

Chapter 12

Intersecting Learning Analytics and Measurement Science in the Context of ICT Literacy Assessment

Mark Wilson, Kathleen Scalise, and Perman Gochyyev

Abstract This chapter reviews the state-of-play in overlap between learning analytics (LA), specifically data mining and exploratory analytics, and the field of measurement science. First, some basic ideas are introduced in a broad way. Then a current definition of LA is introduced, and main ideas of the area are discussed. Second, the logic of measurement science is reviewed, as instantiated through the BEAR Assessment System (BAS; Wilson, Constructing measures: an item response modeling approach. Lawrence Erlbaum Assoc, Mahwah, 2005), and illustrated in the context of an LA example. An example based in the context of ICT Literacy is presented, showing how complex digital assessments can be designed through BAS with attention to measurement science, while LA approaches can help to score some of the complex digital artifacts embedded in the design. With that background, ways are suggested through which the two approaches can be seen to support and complement one another, leading to a larger perspective. This chapter concludes with a discussion of the implications of this emerging intersection, and a survey of possible next steps.

M. Wilson (✉)

University of California, Berkeley, CA, USA

University of Melbourne, Parkville, VIC, Australia

e-mail: markw@berkeley.edu

K. Scalise

University of Oregon, Eugene, OR, USA

P. Gochyyev

University of California, Berkeley, CA, USA

© Springer International Publishing AG 2018

E. Care et al. (eds.), *Assessment and Teaching of 21st Century Skills*,
Educational Assessment in an Information Age,

DOI 10.1007/978-3-319-65368-6_12

Learning Analytics

A popular definition of learning analytics was adopted by the Society for Learning Analytics Research in 2011:

Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs.

Two other aspects should be considered: (i) how to interpret results, and (ii) how to choose data types and algorithms. It is important to reflect on the interpretation of the data analysis, over and above the results (Wilson 2005; Wilson et al. 2012, *in press*).

Surprisingly, this phase of interpretation is not a part of the LAK/SoLAR definition; that is “collection, analysis and reporting” does not explicitly include this critical aspect. A blind-spot like this can lead to the disastrous situation where, once results are reported, it is assumed that their meaning is self-evident.

Sound interpretation is facilitated by an evidence-based framework—an example is the four measurement principles of the BEAR Assessment System, described below (Wilson 2005; Wilson et al. 2012). In order to make useable claims about the learner, such a framework must be designed to link between the goals of the analysis back to the results (Mislevy et al. 2003; Wilson and Sloane 2000). In addition, a scientific interpretation must encompass the uncertainty, or range of error, present in the results.

This needs to have a sound evidentiary argument to support interpretations about learning and can be couched as either a posteriori (following the analysis) or a priori (in advance of the analysis). The a posteriori approach is an exploratory approach, which, for scientific purposes, will need confirmation by a second round of data collection. It is commonly called “data mining” (Papamitsiou and Economides 2014) or, sometimes, “machine learning,” as well as “unsupervised learning” (Russell and Norvig 2009), in contrast to the supervised learning concept described in the next paragraph, where models are set up with theoretical structures and/or empirical data. When there is a need to explore the patterns in the data sets in a context where not much is known, such exploratory approaches can be very useful.

In contrast, an a priori approach begins with a strong theory and/or prior empirical information, and is thus labeled as a confirmatory learning analysis. Russell and Norvig (2009) called it “supervised learning,” meaning that characteristics of the LA learning algorithms such as factors, weights, and network structures, are pre-set at some level of detail, based on prior data or theory.

Measurement Approach

For our analysis we use four principles of good assessment and measurement practice which are part of the BEAR¹ Assessment System (BAS: Wilson 2005). BAS delineates techniques used in the construction of high-quality assessments (see Fig. 12.1). The four principles (Wilson 2005), expressed in the context of technology-enhanced assessments and learning (Scalise et al. 2007) are:

- Principle 1: Assessments should be based on a developmental perspective of student learning.
- Principle 2: Assessments in learning should be clearly aligned with the goals of instruction.
- Principle 3: Assessments must produce valid and reliable evidence of what students know and can do.
- Principle 4: Assessment data should provide information that is useful to teachers and students to improve learning outcomes.

In Principle 1, the concept of a developmental perspective of student learning entails that we consider how student understanding of particular concepts and skills develops over time, rather than taking a one-shot view. This perspective requires a definition of what students are expected to know at particular points in their development, incorporated into a theoretical framework of how that learning unfolds as the student makes progress.

For Principle 2, the concept of establishing a good match between what is taught and what is assessed means that the goals of learning and the assessments should be directly related. This is the opposite of the situation where teachers interrupt their

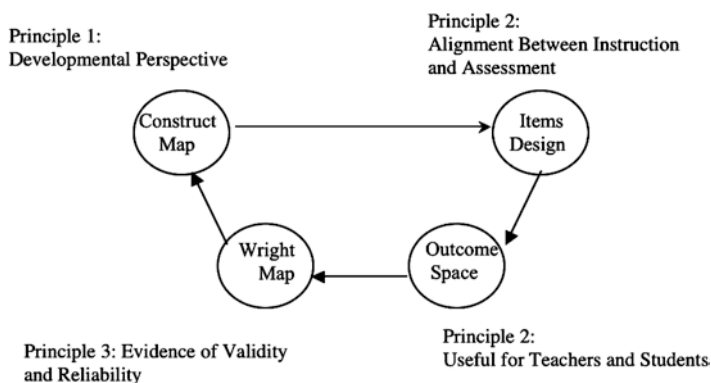


Fig. 12.1 A diagram of BAS, showing both the principles and four building blocks of measurement

¹BEAR Center = Berkeley Evaluation and Assessment Research Center.

regular curriculum progress to “teach the test” that students will encounter on summative tests.

Principle 3 addresses issues of technical quality in assessments. Numerous technology-enhanced learning assessment procedures are gaining “currency” in the educational community by making inferences about students that are supported by evidence for the validity and reliability of those inferences. Reliability concerns the consistency of results and validity relates to whether an assessment measures what it is intended to measure. To make results useful across time and context, these issues must be addressed in any serious attempt at technology-based measures.

Principle 4 is perhaps the most critical: learning assessment systems must provide information and interpretations that are useful for improving learning outcomes. Teachers must have efficient means to explain resulting data and make appropriate inferences. Students also should be participants in the assessment process, and their assessments should be designed to encourage the development of metacognitive skills that will further the learning process. If teachers and students are to be held accountable for performance, they need a good comprehension of what students are expected to learn and of what counts as sound evidence of student learning. Teachers are then in a better position, and a more central and responsible position, for presenting, explaining, analyzing, and defending their students’ performances and outcomes of their instruction.

These four principles summarize a way to understand the advantages and disadvantages of assessments, how to use such assessments, and how to apply these methods to develop new instruments or adapt old ones (Wilson 2005). These four principles match four “building blocks” (see Fig. 12.1) that make up an assessment—the construct map, the design plan for the items, the outcome space, and the statistical measurement model or algorithms to be used to compile and analyze patterns in the data.

Bringing the Two Perspectives Together

So here is the intersection of measurement technology and information technology. In the context of applying learning analytics to educational assessment, can the two perspectives work together to achieve something that is more than the sum of the two parts? To help answer this question, we next take up a brief example that incorporates the two in the context of looking at the assessment of collaborative learning in digital interactive social networks (Wilson and Scalise 2015).

The example we will use here is taken from the Assessment and Teaching of Twenty-First Century Skills project (ATC21S), and as both the project and the example have been described earlier in this Volume (Wilson et al. 2018), we will not describe them here, but assume that the reader has read that chapter.

The ATC21S demonstration scenario is conceived as a “collaboration contest,” or virtual treasure hunt. The Arctic Trek scenario conceptualises social networks through ICT as an assembly of different tools, resources and people that together

Norwegian Bay	190 (1998)	102-278					4	4	Data deficient	Declining	Very high
Southern Beaufort Sea	1526 (2006)	1210-1842					44	80	Reduced	Declining	Moderate
Southern Hudson Bay	900-1000 (2005)	396-950 (ON) 70-100 (James Bay)					35	61	Not reduced	Stable	Very high

Clue 3
Find the table that tells how many there are of me. How many colors are used to describe the bear population? (use status, trend and risk columns)

Answer: I looked and counted the colors 5 -Alex, Team Captain
I think there should be six because the "data deficient" is also a color. -Andy

Fig. 12.2 Example of student collaborative chat in Arctic Trek task

build a community in a relevant topic. In this task, students in small teams explore tools and approaches to unravel clues through the Go North site, by visiting information about scientific and mathematics expeditions of actual scientists.

In the Arctic Trek task challenge shown in Fig. 12.2, students must identify the colors that are used to describe the bear population in the table, a part of which is shown at the top. The highlighted chat log of students at the bottom of Fig. 12.2 (which actually takes the form of a collaborative laboratory notebook) indicates that students are indeed communicating to identify what is signal versus noise in the supplied information. The colors in the text are the colors shown in the columns on the right of the table. Requiring both identifying signal versus noise in information and interrogating data for meaning, this performance can be mapped into the ICN3 level (“Proficient builder”) of the ICN strand (Wilson et al. 2018). For further examples of activities and items from the Arctic Trek scenario, see Scalise (2018).

The connection between measurement science and learning analytics can be made in two ways in the context of this example. First, the statistical analysis approach used to estimate scores in measurement science is generically called a “measurement model.” It serves as an algorithm to gather the results together and make inferences about learners. Other fields such as computer science that have come to learning analytics from a different historical basis often use a different vocabulary to describe such algorithms. For instance the Rasch model often used in educational assessment from a computer science perspective would be considered as an LA algorithm employing a multilayer feed-forward network (Russell and Norvig 2009) with g as the Rasch function (a semi-linear or sigmoidal curve-fitting function), in which weights (item discrimination) are constrained to one for all inputs, and the item parameters estimated are the thresholds on each item node (item difficulty).

Secondly, the critical point we want to illustrate in this section is that additional tools from learning analytics can be added to or embedded within the traditional measurement model. Below we show an example of such embedding through an automated scoring engine. The scores produced by a scoring engine can be merged into a data set to be analysed by a measurement model. As an example, some of the complex student work products from the Arctic Trek module were also analysed

under a learning analytics approach called “sentiment analysis” which involves predictions of team success in the collaborative notebooks.

In this example, some notebooks that were identified to be used for a training set of the LA engine were initially hand-scored using traditional tools such as rubrics and exemplars. A collection of 28 hand-scored notebooks, which were the work products from approximately 112 students, provided this training set. The training set was then analysed by RapidMiner (Hofmann and Klinkenberg 2013) for the LA sentiment analysis approach.

Sentiment analysis in RapidMiner is an LA technique that aims to extract information from large full-text data sources such as online reviews and social media discussions. It can be used to interpret and optimize what is being thought, said, or discussed about a company or its products—or in this example, it is used to analyze what is being discussed in a collaborative learning situation the ATC21S Arctic Trek task.

The main idea in sentiment analysis is to classify an expressed opinion in a document, in a sentence or an entity feature, as positive or negative. In this example, “positive” means that the notebook shows some good evidence of learning in networks, based on the construct conceptualization described above. To calibrate the engine, first, both positive and negative “scores” of the task results are analysed—or in other words, a training set of scored collaborative notebooks are provided to the engine.

The engine first stems the words into root words. Then, a vector word list and a model are created. Using the training set, the model compares each word in the given notebook being considered with that of words that come under different predictions stored earlier. The notebook prediction is estimated based on the majority of words that occur under a polarity (i.e., a trend direction toward a negative or positive prediction). In this way, sentiment analysis is an artificial intelligence technique based on a “bag of words” (Russell and Norvig 2009). More sophistication can be added to the sentiment analysis data mining engine to include a variety of relationships between words, if desired, or data adjustments such as spelling corrections, “black lists” and “white lists” that are addendums or eliminations from the data dictionary, etc. An example of the sentiment analysis design window is shown in Fig. 12.3. The components of the full analysis for the Arctic Trek sentiment analysis engine used here are shown in Fig. 12.4.

Following the establishment of the training set, four additional collaborative notebooks were added to the work product data set for the sentiment analysis. These additional notebooks were not used for the sentiment analysis in the first instance. Rather, the LA engine was used to generate the prediction for each of the four notebooks. However, the four notebooks were also hand-scored in advance using the same human scoring approaches as for the other notebooks. The point was to see if the LA engine could match and even potentially add to the results generated by the hand-scoring.

Thus, this could provide some evidence that an LA sentiment analysis engine (in this case, RapidMiner) might effectively be incorporated into the measurement science approach. This could help to satisfy the measurement principle of usability by teachers and students, since an effective LA engine might eliminate some of the

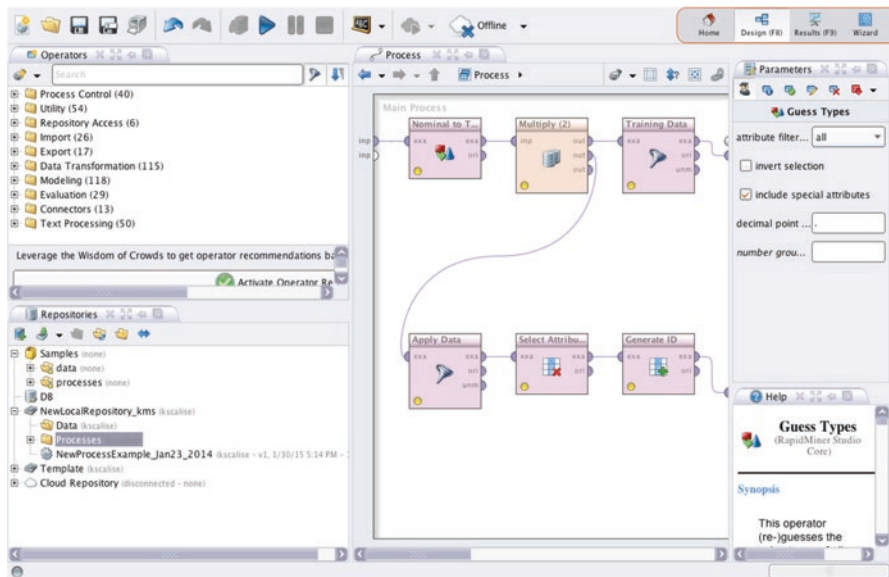


Fig. 12.3 Sentiment analysis design window for ATC21S example

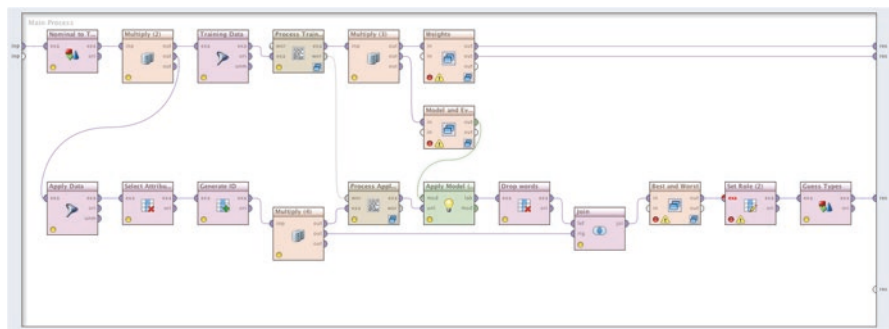


Fig. 12.4 Sentiment analysis component elements for LA engine in Arctic Trek

need for extensive hand-scoring. And, this could be applied to other, more complex learning and assessment activities.

The four notebooks selected were a small but very purposive sample for the engine to score. Only one notebook was high scoring according to the human rating (see Table 12.1). A second notebook with a low hand-score illustrated a similar level of text complexity but without as much substantively correct information and with little evidence of collaboration. Two additional notebooks were scored—they represented sparser and less correct scripts, with poor evidence of effective *learning in networks* practices, according to the construct ideas described above. All notebooks were supplied to the engine in their original formats, without editing or correction.

Table 12.1 Sentiment analysis results for Arctic Trek four notebooks

Notebook ID number	Sentiment ranking (pos/neg)	RapidMiner “Score”	Hand-score ranking and notes
A (original case number 32)	Positive	78.0	1 (only notebook of the four judged as high-scoring, illustrated strong elements of collaboration)
B (original case number 11)	Negative	46.0	2 (low-scoring notebook but with some beginner elements of collaboration; text complexity similar to notebook A above)
C (original case number 14)	Negative	39.0	3 tie (low-scoring notebook, few if any relevant elements of collaboration visible)
D (original case number 13)	Negative	35.0	3 tie (low-scoring notebook, few if any relevant elements of collaboration visible)

One caveat concerning the limitations that should be noted in advance of reporting the results is that this is a very small data set intended only to serve as an illustrative example and a larger set would be needed to provide a more formal example. Thus, this example should not be considered conclusive evidence of the sentiment engine here as being effective or ineffective for such purposes. Rather it should be considered as being illustrative of the general topic: the potential for positive interaction of measurement science and learning analytics. Typically, collaborative data sets based on teams of four result in fewer unique work products than result from individual assessments. A larger data set of 150–175 notebooks, (i.e., about 600–900 students), if composed of collaborative teams of four students per notebook, would be more desirable for training an engine. Furthermore, it should be noted that if other collaborative data sets were available, other LA techniques might be more desirable (Chi et al. 2008; Pirolli 2007, 2009; Pirolli et al. 2010; Pirolli and Wilson 1998).

A very small example of the results of the sentiment analysis is shown in Table 12.1. It shows that the LA sentiment engine in this case was able to rank the four notebooks in the same order as the hand-scoring. The highest scoring notebook was rated considerably higher than the next ranked notebook, in spite of the similar text complexity between the two notebooks. Furthermore, RapidMiner was also able to do a reasonable job of awarding “partial credit,” establishing a score substantially higher for the top notebook, but also ranking the next notebook somewhat higher than the other two, as had been the case for the human ratings. The notes in the hand-score ranking column provide some interpretive context for teachers and students and could be applied to the LA results as well, and mapped to the construct information described above.

If the improvement of twenty-first century skills such as digital collaboration for learning in social networks is a goal, it is important to help teachers understand what a successful performance looks like in a collaborative digital space. In addition, providing tools that populate the intersection of measurement science and LA, as described in this chapter, can help to inform teaching so that teachers know how such skills can be effectively assessed.

Discussion and Conclusion

What Measurement Can Learn from LA

Learning Analytics has evidenced a “brave-new-world” view in taking advantage of the new sources and large scope of data that have become available in this digital age. This has hugely expanded the types and volume of data available to education, and has opened unforeseen possibilities, from moment-to-moment data collection in educational settings to fine-grained records of interactive settings, such as one-on-one conversations and classroom discussions, to the data representation of objects that were previously not available to quantitative analysis, such as syntactic and content representations of document, student products, and so forth.

Not only is it the data being collected that is changing, it is the speed of collection, and the possibility of intelligent computer-generated feedback that opens up significant possibilities for education. Educators no longer have to wait for the data analysts to spend days (or weeks) analyzing the data and preparing reports. Teachers can have effectively instantaneous feedback, once the student has responded—it is this that holds the greatest promise. The impact of classroom assessment on student success has been well documented in a classic meta-analysis by Black and Wiliam (1998). But this historic level of impact had little to do with measurement since, in the past, the classroom environment was too fast-paced for the decidedly careful pace of traditional educational measurement. Partly by virtue of its usual funding sources (policy-level decision-makers), and partly due to the lack of appropriate technology, as noted above, measurement has been focused on large-scale samples of sparse data for each sampled student. What was useful for administrative and program evaluation purposes had no place in the most important site of educational change and improvement—the classroom. While early measurement scientists often had a strong domain grounding in what they were trying to measure (e.g. psychologists trying to measure psychological traits), measurement science has become its own sub-discipline, and much of the domain expertise has been lost directly by the psychometricians (Mislevy 2016). In contrast, LA researchers have worked to build strong, diverse teams that bring domain expertise back into play in ways from which measurement science can learn. These teams can tackle much more complex work products and data streams, but only because they ensure that they have educational professionals and domain analysts for the given area of interest working on the team.

What LA Can Learn from Measurement

The discussion above notes several potential strengths of the measurement approach as a framework for LA. First, whenever someone interprets LA student performance results, they are making certain assumptions. Over many years and across a wide range of contexts, the nature of these assumptions has been considered and

contested within the domain of the science of measurement. In the discussion above, we have emphasized the critical importance of having a scientific theory that is the basis for the interpretation of these results (the construct map in the context of the BAS—although, of course, there could be many other such bases). Equally, there needs to be an understanding of how the data generation model relates back to this scientific theory (this was embodied in the item design and the outcome space in the BAS). And, in order to evaluate how well the accumulated evidence relates to the hypothesised scientific construct, it is important to have a statistical model for estimation so that uncertainty can be included in the resulting outcomes (which is one aspect of the measurement model in the BAS).

In addition, Quality Control considerations need to be invoked, which are expressed in the measurement approach through the concepts of validity and reliability evidence (e.g., AERA/APA/NCME 2014), which constitute the grounds on which to be assured that the interpretations that analysts would like to make of the LA results are indeed valid.

No abundance of data (i.e., “big data”), nor frequency of responses, nor novelty of data-format, will eliminate the need for these issues to be considered and responded to. At the initial stages of implementation, it may be acceptable to ignore this need, but long practice in many different domains has told us that such ignorance is fraught with risk, not just for the Learning Analysts and their findings, but also for the students and teachers who rely on them.

What LA and Measurement Can Do Together

Perhaps even more important than what the two approaches can learn from each other is that they can benefit by working together. The small example above shows some of the overlaps and complementarities that can be seen to exist between the two approaches (with a little bit of cross-disciplinary insight). Our principal argument above is not based on necessary oppositions between the two, but rather on how they can be seen to offer ways to extend each other.

Learning Analytics and Measurement Science

Considering the discussions in this chapter, one can perceive new research directions at the intersection of LA and measurement. First, from the direction of how interactions with LA can improve and expand measurement science, we noted the following possibilities:

- (a) Measurement science needs to adapt its methods to the new directions that LA takes as standard, in particular to the gathering and analysis of new types of data relating to student behaviors beyond the standard measurement science formats

of the test and the questionnaire/survey—for instance to incorporate not just student “answers,” but also their many steps and actions that lead to those answers.

- (b) Measurement science also needs to explore the broader horizon of being able to examine real-time segments of student educational experiences—not just a single “test” event in a single classroom in a single year—by having access to the whole range of IT-enabled data that will be available regarding students. The very size of LA data sets is also a challenge to standard measurement science—the typical techniques of statistical analysis will have to give way to more flexible and faster algorithms and means of communicating results.

Second, thinking about how interactions with measurement science can improve and expand LA, one can see several possibilities. One possibility will include new LA algorithms and aggregation approaches. These are likely to be situated in data density, but they will also rely on more pattern finding and probably noisier patterns, with more construct irrelevant variance, included in less structured but larger data sets. A good direction for assessing efficacious algorithms and methods of classification and feedback specifically for educational applications will be to search for methods that add to the explained variance of models already employed in measurement science. As LA matures beyond a focus on predictive validity to the establishment of well-accepted procedures for quality and the adherence to strong measurement standards, new research directions will emerge in the science of LA assessment. These are likely to include technical studies and simulations to understand and address (a) reliability and precision information for LA, (b) assessment form creation, (c) linking and equating, (d) adaptive administrations, (e) the evaluation of data-generation assumptions, and (f) the checking of data-model fit. As LA opens up more opportunities for deeper assessment of hard-to-measure constructs that are instructionally relevant, the interpretive focus of LA will become more prominent. LA will need to add expertise regarding validity evidence for the interpretation of its outcomes: measurement science has had over 100 years of experience in this, and it will be much more efficient for LA to learn from that experience than to repeat that century of effort and thinking.

Contemplating this from both sides, an important area of research emerges related to improving and informing instruction. Research questions to be asked include:

- (a) How and whether teaching and feedback opportunities can enrich student learning outcomes, and
- (b) Whether they can address that need for *all* students, including disadvantaged students.

Technology can help to level the playing field and close achievement gaps, but it can also further marginalize some populations.

Thus, we see a need for new research and development projects that combine the two approaches. Such projects must provide for a wide dissemination of research outcomes and products in order to reach the many widely distributed fields of appli-

cation, which often do not share the same resource spaces. Joint publication of books that combine the approaches and synthesize approaches would be helpful. And advanced training programs are needed that combine the two, both for graduate students and for working professionals and academics.

In conclusion, as we enter a brave new world of digitally-extended data collection, we need to match the fearlessness of LA with the strength and re-assurance of measurement science.

References

- American Educational Research Association, American Psychological Association, National Council for Measurement in Education. (AERA, APA, NCME, 2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. doi:10.1080/0969595980050102.
- Chi, E. H., Pirolli, P., Suh, B., Kittur, A., Pendleton, B., & Mytkowicz, T. (2008). *Augmented social cognition*. Palo Alto: Palo Alto Research Center.
- Hofmann, M., & Klinkenberg, R. (2013). *RapidMiner: Data mining use cases and business analytics applications*. Boca Raton: CRC Press.
- Mislevy, R.J. (2016). Discussion of learning analytics.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). *A brief introduction to evidence-centered design, CRESST Technical Paper Series*. Los Angeles: CRESST.
- Papamitsiou, Z., & Economides, A. A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, 17(4), 49–64.
- Pirolli, P. (2007). Cognitive models of human-information interaction. In F. T. Durso (Ed.), *Handbook of applied cognition* (pp. 443–470). New York: Wiley.
- Pirolli, P. (2009). *An elementary social information foraging model*. Paper presented at the CHI 2009, ACM Conference on Human Factors in Computing Systems, Boston, MA.
- Pirolli, P., & Wilson, M. (1998). A theory of the measurement of knowledge content, access, and learning. *Psychological Review*, 105(1), 58–82.
- Pirolli, P., Preece, J., & Shneiderman, B. (2010). Cyberinfrastructure for social action on national priorities. *IEEE Computer*, 43(11), 20–21.
- Russell, S., & Norvig, P. (2009). *Artificial intelligence, a modern approach* (3rd ed.). Upper Saddle River: Prentice Hall.
- Scalise, K. (2018). Next wave for integration of educational technology into the classroom: Collaborative technology integration planning practices. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and teaching of 21st century skills: Research and applications*. Dordrecht: Springer.
- Scalise, K., Bernbaum, D. J., Timms, M. J., Veeragoudar Harrell, S., Burmester, K., Kennedy, C. A., & Wilson, M. (2007). Adaptive technology for e-learning: Principles and case studies of an emerging field. *Journal of the American Society for Information Science and Technology*, 58(14), 001–015.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah: Lawrence Erlbaum Assoc.
- Wilson, M., & Scalise, K. (2015). Assessment of learning in digital networks. In P. Griffin & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach* (pp. 57–81). Dordrecht: Springer.

- Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education, 13*(2), 181–208.
- Wilson, M., Bejar, I., Scalise, K., Templin, J., Wiliam, D., & Torres Iribarra, D. (2012). Perspectives on methodological issues. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills*. New York: Springer.
- Wilson, M., Scalise, K., Gochyyev, P. (in press). ICT literacy – learning in digital networks. In R. W. Lissitz, & H. Jiao (Eds.), *Technology enhanced innovative assessment: Development, modeling, and scoring from an Interdisciplinary perspective*. Charlotte: Information Age Publisher.
- Wilson, M., Scalise, K., & Gochyyev, P. (2018). ICT literacy in digital networks. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and teaching of 21st century skills: Research and applications*. New York: Springer.