sao Pedro, M.A., Baker, R.S.J.D., Gobert, J.D., 2013. What different kinds of stratification can reveal about the generalizability of data-mined skill assessment models. In: Suthers, D., Verbert, K., Duval, E., Ochoa, X.

(Eds), Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13). ACM, New York, pp. 190–194.

Segedy, J.R., Loretz, K.M., Biswas, G., 2013. Model-driven assessment of learners in open-ended learning environments. In: Suthers, D., Verbert, K., Duval, E., Ochoa, X. (Eds.), Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13). ACM, New York, pp. 200–204.

Shibani, A., Koh, E., Hong, H., 2015. Text mining approach to automate teamwork assessment in group chats. In: Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15). ACM, New York, pp. 434–435.

Shum, S.B., Crick, R.D., 2012. Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. In: Buckingham Shum, S., Gasevic, D., Ferguson, R. (Eds.), Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12). ACM, New York, pp. 92–101.

Swan, K., Shen, J., Hiltz, S.R., 2006. Assessment and collaboration in online learning. J. Async. Learn. Network. 10, 45–62.

Tempelaar, D.T., Heck, A., Cuypers, H., van der Kooij,, H., van de Vrie, E., 2013. Formative assessment and learning analytics. In: Verbert, Katrien, Duval, Erik, Ochoa, Xavier (Eds.), Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13), Dan Suthers. ACM, New York, pp. 205–209.

Triantafillou, E., Georgiadou, E., Economides, A.A., 2008. The design and evaluation of a computerized adaptive test on mobile devices. Comput. Educ. 50 (4), 1319–1330.

Vozniuk, A., Holzer, A., Gillet, D., 2014. Peer assessment based on ratings in a social media course. In: Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14). ACM, New York, pp. 133–137.

Whitelock, D., 2011. Activating assessment for learning: are we on the way with Web 2.0? In: Lee, M.J.W., McLoughlin, C. (Eds.), Web 2.0-Based-E-Learning: Applying Social Informatics for Tertiary Teaching. IGI Global, Hershey, Pennsylvania, pp. 319–342.

Whitelock, D.M., Brasher, A., 2006. Developing a roadmap for e-assessment: which way now? In: Danson, M. (Ed.), Proceedings of the 10th CAA International Computer Assisted Assessment Conference. Professional Development, Loughborough University, Loughborough, UK, pp. 487–501.

Whitelock, D., Twiner, A., Richardson, J.T.E., Field, D., Pulman, S., 2015. OpenEssayist: a supply and demand learning analytics tool for drafting academic essays. In: Proceedings of the 5th International Conference on Learning Analytics and Knowledge (LAK '15). ACM, New York, pp. 208–212.

Wilson, K., Boyd, C., Chen, L., Jamal, S., 2011. Improving student performance in a first-year geography course: examining the importance of computer-assisted formative assessment. Comput. Educ. 57 (2), 1493–1500.

Worsley, M., Blikstein, P., 2013. Towards the development of multimodal action based assessment. In: Suthers, D., Verbert, K., Duval, E., Ochoa, X. (Eds.), Proceedings of the 3rd International Conference on Learning Analytics and Knowledge (LAK '13). ACM, New York, pp. 94–101.

Xing, W., Wadholm, B., Goggins, S., 2014. Learning analytics in CSCL with a focus on assessment: an exploratory study of activity theory-informed cluster analysis. In: Proceedings of the 4th International Conference on Learning Analytics and Knowledge (LAK '14). ACM, New York, pp. 59–67.