Chapter 12
The Bi-directional Effect Between Data and Assessments in the Digital Age

Abelardo Pardo and Peter Reimann

Abstract Assessment and feedback, two important factors of any learning experience in higher education, are being significantly disrupted by the emergent ecosystem of increasingly technology-mediated learning environments. Digital learning experiences produce data sets with highly detailed accounts of the interactions among participants. This information brings an unprecedented potential to move the focus of assessment and feedback away from the result to the process by which such result is attained. This new focus may have a profound effect in the process to influence how students engage with their work, its comparison with an appropriate standard, and to increase their self-evaluative capacity. But at the same time these data sets pose substantial challenges on how to integrate their presence in the design, deployment and refinement of learning experiences. In this chapter, we describe the main elements that need to be considered to translate these rich data sets into actions and design aspects that achieve a positive effect in the student experience.

12.1 Introduction

The increasing mediation of technology in learning experiences has translated into widespread availability of detailed data traces. Almost every computer program supporting student learning has the potential of recording detailed logs of user interactions. These traces can be used to offer a more detailed understanding of the learning process and identify potential improvements (Macfadyen and Dawson 2010).

Despite the increasingly pervasive presence of data in learning, the connection between data and assessment remains largely under/unexplored. Intelligent tutoring systems provide fully automated support for assessment and teaching, but they

A. Pardo (✉)
UniSA STEM, University of South Australia, Adelaide, Australia
e-mail: Abelardo.Pardo@unisa.edu.au

P. Reimann
The University of Sydney, Centre for Resarch on Learning and Innovation, Sydney, Australia
e-mail: peter.reimann@sydney.edu.au

© Springer Nature Switzerland AG 2020

M. Bearman et al. (eds.), *Re-imagining University Assessment in a Digital World*, The Enabling Power of Assessment 7,
https://doi.org/10.1007/978-3-030-41956-1_12

typically require sophisticated models of the student, domain knowledge, and reasoning patterns (Reimann et al. 2013). However, as pointed out by Baker (2016), substantial research is needed to explore how data can enhance teacher intelligence, instead of trying to replace it.

Mediation through technology now produces data that can be used to observe how students collaboratively write a document, evolve and improve a design concept, or capture the result of a brainstorming session. Assessment should include both the collection of information about the learner, but also acting on this information in relation to the educational goals (Shute and Rahimi 2017). The increasing presence of technology in learning environments opens new avenues to provide measurements of the overall student experience (Shute et al. 2016). Thus, the design and deployment of learning environments in general and assessment in particular need to be revisited to explore how to better exploit the use of data to enhance the overall intelligence of the process.

This chapter explores two dimensions to revisit assessment in the presence of comprehensive data capturing. The first one assumes that technology may widen both the repertoire and scaling possibilities for assessment. The second dimension is based on the possibility of observing the process through which the evidence for assessment is created.

12.2 Background

Assessment is generally accepted as a ubiquitous part of the student experience, one with a significant influence on how students approach their learning experience. This importance has led to a significant body of knowledge about various aspects of assessment such as their theoretical underpinning, the production and collection of evidence, the methods to for interpreting performance, or their effect in student learning (Pellegrino 2018). Additionally, assessment poses various challenges related with aspects such as robustness, scalability, alignment with learning objectives, or lack of student understanding. Rust et al. (2005) proposed an assessment model based on a social constructivist approach with special emphasis on the alignment between assessment and the other components of a learning design—the clear definition of criteria and the creation of an effective feedback process. We posit that each of these elements can be revisited in the context of a technology-mediated learning environment and the availability of detailed data about the interactions occurring in such context.

The increasing presence of technology mediating learning experiences has transformed the context in which assessment takes place. Shute et al. (2016) point to the advancements in technology and learning sciences as two conditions that are prompting the community of experts to reconceptualize assessment as a whole. Advances in learning sciences now require considering the affective and emotional context in which assessment and feedback processes occur. Analogously, advances in technology have redefined how communication takes place among those

participating in a learning experience. Technological platforms create detail traces of the events occurring in these contexts that can be processed by software programs to detect frequently occurring patterns. These patterns can be assimilated to specific student behaviour and automatically adapt the learning environment or provide suggestions to learners. Far from reducing the already existing tension between efficiency and accuracy of assessment, these changes have instead increased the need to explore the trade-off between these two aspects. In short, there is mounting pressure to increase the accuracy and efficiency of how student success is measured; the use of data collected using technology mediation is one possible avenue to address it.

In a review of computer-based assessment in the context of elementary and secondary education Shute and Rahimi (2017) categorized the presence of technology into those used as a supplement in the classroom, those that are web-based, and those that are data-driven and continuous. This last category emerges as a consequence of the use of data-driven approaches to assessment. There have been comprehensive studies attesting the need to explore how assessments are used in so-called Technology-Rich Environments (TRE; Bennett et al. 2007) to measure aspects such as problem-solving skills. Technology is also blurring the differences between *assessment of learning* and *assessment for learning* (Bennett 2011). In the context of this chapter we will assume that a TRE is a learning environment that is highly mediated by technology **and** such mediation provides comprehensive records or data traces. These data traces can be used for attaining a measure of a learning facet and to act on that information, for example by providing students with additional information for attaining the learning goals.

In parallel with the appearance of TREs, the areas of Educational Data Mining and Learning Analytics have emerged, providing techniques to increase the understanding of learning processes and providing the knowledge to improve them (Berland et al. 2014). Educational Data Mining (EDM) proposes the use of algorithms to process the vast amount of data that is captured when learning occurs in TREs and use that information to guide quantitative research and practice. The distinctive aspect for assessment is the capturing of detailed data traces to gain further understanding of artefacts created by learners, emotional states during learning experiences, etc. For example, Heffernan and Heffernan (2014) proposed to design and deploy questions, answers, web-based videos and hints, to support students while learning mathematics. The data captured during the interactions allows for the study of how students react in the presence of difficulty, how often they ask for help, and their trajectory while attaining the learning goals. Another example of the possibilities of data capturing in TREs looks at the detection of affective states (boredom, confusion, frustration, etc.) using Natural Language processing tools while students interact with a math tutor (Slater et al. 2017).

Complementary to EDM, the area of Learning Analytics proposes the use of data about learning experience to increase the understanding of how learning occurs and to support their improvement. The focus is not only on the algorithmic part of the problem, but on how the use of data can be properly integrated in the ecosystem of design, deployment, student support, and ethical and privacy issues (see Lang et al. 2017 for a comprehensive description of the area). The use of data in this context

has proven to be useful to understand emerging social structures (e.g., Bakharia and Dawson 2011; Ferguson and Buckingham Shum 2012), detect and support students at risk (e.g., Krumm et al. 2014; Macfadyen and Dawson 2010; Waddington et al. 2016), or ethical and privacy issues (Drachsler and Greller 2016; Kitto et al. 2018; Prinsloo and Slade 2015). For the sake of argument in this document, the areas of Educational Data Mining and Learning Analytics will be referred as the single *educational analytics*.

The overlap between assessment and educational analytics is obvious: on the one hand, there is the need for assessments to be reconceptualized in learning contexts with substantial technology mediation. On the other hand, there is also a need to widen assessment beyond the attainment of learning goals (summative assessment, or assessment of learning) towards assessing the processes that lead to learning. For example, taxonomies proposed for the design of assessment items in e-learning environments (Scalise and Gifford 2006) still resort to the description of conventional items in the new context without accounting for the new knowledge that can be derived from captured data. More refined proposals such as the four levels of integration between technology and assessments (DiCerbo and Behrens 2012) consider the use of technology to accumulate information from a variety of digital activities and acknowledges dashboards as an example of that type of assessment. But the design of assessments that go beyond the accumulation of information and combine aspects such as usage patterns, engagement patterns, predictive models, etc. are not yet a reality.

Similarly, educational analytics methods are mainly focussed on how to capitalize on the vast amount of information collected, how to distil knowledge from this data, but the connection with robust assessment paradigms is still weak. Collecting detailed information about students allows for the personalisation of processes such as feedback (Pardo et al. 2018), but this area needs to deepen the study of how knowledge derived from data traces can be properly integrated in a learning design and inform assessment tasks.

We envision two areas through which this connection between assessment and educational analytics can be articulated. The first one is around the notion of new assessment techniques deployed at scale. Comprehensive data collection may offer instructors and learners the possibility of exploring assessment aspects that go beyond conventional boundaries such as courses, topics, or even degrees. Learning is becoming a life-long endeavour and comprehensive data collection may offer insights never observed before. Analogously, some assessment typically restricted to contexts with a small number of participants can now be scaled through technological support. Algorithms may support large communities of learners to assess their contributions, suggest adequate next steps, or recommend peers for interaction, typically restricted to small student cohorts.

The second area to connect assessment and educational analytics is to widen the focus of assessment from final artefacts to processes. Assessments typically focus on specific artefacts that are the result of learners engaging with a task. Technology and comprehensive data capture may pave the way to increase our understanding of these processes and perhaps evolve towards a *correct by construction* type of

assessment where learners are guided throughout the artefact-creation process to accomplish the learning goals.

12.3 Widening Assessment Repertoire and Scaling

ICT and educational analytics can extend the assessment repertoire in a number of ways: by (i) offering established assessments more efficiently and more frequently to large numbers of students; by (ii) extending the kinds of assessments; and by (iii) extending what gets assessed. The use of learning management systems (LMSs) has contributed greatly to extending the reach and frequency of assessments in Higher Education, thus making it easier to employ assessment as a tool for learning, not only of learning (Pellegrino 2018). Multiple-choice tests and many forms of assignments—for individual and group work—can be deployed with ease, across hundreds and thousands of students. Relevant technologies and processes, including MOOCs, are by now in place in higher education around the world. A main benefit of educational analytics lies in the potential to analyse relations between assessments across courses and over time (semesters, years), thus exploiting the potential of big data on learning. For example, Poquet et al. (2018) used data about how users participate in a discussion forum in a MOOC over ten iterations of the course. The results identified three patterns of engagement that appear consistently throughout the editions, and they are maintained even when the number of participants decreases significantly. The type of interactions in the forum also evolved towards a higher focus on tasks related to the course and patterns similar to conventional Q&A spaces. A further example of an analysis of relations across assessments is found in Chap. 11 (Rogaten et al. this volume).

Of particular relevance for the second type of widening is the use of ICT for integrating assessment with learning activities. The guiding vision is the TREs in which all relevant learning (inter-)actions are captured and can be interpreted as such because the environment is designed accordingly (Shute et al. 2016). A typical example is the patient simulator for medical education, described in (Blanchard et al. 2012). TREs are much less prevalent in Higher Education than learning management systems. Because of the time and effort required to develop such environments, which are always domain-specific, they get developed for learning domains that are fairly stable and highly relevant. The main role of educational analytics is similar to the first case: integration of data from different environments and courses and comparative analysis across specific applications. The latter task is made easier by storing data in repositories such as the CMU data shop (pslcdatashop.web.cmu.edu). This repository provides access to data sets stemming from the use of intelligent tutoring systems and facilitates the analysis of these data from a number of perspectives (Koedinger et al. 2010).

The third form of widening pertains to assessing properties of learning and the learners that have not been the subject of assessment in higher education, from self-regulation to graduate qualities and “21st century skills”. The two main strategies to

get to measurements are the use of psychometric tools such as scales for self-regulation and the use of analytics to identify and track indicators in students' activities. For instance, Fincham et al. (2018) proposed the use of data capture about how students engage in a learning experience to first obtain indicators of *study sessions* and then use them to identify different study tactics and how they evolve over the duration of a course adopting a flipped classroom design strategy. Although there is not a clear notion of a *correct* or *incorrect* study tactic, the information extracted from the traces provides a nuanced account of how students approach a learning experience and the potential to support them more effectively throughout the process. The key elements to use these traces effectively are (1) the connection with relevant elements of the learning design, and (2) the derivation of actions that promote aspects already present in learning designs such as self-reflection, regulation, goal setting, etc. The first element, the connection with elements of the learning design requires an explicit relation between data traces and the type of interactions that are desired within a learning experience. For example, if learners are given a set of exemplars to analyse and then discuss, the data capturing process should be aligned with these two steps and identify which exemplars are being accessed and which events occur in the discussion space. The wide variety of possible scenarios makes this relation highly sensitive to the context, but when achieved it guarantees that the indicators identified by software programs can be interpreted within the context of a task in the design. The second element, the derivation of actions to support learners requires a tight integration between the conclusions derived from interpreting the traces and the type of actions. For example, elements within the learning context can be labelled as to support certain aspects and therefore be part of a set of *recommendations* automatically provided to the students when certain patterns are identified. A hybrid approach in terms of automation could require the intervention of the instructor (expert) to provide coaching advice on how to adopt a more appropriate learning strategy. The key element in this context is to align the comprehensive collection of data with existing models of what constitutes good learning and assessment practices.

In these scenarios educational analytics plays a role not only in integrating data sets into big data and performing comparative analyses but promoting new approaches for assessment design and engineering. This is because the student activities relevant for assessment take multiple forms and are distributed over multiple contexts. For instance, when assessing the development of collaboration skills, evidence for skill proficiency can consist of self-reports, tests, psychometric scales, and a range of observations captured in log files and in portfolios in a variety of databases. This task is made the more challenging as university teachers (other than those teaching in a liberal arts college perhaps) are by and large not experts in the development and assessment of student qualities that are not subject-matter specific. This and the fact that educational analytics practitioners are also not experts in—for a lack of better word—general pedagogy has contributed to an approach to assessment design that does not build on knowledge about the nature and the development of such capacities.

In this knowledge-lean approach, the intrinsically complex measurements of the quality of collaboration can be approximated through indicators that are mainly identified by way of data mining. For instance, assume we know students' scores on a psychometric scale for self-regulation and have all their data from the learning management system they use. In such circumstances, new indicators for self-regulation can be identified by correlating the scores from the scale with variables derived from the LMS data—such as regularity of contributions, adherence to deadlines, etc. An example of this combination of results was described by Ellis et al. (2017). The conceptual framework of student approaches to learning (Pintrich 2004) was combined with the Revised Study Process Questionnaire (R-SPQ) to collect data self-reported by the students (Biggs et al. 2001). The results were processed using Exploratory Factor Analysis and combined with data derived from interactions with resources in a blended learning environment. The authors show the increase in variance explained when both data sources are combined.

The knowledge-lean approach, while being practical, has a number of disadvantages, owing to its inductive-correlational nature. When statistical significance is used as the criterion, then the selection of indicators depends on the number of students; the size of correlations depends on variability (hence, processes that are necessary but show less variability will be missed); and including less or more variables might change the correlations significantly. Another disadvantage is that knowing that two variables correlate does not necessarily lead to actionable knowledge. For instance, raising student's adherence to timelines may not lead to higher self-regulation; it may as likely have no effect or a negative effect on self-regulation. Knowledge-rich approaches to assessment engineering, such as provided with the Cognitive Design Systems framework (Embretson 1998) are better suited to develop valid indicators. However, they require more effort and the combination of domain expertise with pedagogical and assessment expertise. A major advantage we see in their deployment is the closure of the gap between assessment and learning processes (Shute et al. 2016, p. 52).

12.4 Observing the Process

The second assessment aspect that will be significantly impacted by the ubiquity of data sets is the capacity to observe the process of creating artefacts for assessment. In typical assessment scenarios, learners are given the description of a task, a set of assessment criteria and they have to produce an artefact that is going to be assessed with respect to those criteria. A more holistic view of this process includes the deployment of feedback and the interplay of existing knowledge and beliefs, goal setting, strategies to achieve these goals, etc. Existing models such as, for example, self-regulated learning (Winne 1997, 2014) consider all elements in a common context that include explicitly the notion of task and performance (see Winne 1997, p. 399). But the existence of comprehensive data collection and analysis methods requires a reconceptualization of the relationship between learners, assessment and

feedback (Pardo 2018). A new landscape is emerging in which assessment tasks can be conceived as a learning trajectory that may include frequent interactions with multiple agents (human and non-human) informed by data collection. Various elements in this trajectory such as task descriptions, clarifications, relevant resources, or even strategic steps, may be adjusted or personalised depending on the captured contextual data. The increased relevance of this trajectory may be because data provides unprecedented levels of detail to be deployed at scale (large number of learners) and consequently may lead to more comprehensive learner support.

A representative example of the combination of technology, human intervention, and a tight integration with existing learning design is provided by reflective writing assignments (Gibson et al. 2017). Natural language processing technology was used to analyse the rhetorical moves of 120 submissions from 30 students. The study highlighted the need to interpret textual data in the context of the assignment and a theoretical framework to capture the structure of such texts. But once these indicators are obtained, it was equally important to frame the feedback in adequate terms to effectively support learners towards improvement of their submissions.

In these contexts, assessment designers not only need to contemplate the production of criteria to evaluate the final artefact, but also need to develop rules to analyse the indicators derived from the creation process, relate these observations to actions and deploy them in the assessment scenario. This new way of assessment design is complex. The process of connecting data with actionable items is influenced by the learning design and heavily mediated by the learner's affective reactions, motivation, self-regulation, etc.

Additionally, the use of indicators derived from data traces requires us to reconsider the notion of validity of assessment instruments (Hickey et al. 2000; Maier et al. 2016). An assessment instrument is considered valid when it provides a sound distinction between different levels of attainment of the learning goals. Indicators derived from captured data are highly unlikely to provide such level of validity. Typically, the indicators will *correlate* with other aspects of the learning experience, thus providing only a partial or approximate measure of the process. In other words, indicators derived from observing how students engage with an assessment task will contain an inherent level of uncertainty (Macfadyen and Dawson 2012).

Assessment criteria need also to be revisited when considering the shift of focus towards the process to create the artefact. Under this new lens, scoring the final outcome is just one out of multiple elements of the assessment process. Criteria now need to morph into more sophisticated statements describing the aspects to observe during the process of creating an artefact, the way observations are combined to derive insights, the actions that are derived from these insights, and how such actions are deployed in the learning experience. Technology mediation opens the possibility of reconsidering the overall process as a continuously iterative reflective loop involving learners and instructors that simultaneously promotes the capacity of learners to assess their work, but also to achieve the desired outcomes. We envision the presence of technology as the catalyzer of assessments that blur the distinction between formative and summative.

This shift in focus is even more relevant when trying to assess the so-called *higher order skills*. Although arguably these skills have always been desirable and promoted in traditional learning experiences, the existence of data extracted from technology mediation widens the range of indicators to assess them. For example, concept or mind mapping techniques are typically used to promote critical and analytical skills (Davies 2011). There is a wide variety of software tools that help students create visual representation of concepts and relations. Observing students throughout the process of constructing these diagrams offers the opportunity to support them towards the creation of useful maps.

12.5 The Effect on Strategies for Assessment Design

The new conceptualisation of assessments in settings with comprehensive data capturing influences design and deployment. There are various aspects when designing an assessment that can be reconsidered in the presence of data. We envision a bi-directional influence between data and how assessments are designed. One of these design aspects is a potentially wider consideration of learning goals. Rather than stating the learning goals, an assessment may also identify the skills, strategies or attitudes that would facilitate the attainment of such goals. The data captured in relation to these skills may then be available to instructors and learners to increase their awareness of the whole process. This relation is an example of how assessment definition may influence how data is captured.

Analogously, if learning experiences are deployed in contexts that provide detailed observations of elements relevant to the attainment of the learning goals, the assessment may be reframed to take these observations into account. For example, if a platform allows student to specify their goals and strategic steps to tackle a problem, the assessment may be redesigned to include the provision of suggestions or support actions to take full advantage of this feature. Researchers are already obtaining increasingly reliable measures of learner affect (Bosch et al. 2015); assessments can now be designed to take these observations into account and personalise their form depending on these values. For example, a student known to have *test anxiety* may be identified in advance and given a combination of alternative tasks to obtain the required valid measure of attainment. Identifying this type of aspects undoubtedly has ethical ramifications that need to be taken into consideration when assessing the type of indicators as well as the conclusions derived from these observations. But the relationship between how assessments are conceived and analysed is bi-directional. Obtaining detailed accounts of how assessments are delivered may prompt reflection about features of the learning design. For example, identifying a large percentage of students with test anxiety may prompt a reconsideration of the assessment items in a design.

An additional aspect to be considered is how assessment results are reported. As discussed in the previous section, the indicators derived from the data captured in a learning environment may offer a partial view of how learners engage with a task,

but even in the presence of uncertainty, and being fully aware of their limitations, they could also be part of the assessment result that are distributed among stakeholders. This modified report containing indicators would also contribute to increase the transparency of the overall procedure as a mechanism to increase its reuse and adaptation to other contexts.

12.6 Conclusions

The availability of data captured by technology when mediating a learning experience is an element that is transforming numerous aspects of education. But as is already happening in other areas, the presence of data is not enough to produce tangible improvements. Data needs to be analysed, situated in the proper context, and connected with existing elements in a learning design. Assessment is no exception to this observation, quite the contrary. This chapter has explored two aspects of assessment that are heavily influenced by this change. The first one is that data availability is clearly widening the repertoire of assessment tasks that are feasible, but also the scale at which these can be deployed. Massive Open Online Courses are examples of how assessments can be shaped and the trade-offs that emerge under these demands. Systematic data collection allows for the creation and refinement of models to characterise how students participate in learning experiences. These models can be used to predict goal attainment, detect strategies, and ultimately provide more effective student support.

The second aspect that will be significantly influenced by the presence of data is the capacity to observe the process leading to the production of the artefact subject to assessment. This shift in focus away from the finished product and to the process leading to it opens the possibility to consider a holistic view including student prior knowledge, beliefs, goals, strategies, affect, etc. Although at the risk of significantly increasing their complexity, new assessment tasks may include criteria and instructions to extract insights from existing data and deploy actions to support students towards goal attainment.

Although these two aspects paint a promising horizon of improvements, there are numerous caveats that are intrinsic to these approaches need to be carefully considered. Although technology may provide the aforementioned detailed data traces, they may still not be enough to describe what exactly is happening in a learning experience. The best approach to account for this limitation is to frame this as an *approximation problem*. Technology is now offering us the possibility to have a more approximate idea of how learning experiences unfold, increase our level of understanding and therefore our possibility to improve them. This increase in perception cannot be derived from the mere presence of detailed data. Context is essential, thus the need for a solid connection between data, analytics methods and the learning design. Such connection is not trivial to establish and may significantly raise the overall complexity of the problem.

Aside from these caveats, the deployment of any paradigm that involves comprehensive collection of personal data, needs to be implemented under the strictest rules to observe privacy, ethics and transparency. These aspects not only apply to how data is captured and managed, but they need to be considered in the early stages of the design. The potential improvements in learning experiences cannot be an excuse to ignore these aspects.

In addition to these aspects, as with any context in which the use of data is introduced as a key element, it is desirable for all stakeholders to have solid data literacy skills. If learners receive detailed accounts of their progress, they need to be able to interpret the data and extract useful knowledge. These skills are even more important in the case of instructors as they are typically required to interpret data in the context of the overall application and assess both individual and population measures. The most powerful educational analytics solution would have a significantly reduced impact in the hands of stakeholders with poor data literacy skills.

As all these elements point out, the irruption of data in the assessment space has the potential of improving learning experiences, but at the same time requires a reconceptualization of the design procedures in which conventional assessment design techniques and the presence of data are continuously influencing each other.

References

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education, 26*(2), 600–614. https://doi.org/10.1007/s40593-016-0105-0.

Bakharia, A., & Dawson, S. (2011). *SNAPP: A bird's-eye view of temporal participant interaction.* Paper presented at the 1st international conference on learning analytics and knowledge, Banff, Canada.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25. https://doi.org/10.1080/0969594x.2010.513678.

Bennett, R. E., Persky, H., Weiss, A. R., & Jenkins, F. (2007). *Problem solving in technology-rich environments. A report from the NAEP technology-based assessment project, research and development series (NCES 2007-466)* (p. 196). Washington, DC: National Center for Education Statistics.

Berland, M., Baker, R., & Blikstein, P. (2014). Educational data mining and learning analytics: Applications to constructionist research. *Technology, Knowledge and Learning, 19*(1), 205–220. https://doi.org/10.1007/s10758-014-9223-7.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-SPQ-2F. *British Journal of Educational Psychology, 71*(1), 133–149. https://doi.org/10.1348/000709901158433.

Blanchard, E. G., Wiseman, J., Naismith, L., & Lajoie, S. P. (2012). *A realistic digital deteriorating patient to foster emergency decision-making skills in medical students*. Paper presented at the IEEE international conference on advanced learning technologies, Rome, Italy.

Bosch, N., D'Mello, S., Baker, R. S., Ocumpaugh, J., Shute, V. J., Ventura, M., … Zhao, W. (2015). *Automatic detection of learning centered affective states in the wild*. Paper presented at the 20th international conference on intelligent user interfaces, Atlanta, Georgia.

Davies, M. (2011). Concept mapping, mind mapping and argument mapping: What are the differences and do they matter? *Higher Education, 62*(3), 279–301.

DiCerbo, K. E., & Behrens, J. T. (2012). *From technology-enhanced assessment to assessment-enhanced technology*. Paper presented at the National Council on Measurement in Education 2012 Annual Meeting, Vancouver, Canada.

Drachsler, H., & Greller, W. (2016). *Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics*. Paper presented at the 6th international conference on learning analytics & knowledge, Edinburgh, UK.

Ellis, R. A., Han, F., & Pardo, A. (2017). Improving learning analytics – Combining observational and self-report data on student learning. *Educational Technology & Society, 20*(3), 158–169.

Embretson, S. E. (1998). A cognitive design system approach to generating valid tests: Application to abstract reasoning. *Psychological Methods, 3*(3), 380–396.

Ferguson, R., & Buckingham Shum, S. (2012). *Social learning analytics: Five approaches*. Paper presented at the 2nd international conference on learning analytics and knowledge, Vancouver, Canada.

Fincham, E., Gašević, D., Jovanović, J., & Pardo, A. (2018). From study tactics to learning strategies: An analytical method for extracting interpretable representations. *IEEE Transactions on Learning Technologies, 12*, 59–72. https://doi.org/10.1109/TLT.2018.2823317.

Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., & Knight, S. (2017). *Reflective writing analytics for actionable feedback*. Paper presented at the 2nd international conference on learning analytics and knowledge, Vancouver, Canada.

Heffernan, N. T., & Heffernan, C. L. (2014). The ASSISTments ecosystem: Building a platform that brings scientist and teachers together for minimal invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470–497.

Hickey, D. T., Wolfe, E. W., & Kindfield, A. C. H. (2000). Assessing learning in a technology-supported genetics environment: Evidential and systemic validity issues. *Educational Assessment, 6*(3), 155–196. https://doi.org/10.1207/s15326977ea0603_1.

Kitto, K., Buckingham Shum, S., & Gibson, A. (2018). *Embracing imperfection in learning analytics*. Paper presented at the 8th international conference on learning analytics and knowledge, Sydney, Australia.

Koedinger, K. R., Baker, R. S. J. D., Cunningham, K., Skogsholm, A., Leber, B., & Stamper, J. (2010). A data repository for the EDM community: The PSLC DataShop. In C. Robero, S. Ventura, M. Pechenizkiy, & R. Baker (Eds.), *Handbook of educational data mining* (pp. 43–56). Boca Raton: Chapman & Hall/CRC.

Krumm, A. E., Waddington, R. J., Teasley, S. D., & Lonn, S. (2014). A learning management system-based early warning system for academic advising in undergraduate engineering. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 103–119). New York: Springer.

Lang, C., Siemens, G., Wise, A., & Gašević, D. (Eds.). (2017). *Handbook of learning analytics*. Vancouver: Society for Learning Analytics Research.

Macfadyen, L. P., & Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & Education, 54*(2), 588–599. https://doi.org/10.1016/j.compedu.2009.09.008.

Macfadyen, L. P., & Dawson, S. (2012). Numbers are not enough. Why e-learning analytics failed to inform an institutional strategic plan. *Journal of Educational Technology & Society, 15*(3), 149–163.

Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers & Education, 95*, 85–98. https://doi.org/10.1016/j.compedu.2015.12.002.

Pardo, A. (2018). A feedback model for data-rich learning experiences. *Assessment & Evaluation in Higher Education, 43*(3), 428–438. https://doi.org/10.1080/02602938.2017.1356905.

Pardo, A., Jovanović, J., Dawson, S., Gašević, D., & Mirriahi, N. (2018). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*, 128–138. https://doi.org/10.1111/bjet.12592.

Pellegrino, J. W. (2018). Assessment of and for learning. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 410–421). New York: Routledge.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407. https://doi.org/10.1007/s10648-004-0006-x.

Poquet, O., Lim, L., Mirriahi, N., & Dawson, S. (2018). *Are MOOC forums changing?* Paper presented at the 8th international conference on learning analytics and knowledge, Sydney, Australia.

Prinsloo, P., & Slade, S. (2015). *Student privacy self-management: Implications for learning analytics*. Paper presented at the 5th international conference on learning analytics and knowledge, Poughkeepsie, USA.

Reimann, P., Kickmeier-Rust, M., & Albert, D. (2013). Problem solving learning environments and assessment: A knowledge space theory approach. *Computers & Education, 64*, 183–193. https://doi.org/10.1016/j.compedu.2012.11.024.

Rogaten, J., Clow, D., Edwards, C., Gaved, M., & Rienties, B. (This volume). Are assessment practices well aligned over time? A big data exploration. In M. Bearman, P. Dawson, J. Tai, R. Ajjawi, & D. Boud (Eds.), *Reimagining assessment in a digital world* (Chapter 11). Dordrecht: Springer.

Rust, C., O'Donovan, B., & Price, M. (2005). A social constructivist assessment process model: How the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education, 30*(3), 231–240. https://doi.org/10.1080/02602930500063819.

Scalise, K., & Gifford, B. (2006). Computer-based assessment in E-learning: A framework for construction "intermediate constraint" questions and tasks for technology platforms. *The Journal of Technology, Learning, and Assessment, 4*(6).

Shute, V. J., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning, 33*(1), 1–19. https://doi.org/10.1111/jcal.12172.

Shute, V. J., Leighton, J. P., Jang, E. E., & Chu, M. (2016). Advances in the science of assessment. *Educational Assessment, 21*(1), 34–59. https://doi.org/10.1080/10627197.2015.1127752.

Slater, S., Ocumpaugh, J., Baker, R., Almeda, M. V., Allen, L., & Heffernan, N. (2017). *Using natural language processing tools to develop complex models of student engagement*. Paper presented at the 7th international conference on affective computing and intelligent interaction, San Antonio, USA. https://doi.org/10.1109/ACII.2017.8273652.

Waddington, R. J., Nam, S., Lonn, S., & Teasley, S. D. (2016). Improving early warning systems with categorized course resource usage. *Journal of Learning Analytics, 3*(3), 263–290. https://doi.org/10.18608/jla.2016.33.13.

Winne, P. H. (1997). Experimenting to bootstrap self-regulated learning. *Journal of Educational Psychology, 89*(3), 397–410.

Winne, P. H. (2014). Issues in researching self-regulated learning as patterns of events. *Metacognition and Learning, 9*(2), 229–237. https://doi.org/10.1007/s11409-014-9113-3.

Abelardo Pardo is Professor and Dean of Programs at UniSA STEM at the University of South Australia. His research interests include the design and deployment of technology to increase the understanding and improve digital learning experiences. More specifically, his work examines the areas of learning analytics, personalized active learning, and technology for student support.

Peter Reimann is Professor in the Sydney School of Education and Social Work at the University of Sydney where he co-directs the Centre for Research on Learning and Innovation (CRLI). His primary research area is the study of learning with emphases on educational computing, multimedia-based and knowledge-based learning environments, e-learning, and the development of evaluation and assessment methods for the effectiveness of computer-based technologies. Peter is

well-known for his work on learning by self-explanation, on augmentation approaches to support computer-mediated group learning, and for his contributions to computer-based analysis of temporal data. His recent book *Measuring and visualising learning in the information-rich classroom* has been published by Routledge.