

Assessment of Learning in Digital Interactive Social Networks: A Learning Analytics Approach

Mark Wilson, Perman Gochyyev
University of California, Berkeley

Kathleen Scalise
University of Oregon

Abstract

This paper summarizes initial field-test results from data analytics used in the work of the Assessment and Teaching of 21st Century Skills (ATC21S) project, on the “ICT Literacy — Learning in digital networks” learning progression. This project, sponsored by Cisco, Intel and Microsoft, aims to help educators around the world enable students with the skills to succeed in future career and college goals. The paper begins with describing some expansions to a common definition of learning analytics, then includes a review of the literature on ICT literacy, including the specific development that led to the ATC21S effort. This is followed by a description of the development of a “learning progression” for this project, as well as the logic behind the instrument construction and data analytics, along with examples of each. Data were collected in a demonstration digital environment in four countries: Australia, Finland, Singapore and the U.S. The results indicate that the new constructs developed by the project, and the novel item forms and analytics that were employed, are indeed capable of being employed in a large-scale digital environment. The paper concludes with a discussion of the next steps for this effort.

Acknowledgements: We thank the ATC21S project and its funders for their support for the work reported in this report. We also acknowledge the expertise and creative input of the ATC21S Expert Panel in ICT Literacy: John Ainley (Chair), Julian Fraillon, Peter Pirolli, Jean-Paul Reeffer, Kathleen Scalise, and Mark Wilson. Of course, the views and opinions expressed in this paper are those of the authors alone.

Introduction

The view of learning analytics that this paper is based on starts with the observation that the current practices of schooling are somewhat outmoded in the new global working environment. Today, people work both individually and in groups to share complementary skills and accomplish shared goals—this practice contrasts to schools and assessments where students take tests individually.

Knowledge is applied across disciplinary boundaries in the process of solving real-world problems, but in schools, the school subjects are divided by disciplinary boundaries. Furthermore, at work, problem solving is often complex and ill-structured, in contrast to the simplified problem-types featured in much of school education, and especially school standardized testing. Finally, in the work setting, people have access to enormous resources of information and technological tools, where the challenge is to strategically craft a solution process, which contrasts strongly with the traditional “closed book” analytics of what learners know and can do (CIM, 2008).

As a response to such transitions, the Assessment and Teaching of Twenty-First Century Skills project (ATC21S) was launched in 2009. A goal was to employ new analytical approaches to the assessment of learning such that the challenges above could be better addressed. We will utilize an example of work from this project to illustrate the points we are making in this paper.

A commonly used definition of learning analytics that we will draw on here was proposed by the first international Conference on Learning Analytics and Knowledge (LAK 2011) and adopted by the Society for Learning Analytics Research (Society for Learning Analytics Research, 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.”

While this definition is helpful, two additional aspects are important to consider: the interpretation of results, and the choice of appropriate data types and algorithms.

The paper begins with a description of some expansions to a common definition of learning analytics, then includes a review of the literature on ICT literacy, including the specific development that led to the ATC21S effort. This is followed by a description of the development of a “learning progression” for this project, as well as the logic behind the instrument construction and data analytics, along with examples of each. The paper goes on to describe the data that was collected, and reports on the learning analytics approach that was used. The paper concludes with a discussion of the next steps for this effort.

Bridging the Gap between Evidence and Interpretation: Some Background

First, in this paper we make the point that, for learning analytics, the meaningful interpretation of the data analysis is critical to consider, not simply reporting the results (Wilson, 2005; Wilson et al., 2010; Wilson, Scalise, & Gochyyev, 2014). Interpretation is not directly included in the LAK/SoLAR definition of “collection, analysis and reporting.” This weakness in the definition can lead to the assumption that once results are composed and reported, their meaning for learners and learning outcomes is self-evident.

However, that said, learning analytics as described in the LAK/SoLAR definition constitute a type of educational assessment. As such, meaningful interpretation means having an evidentiary framework designed to connect results clearly and on an empirical basis back to the goals and objectives of the analysis in order to make clear evidentiary claims about the learner (Mislevy, Almond, & Lukas, 2003; Wilson, 2005). It also means being able to understand the uncertainty or range and degree of error likely to be present in the results.

Some groups have begun to establish standards of practice in learning analytics for such 21st century complex data analysis methodologies (Sclater, 2014; Wilson et al., 2012). In this paper, we will present an example that helps establish the coherent evidentiary argument for the learning analytics involved through a framework called a “learning progression.” This framework connects the results to (a)

the data and the learning analytics? questions being asked, and (b) to the techniques for the analytics employed. Other researchers have begun to describe the need for such frameworks when learning analytics goes beyond data analysis alone and is to be used for predictive analytics, actionable intelligence, and decision-making (van Barneveld, Arnold, & Campbell, 2012). In learning analytics, the need to establish a coherent evidentiary argument to support claims about learners can be approached either *a priori* (in advance of the analysis) or *a posteriori* (following the analysis). The *a priori* approach is essentially a theoretical approach, based on a strong theory or prior empirical information (or both), and thus might be considered a confirmatory learning analytic technique.

The *a posteriori* approach can be considered generative, or in other words, exploratory, and in many cases will need to be confirmed by a subsequent data collection and analysis. The exploratory approach is sometimes called by the name “data mining” (Papamitsiou & Economides, 2014). Exploratory approaches can be useful when the research goal is to learn more about the patterns in the data sets in a context where little is yet understood, or where new patterns may become evident that were not previously suspected.

The choice between an exploratory or confirmatory approach is an option, with the choice depending on how much prior theory and/or empirics are available. Put together, these exploratory and confirmatory stages can be seen as a cycle in the evidence chain, as shown in Figure 1. The figure depicts a simple example of a learning analytics interpretive cycle, where entry points can be either confirmatory—entering at the “theory or conceptualization” node—or exploratory—entering at “analysis and results” node—for extant data or at the “empirical data” node when observations will be designed and collected (see below for a discussion of extant and collected data).

No single entry point to the cycle is best among the choices in every situation. The choice can be determined by the intended *purposes* of interpretation, and the current *state of claims* that can be made in a given context. In any particular situation, one relevant question to ask is does the analysis begin with an interpretive framework *a priori*, as in the theory component of the cycle below, or is interpretation made *a posteriori*, as when even the initial interpretive framework is derived from data because little is yet known?

In either case, the same cycle is present but with different points of entry and a different flow to the interacting elements.

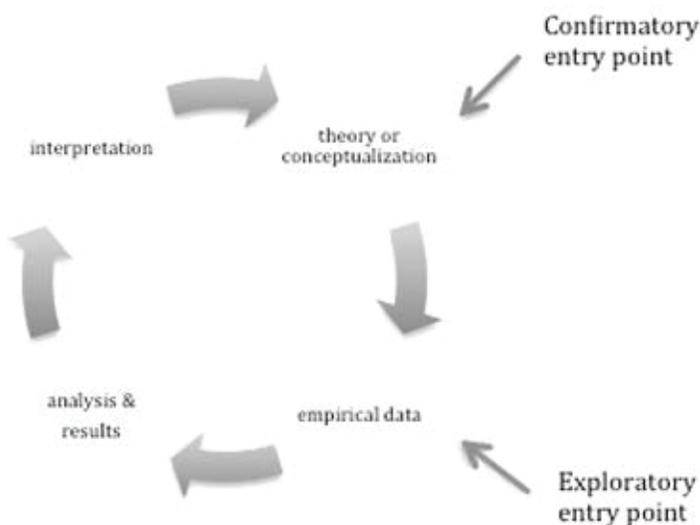


Figure 1. Choice of appropriate data types and algorithms

In terms of data types for which learning analytics by the LAK/SoLAR definition are likely to be useful, in most cases complex data should be involved. If not, other simpler techniques might be better employed (Ferguson, 2012). Complex data can take the form of large data sets (big data), multi-faceted data sets, or

other elements in the data that encode more complex patterns or hard-to-measure constructs not readily identifiable without complex analytic techniques (Scalise, 2012; Wilson et al., 2010).

The unit of analysis itself can also generate data complexity, for instance, when data are collected for individuals but the individuals are collected into groups, programs, institutions, or other collective units, for which data are also collected.

Sometimes data sets can be pre-existing, or extant data sets, as described above. Examples of pre-existing data include downloads from Twitter feeds, click streams in user data, or other online collections generated as a result of processes with various purposes (Baker & Siemens, 2014). At other times, data sets are collected at least in part directly for the purpose of applying learning analytics to the results. Data collection can include, for instance, an adaptive recommender where ratings on prior experiences are solicited for the purposes of prediction of respondent interest in future experiences (Chedrawy & Abidi, 2006; Dagger, Wade, & Conlan, 2005), or evidentiary data collection for educational or professional development, to address personalized or grouped components to support the learner in educational assessment (Brady, Conlan, Wade, & Dagger, 2006; Kennedy & Draney, 2006).

An extension to the LAK/SoLAR definition we propose here is the specification that complex analytic techniques are needed to resolve the multifaceted or complex patterns. The same argument can be made as above for data sets. Complexity should be introduced in the analysis for a coherent evidentiary argument only when necessary. So the usual parsimonious definition should be applied when models or other algorithms are used to fit learner data and resolve patterns.

Finally, it would be helpful if the LAK/SoLAR definition made reference to algorithms, or characteristics of algorithms, that might be useful to apply for aggregating and parsing of patterns, since this is an important consideration in the use of learning analytics (Papamitsiou & Economides, 2014). While it is important to keep the definition general to be inclusive of many useful algorithms that might arise, as a general class the approach typically needs to involve algorithms to automatically process the data, assuming the purposes of interpretation and the complexity of data require algorithmic approaches to the accumulation and parsing of patterns in the data. Algorithms can be statistical in nature, applied as inferential statistical tests or to yield inferential indices as part of the processing, which can help with assessing quality of results (Sclater, 2014). Numerous algorithms in the form of measurement models that take a statistical form for learning outcomes have been created and applied. These are well established in the psychometrics research literature and some of the advanced models as well as basic models can be appropriate to apply in learning analytics to complex 21st century skill settings (Wilson et al., 2012). Algorithms can also process patterns in more descriptive ways, yielding machine-readable results such as categorization or subsetting of respondents (Stanton, 2012). Note that since machine processing is required, however, the data sets at some point have to include machine-readable data. This may be text-based or graphical in nature, or in some other innovative format, depending on the processing requirements of the algorithm and platform, or the data sets may be numeric (Scalise & Gifford, 2006). The desired data characteristics may already be present for a given data set in any particular case, or may require preprocessing. This could include types of scoring, ordering, subsetting, or other types of aggregation. For this, reliable data collection, warehousing and prep can be a problem, so a variety of “clean-up” procedures may be needed. An important stage in learning analytics is reducing construct irrelevant variance including noise, user errors, or out-of-scope entry of data, which should be clarified and validated before conclusions can be drawn (Dringus, 2012).

In light of these sets of clarifications, we suggest a revision to the LAK/SoLAR definition, which we propose as “Learning analytics definition, LAK/SoLAR.v2”:

“Learning analytics is the measurement, collection, analysis, *interpretation*, and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs, by means of a coherent evidentiary argument... Complexity should be introduced in the data and the analysis only when necessary to the development of the evidentiary argument.” Complex data as described here can take the form of large data sets (big data), multi-faceted data sets, and/or other the data elements that encode patterns or hard-to-measure constructs not readily identifiable without advanced analytic techniques.

In the sections that follow, we introduce the ATC21S project, which is the context of our study, and review the literature on ICT Literacy, culminating in a section on the inception of the ATC21S ICT Literacy effort. We then give a very brief description of the BEAR Assessment System mentioned above, as that is used as the organizing principle for the next four sections: The construct map, the items design, the outcome space, and the measurement model. This is followed by a description of the key results from an empirical study of the assessments using a sample of students from four countries: Australia, Finland, Singapore and the United States. The paper finishes with a discussion of conclusions and next steps.

Overview of the ATC21S project

ATC21S was founded, in an unusual collaborative effort, by three information technology companies (Cisco, Intel, Microsoft—CIM). According to the company collaborative:

The economy of leading countries is now based more on the manufacture and delivery of information products and services than on the manufacture of material goods. Even many aspects of the manufacturing of material goods are strongly dependent on innovative uses of technologies. The start of the 21st century also has witnessed significant social trends in which people access, use, and create information and knowledge very differently than they did in previous decades, again due in many ways to the ubiquitous availability of ICT. (CIM, 2008)

It is to be expected that this widespread change will significantly affect the personal and working lives of many people, and hence should have equally large effects on the school systems that educate people for their later lives and careers. These effects may include the nature of and even the names of the subjects that are taught in schools, how those new subjects (and the traditional subjects) are taught and learned, and how education is organized. Thus, the ATC21S project was initiated to develop new assessments in the area of 21st century skills, based on the idea that new assessments could lead the way to these new subjects. The project determined to develop sample assessments in a few specific domains of the 21st century skills, to serve as examples leading to further developments in learning analytics in those areas.

The project recruited a large group of experts in diverse fields to author a set of “white papers” that paved the way for further work. This process is documented in Griffin, McGaw & Care (2012), and two chapters from that volume are germane to the efforts delineated here: the “skills white paper” (Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci & Rumble, 2012); and the “methodology white paper” (Wilson, Bejar, Scalise, Templin, Wiliam, & Torres-Irribarra, 2012). The skills white paper develops a framework for describing and implementing the 21st century skills and how they relate to traditional school subjects. This framework is labeled “KSAVE,” for Knowledge, Skills and Attitudes, Values and Ethics. Two particular 21st century skills were chosen for inclusion in an ATC21S demonstration project—collaborative problem solving and ICT literacy. This second 21st century skill is the focus of this paper, and our particular approach to that construct is described in this paper. The “methodology white paper” describes a method to develop the new assessments (NRC, 2001). The approach chosen is called the BEAR Assessment System (BAS: NRC, 2001; Wilson, 2005; Wilson,

2009a; Wilson & Sloane, 2001), and it will also be briefly introduced here, though the reader is directed to the references just cited for a more complete account.

Brief Literature Review of the Domain

For this project, learning analytics were applied in an *a priori* approach, entering at the theory/conceptualization node of the learning analytics interpretive cycle discussed in the prior section, since considerable information about the domain is already available. Information about the domain is based on prior empirical results and theory briefly summarized in this section.

The idea of learning through digital networks and the use of digital media, in the context of schooling, has been based on the conceptualization of student use of information and communication technology, or “ICT” literacy. To wit, information technologies are seen as resources for creating, collecting, storing, and using knowledge as well as for communication and collaboration (Kozma, 2003).

The definitions of information and communication literacy range from simple—basic digital knowledge and skills—to complex, a wide range of tools and competencies that ATC21S describes as “tools for working” in a digital age (Binkley et al., 2012). These generalized tools include accessing and using technologies to evaluate and utilize information, analyzing media, creating information products, and they include understanding ethical and legal issues involved in being a participant in the knowledge economy.

Efforts to measure ICT practices in schools go back about three decades—this history has recently been reviewed by Wilson, Scalise, & Gochyyev, (2015) so we will not attempt to duplicate that history here. In that paper, the authors described how the concept of ICT literacy has changed a great deal during these years: From a conceptualization as a specific domain of knowledge about computers, to an understanding of it as a domain-general or transversal 21st century skill. They traced the conceptual changes in the idea of ICT literacy in four main steps, as follows.

First, ICT literacy was seen as measurement of a discrete set of knowledge about computers and their use, coalescing into the concept of ICT literacy in the early years of the field. Second, ICT literacy was seen as measuring a broad set of skills that have links to many traditional and non-traditional school subjects, incorporating the move to technology integration in education. Third, ICT literacy was seen as measuring a set of *progress variables*, which are essential tools for the design of curriculum and assessments in terms of competencies. The “competencies” view depicts the need to understand initial ICT knowledge likely to emerge followed by a developing picture of mastery. In this perspective, students are acknowledged as not being one-size-fits-all in ICT literacy but moving toward increasing competency in their virtual skills, knowledge, competency, awareness, and use. Fourth, the measurement is seen as needing to reflect the “network” perspective on ICT—the critical need for building the power of virtual skills through proficiency with networks of people, information, tools, and resources. A new framework for assessing student ICT learning, based on a learning progression and network point of view has been offered in Wilson & Scalise (2014). This paper goes beyond that introduction and presents an initial analysis of the results from a pilot test looking across the multiple dimensions of the construct.

Methods

The ATC21S methodology group has described that, in order to achieve a working hypothesis of such a complex domain, one *a priori* approach by which known information is incorporated is to describe “dimensions of progression,” or theoretical maps of intended constructs, in terms of depth, breadth and how the skills change as they mature for students (Wilson, et al., 2012). For this, the ATC21S project set up an expert panel of ICT experts who turned to the research literature to inform expanded definitions of digital literacy.

The ATC21S panel of ICT experts looked into a wide range of topics in laying out the background for their work in developing the learning progression, including in the areas of augmented social cognition (Chi et al., 2008), applied cognition (Rogers, Pak, & Fisk, 2007), team cognition (Cooke, Gorman, & Winner, 2007), social participation (Dhar & Olson, 1989), cognitive models of human-information interaction (Pirolli, 2007), technological support for work group collaboration (Lampe et al., 2010; Pirolli, Preece, & Shneiderman, 2010), theories of measurement for modeling individual and collective cognition in social systems (Pirolli & Wilson, 1998), topics in semantic representation (Griffiths & Steyvers, 2007), and social information foraging models (Pirolli, 2009).

To take just one example, the research in augmented social cognition field (Chi et al., 2008) describes the emergence of the ability of a group of people to remember, think and reason together. This field investigates how people augment their speed and capacity to acquire, produce, communicate and use knowledge, and to advance collective and individual intelligence in socially mediated environments. This is seen as very relevant to the ICT field, as it is expected that augmented digital and virtual settings will be increasingly common learning environments for students in the 21st Century.

The ATC21S panel of experts identified a set of distinctive ICT literacy goals for students. Earlier frameworks had included individual consumer skills, often on a Web 1.0 model of data repositories that could be accessed over the internet by students. A variety of producer skills, in which students needed to manipulate digital assets in new ways—due to the emergence of Web 2.0 technologies—were seen as emerging trends. Lastly, they found that, as described in documents from the National Initiative for Social Participation (NISP) (e.g., Lampe et al., 2010; Pirolli, Preece, and Shneiderman, 2010), the field of ICT literacy was on the cusp of recognizing the importance of digital networks to education, which would require from the students both “social capital” skills and the ability to draw on the “intellectual capital” of groups and teams. Included in intellectual capital are the Web 3.0 skills of “semantics,” or meaning-making through technology, including complex tools such as analytics, the effective use and evaluation of ratings, crowd-sourcing, peer evaluation, tagging, and the ability to judge the credibility and viability of sources.

The expert panel then challenged itself to define, for each of the competencies within the learning progression, increasing levels of sophistication of student competence, that is, to describe what having “more” and “less” of the competency would look like, for age 11-, 13- and 15-year-old students. In other words, as one expert noted, “When someone gets better at it, what are they getting better at?” This might also be thought of as describing how the field of education will document the range of competencies that we might expect as students develop in sophistication.

The BEAR Assessment System

The BEAR Assessment System (BAS) consists of employing data analytics structured around four basic principles of assessment, and four accompanying building blocks that are tools to help develop the assessment. The principles and building blocks are as follows:

Principle 1: Assessment should be based on a developmental perspective of student learning; the building block is a *construct map* of a progress variable that visualizes how students develop and how we think about their possible changes in response to items. Data analytics in this case are structured around the theoretical conception of the construct map.

Principle 2: There must be a match between what is taught and what is assessed; the building block is the *items design*, which describes the most important features of the format of the items—the central issue, though, is how the items design results in responses that can be analytically related back to the levels of the construct map.

Principle 3: Teachers must be the managers of the system, with the tools to use it efficiently and

effectively; the building block is the *outcome space*, or the set of categories of student responses that make sense to teachers. These categories of student responses become the core of quantitative measures for conclusions through the data analytics.

Principle 4: There is evidence of quality in terms of reliability and validity studies and evidence of fairness, through the data analytics; the building block is an algorithm specified as a *measurement model* that provides for a visual representation of the students and the items on the same graph (called a “Wright Map”), and a number of other data analytic tools that are helpful for testing the quality of the measurement. (Wilson, 2005)

How these principles become embedded in the process and the product of the assessment development, and the nature of these building blocks, is exemplified in the account below.

The Construct Maps. Levels in the ATC21S Learning in Digital Communities learning progression framework, as described below, were developed using the BAS approach described in the previous section, and draw substantively from the literature noted above. The levels in each strand follow a similar valuing, starting first with awareness and basic use of tools, and then building to more complex applications. Evaluative and judgmental skills improve. Leadership and the ability to manage and create new approaches emerge as the more foundational skills are mastered.

The skills can be seen as building mastery, in a theoretical sense, but more information from student work is needed to help validate frameworks. Examples of student work can be collected in cognitive laboratories as was done here or in other venues. Such frameworks become especially useful if ranges of performances can be collected that successfully populate the strands and levels of expected performances. Thus a working definition of the competencies is strengthened through exemplification.

Further large-scale data collection and analysis of results also helps to validate, or refine and reconfigure meaningful frameworks. An important consideration in emerging areas such as these is not only what student performances are currently, but what teachers, schools and states would like them to be. Validated working definitions help school systems plan instructional interventions and assessments, and work with teachers for professional development.

The Four Strands. For the ATC21S project effort, ultimately the focus of ICT Literacy was on *learning in networks*, which was seen as being made up of four strands of a learning progression:

- Functioning as a consumer in network;
- Functioning as a producer in networks;
- Participating in the development of social capital through networks;
- Participating in intellectual capital (i.e., collective intelligence) in networks.

The four strands are seen as interacting together in the activity of learning in networks. They are conceptualized as parallel developments that are interconnected and make up that part of ICT literacy that is concerned with learning in networks.

First, functioning as a Consumer in Networks (CiN) involves obtaining, managing and utilizing information and knowledge from shared digital resources and experts in order to benefit private and professional lives. It involves questions such as:

- Will a user be able to ascertain how to perform tasks (e.g. by exploration of the interface) without explicit instruction?

- How efficiently does an experienced user use a device, application, or other ICT strategy to find answers to a question?
- What arrangement of information on a display yields more effective visual search?
- How difficult will it be for a user to find information on a website?

Second, functioning as a Producer in Networks (PiN) involves creating, developing, organizing and re-organizing information/knowledge in order to contribute to shared digital resources.

Third, developing and sustaining Social Capital through Networks (SCN) involves using, developing, moderating, leading and brokering the connections within and between individuals and social groups in order to marshal collaborative action, build communities, maintain an awareness of opportunities and integrate diverse perspectives at community, societal and global levels.

Fourth, developing and sustaining Intellectual Capital through Networks (ICN) involves understanding how tools, media, and social networks operate, and using appropriate techniques through these resources to build collective intelligence and integrate new insights into personal understandings. An initial idea of the relationship amongst these four strands of the learning progression is given in Figure 2. The four columns are the four strands, each with defined levels within them (exemplified in the next paragraph), and the arrows show expected tracks of development among these constructs. An example of an activity that is hypothesized to map into the ICN2 level (“Functional builder”) of the Intellectual Capital strand is demonstrated in the chat log in Figure 3. In this example, students are creatively using tags to make their communication clearer. Specifically, the team members have devised a strategy using different fonts and/or colors to distinguish what each team member is typing.

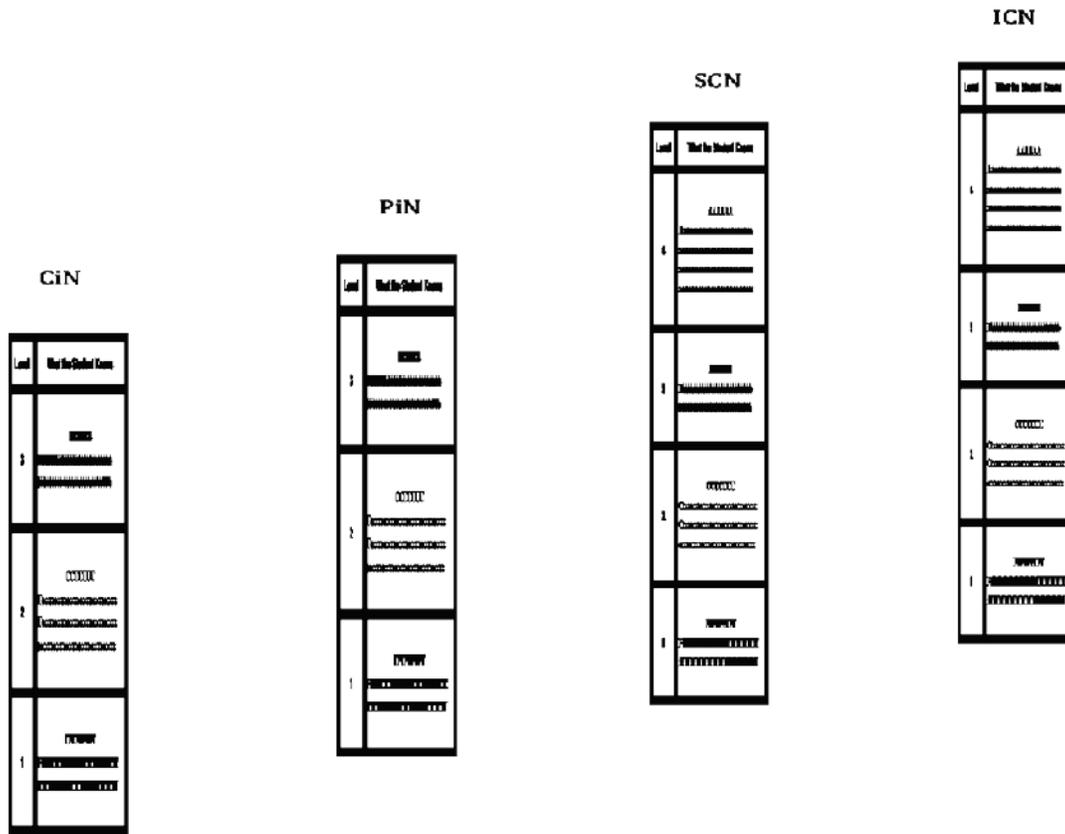


Figure 2. The four strands of ICT Literacy, represented as a four-part learning progression.

```

cool writing WHAT IS THE SECRET CODE FOR THE NOTEBOOK
then go back and write in the box? i knew cause nelson - U R SITTING RIGHT NEXT 2 ME!
IM confused!!!

how did you no it was me nelson janelle ?
okay i need help too! Isabella. NO this is Janelle. I need help on the 6. who is that okay
Nelson. ok
im on number 13 wow
. You said it nelson is any one still here but me NO it me nelson what number
are you on
giovany im on 14

```

Figure 3. Students' chat log.

In Table 1 the levels of the second of these four strands have been described as a hypothesized construct map showing an ordering of skills or competencies involved in each. At the lowest levels are the competencies that one would expect to see exhibited by a novice or beginner. At the top of the table are the competencies that one would expect to see exhibited by an experienced person – someone who would be considered very highly literate in ICT. The construct map is hierarchical in the sense that a person who would normally exhibit competencies at a higher level would also be expected to be able exhibit the competencies at lower levels of the hierarchy. The maps are also probabilistic in the sense that they represent different probabilities that a given competence would be expected to be exhibited in a particular context rather than certainties that the competence would always be exhibited. For a complete account of the construct maps for the four competencies, see Wilson & Scalise (2014).

Table 1. Functioning as a Producer in Networks (PiN)

PRODUCER IN NETWORKS	
	<i>Creative producer</i>
PiN3	<ul style="list-style-type: none"> Team situational awareness in process Optimize assembly of distributed contribution to products Extending advanced models (e.g. business models) Producing attractive digital products using multiple technologies / tools Choosing among technological options for producing digital products
	<i>Functional producer</i>
PiN2	<ul style="list-style-type: none"> Establishing and managing networks & communities Awareness of planning for building attractive websites, blogs, games Organizing communication within social networks Developing models based on established knowledge Developing creative & expressive content artifacts Awareness of security & safety issues (ethical and legal aspects) Using networking tools and styles for communication among people
	<i>Emerging producer</i>
PiN1	<ul style="list-style-type: none"> Produce simple representations from templates Start an identity Use a computer interface Post an artifact

The Items Design. The Berkeley Evaluation and Assessment Research (BEAR) Center at UC Berkeley developed three scenarios in which to place tasks and questions that could be used as items to indicate where a student might be placed along each of the four strands.

Observation formats included numerous innovative types (Scalise & Gifford, 2006). The data set included process data from activities in the tasks, collaboration data from chat logs and other activities, explanations and argumentation from students, and answers to technology-enhanced selected response queries of various types. Some examples are illustrated in the upcoming figures, when each task is introduced.

Each scenario was designed to address more than one strand of the construct, but there were different emphases in how the strands area were represented among the scenarios. Where possible, we took advantage of existing web-based tools for instructional development. These are each briefly described below.

Arctic Trek. One potential mechanism for the assessment of student ability in the learning network aspect of ICT literacy is to model assessment practice through a set of exemplary classroom materials. The module that has been developed is based on the Go North/Polar Husky information website (www.polarhusky.com) run by the University of Minnesota. The Go North website is an online adventure learning project based around arctic environmental expeditions. The website is a learning hub with a broad range of information and many different mechanisms to support networking with students, teachers and experts. ICT literacy resources developed relating to this module focus mainly on the functioning as a Consumer in Networks strand. The tour through the site for the ATC21S demonstration scenario is conceived as a “collaboration contest,” or virtual treasure hunt (see Figure 4 for a sample screen). The Arctic Trek scenario views social networks through ICT as an aggregation of different tools, resources and people that together build community in areas of interest. In this task, students in small teams ponder tools and approaches to unravel clues through the Go North site, via touring scientific and mathematics expeditions of actual scientists. The Arctic Trek task, in which students work in teams, is demonstrated in Figure 5. In that task, students are expected to find the colors that are used to describe the bear population in the table, part of which is shown at the top. The highlighted chat log of students at the bottom of the figure indicates that students are communicating in order to identify 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 & Scalise, 2014). For

further examples of activities and items from the Arctic Trek scenario, see Wilson & Scalise (2014).

Figure 4. A sample screen from the Arctic Trek task: the Welcome screen.



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

Figure 5. Example of students' work analyzing the table.

Webspiration. In the second demonstration task, framed as part of a poetry work unit, students of ages 11-15 read and analyze well-known poems. In a typical school context, we might imagine that the teacher has noticed that his or her students are having difficulty articulating the moods and meanings of some of the poems. In traditional teacher-centered instruction regarding literature, the student role tends to be passive. Often, teachers find that students are not spontaneous in their responses to the poems, but may tend to wait to hear what the teacher has to say, and then agree with what is said. To help encourage students to formulate their own ideas on the poems, we use a collaborative graphic organizer through the Webspiration online tool. The teacher directs the students to use Webspiration to create an idea map collaboratively using the graphic organizer tools, and to analyze each poem they read. Students submit their own ideas and/or build on classmates' thoughts. For examples of activities and items from the Webspiration scenario, see Wilson & Scalise (2014).

Second Language Chat. This scenario was developed as a peer-based second language learning environment through which students interact in learning. Developing proficiency in a second language (as well as in the mother tongue) requires ample opportunities to read, write, listen and speak. This assessment scenario asks students to set up a technology/network-based chat room, invite participants, and facilitate a chat in two languages. It also involves evaluating the chat and working with virtual rating systems and online tools such as spreadsheets. Worldwide, "conversation partner" language programs such as this have sprung up in recent years. They bring together students wishing to practice a language with native speakers, often in far-flung parts of the world. The cultural and language exchanges that result demonstrate how schools can dissolve the physical boundaries of walls and classrooms. Collaborative activities also tap rich new learning spaces through the communication networks of ICT literacy. This task shows how collaborative activities can also provide ample assessment opportunities in digital literacy. For examples of activities and items from the Second Language Chat scenario, see Wilson & Scalise (2014).

The Outcome Space. Each item that was developed was targeted at one or more of the four strands, and the expected range of levels that would be represented in the item responses were also noted. Where the responses are selected from a fixed set (as in a multiple-choice item), the set can be planned ahead of time, but for open-ended tasks and activities, the resulting work product is more complex. Examining the student work product is a something that needs to be empirically investigated. The tabulation is shown in Figure 6. As can be seen, the first three levels were reasonably well-covered, but level 4, which we seldom expect to see for students in this population, had only one instance.

ICT Literacy—Learning in digital communities					
Levels (Progressive)	CONSTRUCT/Learning Outcomes				Total
	Consumer	Producer	Social Capital	Intellectual Capital	
Level 4	N/A	N/A	Web 0 Arctic 1 2LChat 0	Web 0 Arctic 0 2LChat 0	Web: 0 Arctic: 1 2LChat: 0
Level 3	Web 0 Arctic 2 2LChat 0	Web 0 Arctic 2 2LChat 0	Web 0 Arctic 6 2LChat 1	Web 10 Arctic 2 2LChat 1	Web: 10 Arctic: 12 2LChat: 2
Level 2	Web 8 Arctic 6 2LChat 0	Web 4 Arctic 16 2LChat 8	Web 7 Arctic 0 2LChat 6	Web 6 Arctic 7 2LChat 0	Web: 25 Arctic: 29 2LChat: 14
Level 1	Web 2 Arctic 2 2LChat 2	Web 4 Arctic 0 2LChat 6	Web 1 Arctic 0 2LChat 6	Web 2 Arctic 2 2LChat 0	Web: 9 Arctic: 4 2LChat: 14
Total	Web: 10 Arctic: 10 2LChat: 2	Web: 8 Arctic: 18 2LChat: 14	Web: 8 Arctic: 7 2LChat: 13	Web: 18 Arctic: 11 2LChat: 1	120 per age group plus PU/FP/HUE, Practice, & Covariates

*Some CR items (Constructed Response) will measure *up through* the listed level (listed level is top score).

Figure 6. The number of data points from each scenario, and their planned allocation to the levels from each strand.

The Measurement Model: Algorithm Used for Learning Analytics. In this approach to data analytics, a statistical analytic technique called a “measurement model” serves as the algorithm to compile the results and make inferences about learners. Other fields such as computer science that come to learning analytics from a different historical basis often use different vocabulary to describe such algorithms. For instance, the Rasch model often used in educational assessment from a computer science perspective would be considered an algorithm employing a multilayer feed-forward network (Russell & 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 only the thresholds on each item node (item difficulty). The 2PL IRT model, by contrast, is an algorithm employing a multilayer feed-forward network with g as the 2PL function (also a sigmoidal curve-fitting function), in which both weights (item discrimination) and thresholds on each item node (item difficulty) are estimated. The 3PL model is an algorithm employing a multilayer feed-forward network with g as the 3PL function (sigmoidal), in which weights (item discrimination), thresholds on each item node (item difficulty), and a lower asymptote (guessing parameter) are estimated.

For this initial data collection, we chose to use the Random Coefficients Multinomial Logit Model (Adams & Wilson, 1996), which allows polytomous data (as is appropriate for the items in the tasks), and provides the Wright Maps mentioned above (and are shown in Figures 7, 8 and 10 below). Due to small sample size, we ended up with only two scored items in the SCN strand and therefore excluded this strand from the analyses. As a result, we report our findings from three separate unidimensional analyses with 27 items measuring the CiN strand, nine items measuring the PiN strand,

and 22 items measuring the ICN strand. The ConQuest 3.0 estimation software (Adams, Wu & Wilson, 2012) was employed throughout the analyses, which uses RCML as its basis for statistical modeling. See the paper by Adams et al. (1997) for the relevant equations and statistical estimation procedures for the analysis.

Results and Discussion: Coming to an Interpretation

Two of the three scenarios were selected for empirical study with students. The two, the science/math Arctic Trek collaboration contest and the Webspiration shared literature analysis task, were identified by participating countries as the most desirable to study at this time. The science/math and language arts activities were more aligned with the school systems in the countries, which rarely used anything like cross-country chat tools in the classroom, but sometimes did employ math simulations and online scientific documents as well as graphical and drawing tools for student use. In particular, the third task (the Second Language Chat) was described by participating countries, teachers, and schools as a forward-looking, intriguing scenario, but farther away on the adoption curve for school-based technology.

Not all of the planned automated scoring and data analysis for the items in the two piloted scenarios was available for application to this data set, as the total numbers of cases was too small for the empirically-based scoring to be successfully calibrated. This calibration will be completed at a later point, when larger data sets are available. Each of the two scenarios were presented in three forms, for 11-, 13- and 15 year-olds, respectively, with a subset of common items across the three forms.

For a sample of 103 students in our first field test (i.e., those for whom we were able to match their login ids for two scenarios), assessment results from the two scenarios within the overall ICT literacy domain were analyzed using a three separate *consecutive* unidimensional item response (RCML) models.

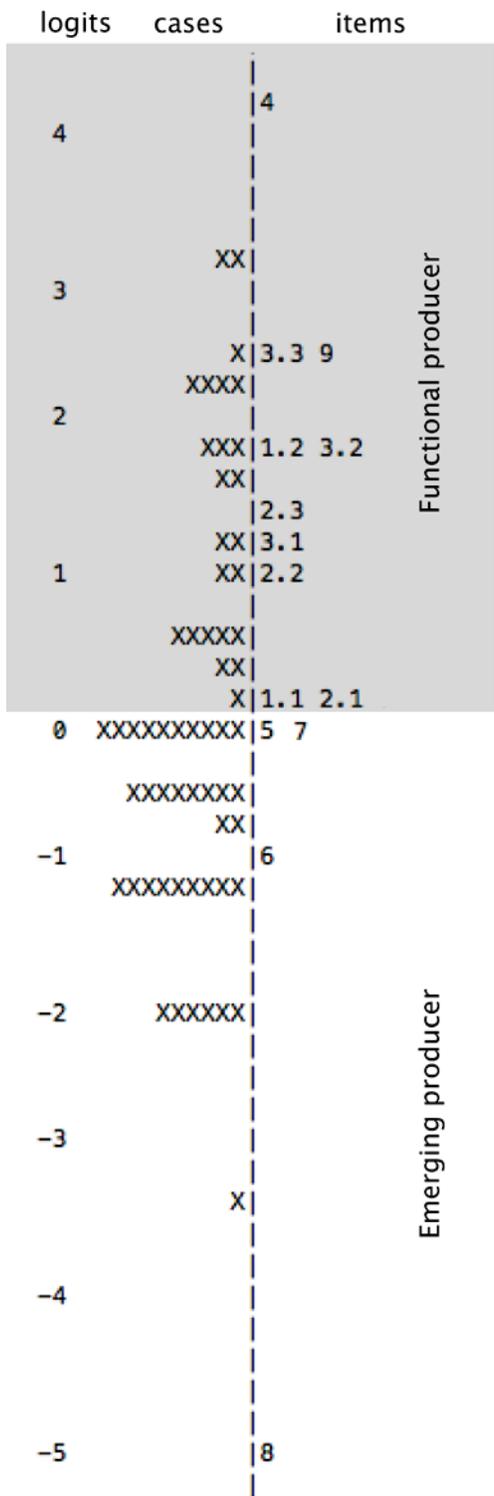
One way of checking if the assumptions and requirements of the confirmatory approach has been met within each of the three consecutive models is to examine the weighted mean square fit statistic estimated for each item within each model. Item fit can be seen as a measure of the discrepancy between the observed item characteristic curve and the theoretical item characteristic curve (Wu & Adams, 2010). ConQuest estimates the residual-based weighted fit statistics, also called infit; ideally, infit values are expected to be close to 1.0. Values less than 1.0 imply that the observed variance is less than the expected variance while values more than 1.0 imply that the observed variance is more than the expected variance. A common convention of 3/4 (0.75) and 4/3 (1.33) is used as acceptable lower and upper bounds (Adams & Khoo, 1996). After deleting items that fit poorly, we found that all of the remaining 58 items for the three consecutive models fell within this range.

Table 2 shows the variances and correlations of EAP scores obtained from the consecutive approach. The closest estimated relation is between the CiN and PiN dimensions, at a correlation of .69.

Table 2. Variances and correlations of EAP scores from three consecutive unidimensional models.

	CiN	PiN	ICN
PiN	0.65		
ICN	0.57	0.57	
variance	0.97 (0.14)	2.54 (0.35)	1.79 (0.25)

Producer
in Social Networks



Reliabilities for each of the three dimensions are .76, .79, .73, for CiN, PiN, and ICN, respectively—all of which are quite high, given the number of items available.

<Figure 7. Wright Map for the Consumer in Social Networks strand with bands representing each level.

Figure 7 shows one of the “Wright Maps” (Wilson, 2005) obtained from the unidimensional model for Consumer in Social Networks. Items are vertically ordered with respect to their difficulties and persons (cases) are vertically ordered with respect to their abilities. Each “X” on the left-hand side represents a small number of students, and the items are shown on the right-hand side using their item numbers. The locations are interpreted as follows, for dichotomous items:

- (a) when a student’s X matches an item location, the probability of that student succeeding on that item is expected to be 0.50;
- (b) when a student’s X is above the item, then the probability is above 0.5 (and vice-versa); and
- (c) these probabilities are governed by a logistic distribution (see Wilson, 2005 for a discussion of this).

Where the items are polytomous, the labeling is a little more complex, for example, in Figure 7, note that Item 5 is represented by two labels: 5.1 and 5.2. The former is used to indicate the threshold between category 0 and categories 1 and 2 (combined); the latter is used to represent the threshold between categories 0 and 1 (combined) and category 2. The interpretation of the probability is equivalent to that for a dichotomous item: that is, when a student’s X matches the 5.1 location, the probability of that student succeeding at levels 1 or 2 on that item is expected to be 0.50; and similarly, when a student’s X matches the 5.2 location, the probability of that student succeeding at only level 2 on that item is expected to be 0.50.

Just as schools have offered a more level playing field for the learning of traditional knowledge areas such as math and reading, so too direct intervention may be needed to bring all students to career and college readiness skills in their digital undertakings. Other 21st century skills should be similarly assessed. Given the importance to the future opportunities that students will have depending on the types of skills and abilities they can exhibit in these areas, the contributions they will need to make through the use of such skills cannot be underestimated (Binkley et al., 2012). Again, more examples that show the

possibility of more fine-grain interpretation at the individual level through such learning analytic approaches will be discussed in an upcoming paper.

The maps are also useful in investigating whether there is a good coverage of abilities by items. Ideally, if permitted a sufficient number of items for each strand, we would hope to see the range of item difficulties approximately match the range of person abilities. This would mean that there are items approximately matching every level of the person ability. This can be seen to be true for the Consumer in Social Networks strand. Figure 7 also shows “banding” for the Wright Map for the Consumer in Social Networks—that is, we have carried out a judgmental exercise to approximately locate where students in each level would be located. We note that, after misfit analysis, only level 3 of item 7 (i.e., 7.3) remains in the item pool to represent the highest level, the discriminating consumer. In contrast, the lower two levels, conscious consumer and emerging consumer, are well-represented. We can see that students in this sample have shown a range of abilities on this strand that spans all three of the hypothesized levels, from Emerging to Discriminating Consumer, although, as we might have expected, there are relatively more in the lower levels than in the higher levels (an observation that will be repeated in the other strands).

Interpretive Phase

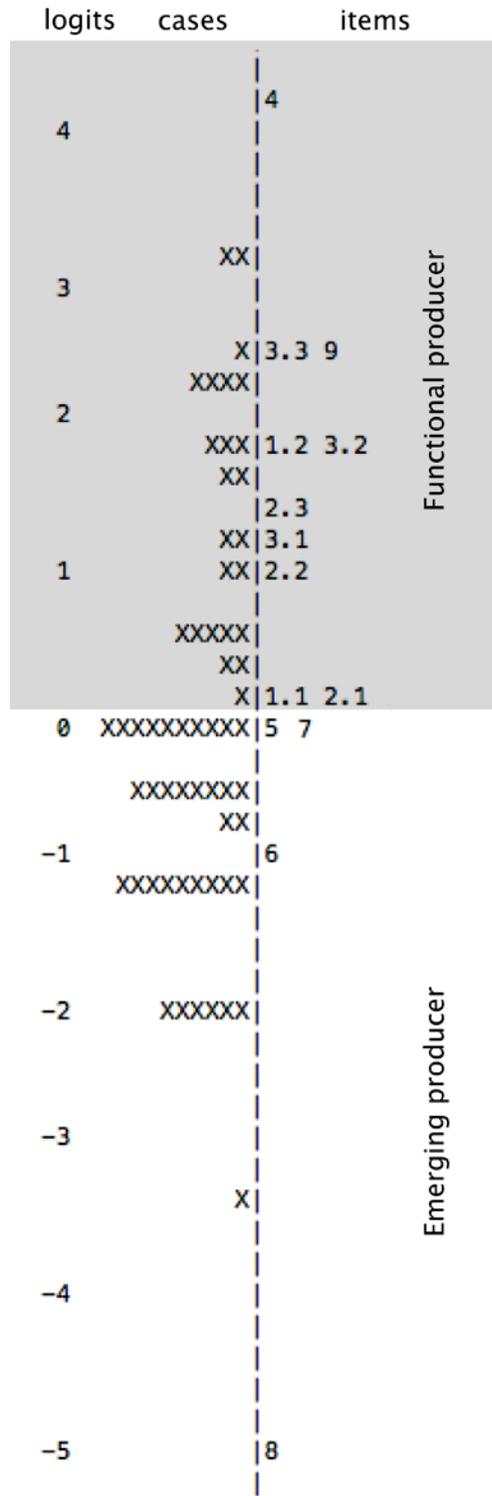
Once the coherence of the four building blocks is in place and results are obtained, interpretation of the learning analytics can take place. This interpretation is done through the maps yielded. For example, the distribution of students over the proposed construct for the Consumer strand can be seen in Figure 7. Most students fell into the Conscious consumer portion of the construct, as shown by the histogram composed of the “X” character on the left portion of the display. This location on the map indicated that the students tended to approach the ICT learning in social network activities with the ability to select appropriate tools and strategies (strategic competence), construct targeted searches, compile information systematically, and at least know that credibility of online information is an issue, though perhaps not how to address this. Some students, represented by the lower X’s in this band, represented some but not full proficiency in these areas while others, as they placed in locations up the histogram, could be expected to employ these behaviors effectively more often and more fully.

However, a few students remained in the Emerging consumer terrain. Here, students could perform basic searches for information and knew that some digital tools existed to do more. But they were often routinely unable to employ digital tools to perform more than the most basic tasks. Some students were challenged even with the notion of clicking a link to navigate to another page, where the information they sought could be found. Others exhibited only the concept of linear reading, where they would attempt to read through each web page they reached from top to bottom. Without the concept of scanning or searching for information within a page, they were unable to progress very far with the activities, or contribute effectively to collaborations.

Also, students in the Emerging terrain often exhibited no notion of questioning sources or considering the credibility of information. These students would need substantial support with their digital literacy skills to take part in virtual or distributed collaborative learning.

By contrast, a few students showed behaviors across settings that would propel them into the upper category of Discriminating consumer proficiencies. These students were showing leadership ability with their ICT collaboration skills. They could not only question but also judge credibility of sources. They often were able to integrate information into larger views of coherent knowledge, construct searches suited to personal circumstances, and then filter, manage, and organize information much more effectively than other students. Because they had a tendency to seek expert knowledge and select optimal tools, their collaboration skills and their teams benefitted. Groups in which they engaged might take leaps forward based on such participants. A notion of how to combine teams so that they represented a range of support and incorporated peer learning would probably be helpful to support the wide range of skills.

Producer
in Social Networks



A more fine-grained interpretation at the level of each student can also take place using these maps. This is needed for clear decision-making about current learning outcomes and future needs if intervention will take place at the student level. For this purpose, a student at any location on the map can be selected. Based on their map location, inferences can be made about what they have mastered in this part of the complex terrain of 21st century skills. Such examples at the individual level will be discussed in an upcoming paper.

For the Producer strand of the construct, the group of students surveyed tended to be more evenly split between the Emerging and Functional producer traits. This positioning? indicates that though living in the era of Web 3.0, when students may be immersed in a Maker culture, some students have considerably more skills toward producing digitally and collaboratively online than do other students. About half the students could produce simple representations from templates or start an online identity, but were challenged to go much beyond this to participate in an effective collaboration, problem solve, or produce a product. By contrast, the other approximately half of students were able to go far beyond this. They showed the ability to organize communication effectively in social networks, use collaborative networked tools to develop creative and expressive artifacts, plan for shared solutions, and also demonstrate? an awareness of safety and security online in the process of their production activities.

So the message here, should it be verified by larger data sets, is a strong word of caution to programs, schools or instructors. Educators may presume every K-12 student today is a digital native. It is true that the digital divide may have eased in many locations and that students in a number of primary and secondary classrooms today may have grown up surrounded by many new technologies. However, the matter of access is not the sole determining factor in student development of 21st century skills and abilities (Binkley et al., 2012).

<Figure 8. Wright Map for the Producer in Social Networks strand with bands representing each level.

Figure 8 shows the banded Wright Map for the Producer in Social Networks strand. All of the items in this strand ended up in their hypothesized band. Items in the Functional Producer band require students to develop creative, expressive and/or useful content artifacts and tools. For instance, item 2 (shown in Figure 9 below) and item 3 require students to create a graph representing data. In item 1, we asked students to summarize the data in a table. We hypothesized that all of these three items would end up in the upper

band of the Wright Map and our hypothesis was confirmed. The two most difficult items in this band, as we hypothesized, are items 4 and 9. In both of these items we asked students to upload files to the webpage. Items in the Emerging Producer band require students to post an artifact and/or perform basic production tasks. In items 5, 6 and 7, for instance, we asked students to copy/paste the readily available content. In item 8, we asked students to connect at least one node in the concept map. All of these items ended up in the lower part of the Wright Map, as we hypothesized before the test.

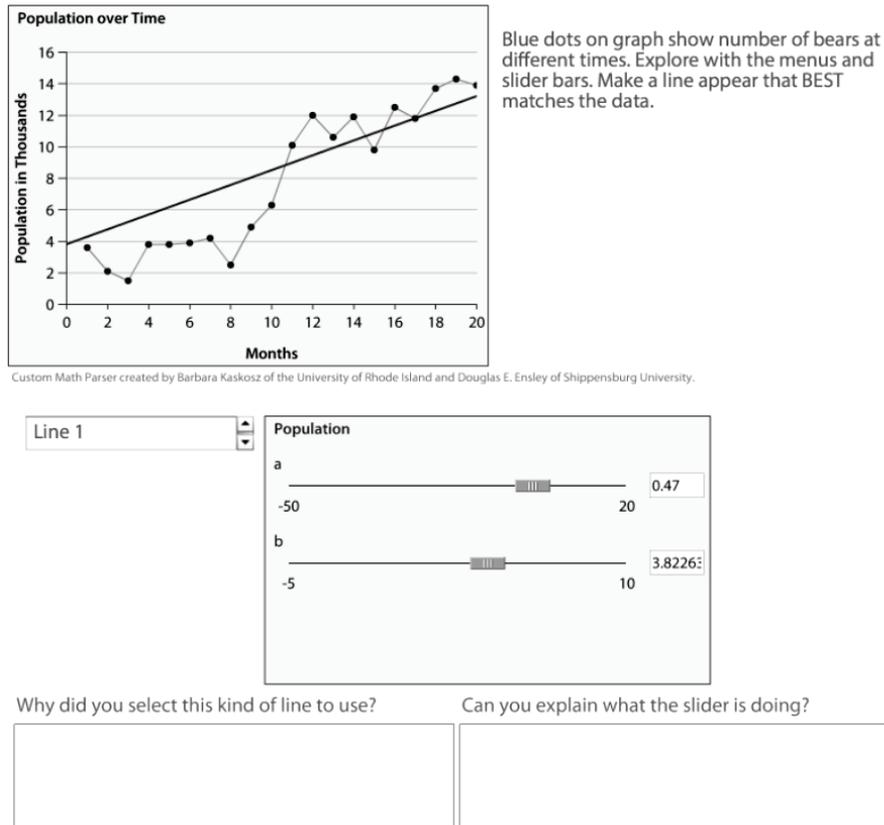


Figure 9. Task for item 2

We do not display a banded Wright Map for the Developer of Social Capital strand, as there are only two items remaining in this strand, both of which are aimed at measuring at third level (Proficient connector) of the construct. Finally, we do not display and discuss the Wright Map results for the Participant in the Intellectual Capital strand, as it will appear in a future paper.

Conclusion

In conclusion, this paper suggests a modified definition for learning analytics. The definition underscores the need for meaningful interpretation of learning results and also incorporates key ways that learning analytics are employed with complex data and algorithms.

Measuring intricate constructs through digital analytics, as displayed here, can help us to appreciate the new ways that students will be required to think and work compared to previous decades. This paper describes a domain modeling process for assessing ICT literacy through the BEAR

Assessment System, along with examples of task development, and results from implementation of a first field test in four countries.

This initial field test has demonstrated through the learning analytics approach that, for this data set, students in the 11-15-year age group show widely differing knowledge and skills in these areas. Some students are only just beginning to take initial steps toward digital competence while other students show levels of mastery that would likely challenge skilled adult samples, such as collaborating easily to create professional-level commentaries or reports. Doubtless the wide range is due at least in part to the absence of formal teaching and opportunities to learn these skills in many schools, which results in an increasing gap between the haves and the have-nots. It looks as though this is particularly true in the domain of ICT literacy, which we take to be a gatekeeper for other skills and achievements in the future.

While these techniques and approaches are illustrated through one example project, they potentially are broadly applicable to many learning domains. The takeaway lesson may be that effectively educating students for 21st century skills may also require offering them 21st century supports for learning—including improved approaches to examining patterns and evidence of learning in complex settings.

References

- ACARA. (2012). Australian Curriculum, Assessment and Reporting Authority, National Assessment Program, ICT Literacy, Years 6 and 10 Report 2011.
- Adams R. J., & Khoo, S-T. (1996). *Quest* [computer program]. Melbourne, Australia: ACER Press.
- Adams, R.J., & Wilson, M. (1996). Formulating the Rasch model as a mixed coefficients multinomial logit. In, G. Engelhard & M. Wilson (Eds.), *Objective measurement III: Theory into practice*. Norwood, NJ: Ablex.
- Adams, R.J., Wu, M., & Wilson, M. (2012). *ConQuest 3.0* [computer program]. Hawthorn, Australia: ACER.
- Ainley, J., Fraillon, J., Schulz, W., & Gebhardt, E. (2014). *Measuring Changes in ICT Literacy over Time*. Paper presented at the National Council on Measurement in Education, Philadelphia, PA.
- Baker, R. S., & Siemens, G. (2014). Educational data mining and learning analytics. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 253-274).
- Binkley, M., Erstad, O, Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First century skills. In Griffin, P, McGaw B. & Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Dordrecht, Germany: Springer.
- Brady, A., Conlan, O., Wade, V., & Dagger, D. (2006). *Supporting Users in Creating Pedagogically Sound Personalised Learning Objects*. Paper presented at the Adaptive Hypermedia and Adaptive Web-based Systems, Dublin, Ireland.
- Briggs, D. C., & Wilson, M. (2003). An introduction to multidimensional measurement using Rasch models. *Journal of Applied Measurement*, 4(1), 87–100.

- Chedrawy, Z., & Abidi, S. S. R. (2006). *An Adaptive Personalized Recommendation Strategy Featuring Context Sensitive Content Adaptation*. Paper presented at the Adaptive Hypermedia and Adaptive Web-Based Systems, 4th International Conference, AH 2006, Dublin, Ireland.
- Chi, E. H., Pirolli, P., Suh, B., Kittur, A., Pendleton, B., & Mytkowicz, T. (2008). *Augmented Social Cognition*. Palo Alto, CA: Palo Alto Research Center.
- Cisco, Intel, Microsoft (CIM). (2008). *Transforming Education: Assessing and Teaching 21st Century Skills*. Authors. Retrieved from <http://atc21s.org/wp-content/uploads/2011/04/Cisco-Intel-Microsoft-Assessment-Call-to-Action.pdf>
- Cooke, N. J., Gorman, J. C., & Winner, J. L. (2007). Team Cognition. In F. T. Durso (Ed.), *Handbook of Applied Cognition* (pp. 239-268). New York: John Wiley & Sons, Ltd.
- Dagger, D., Wade, V., & Conlan, O. (2005). Personalisation for All: Making Adaptive Course Composition Easy. *Educational Technology & Society*, 8(3), 9-25.
- Dhar, V., & Olson, M. H. (1989). Assumptions Underlying Systems that Support Work Group Collaboration. In M. H. Olson (Ed.), *Technology Support for Work Group Collaboration*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Networks*, 16(3), 87-100.
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5/6), 301-317. doi: 10.1504/IJTEL.2012.051816
- Ferrari, A. (2012). *Digital competence in practice: An analysis of frameworks*. Seville, Spain: Institute for Prospective Technological Studies, European Commission.
- Fraillon, J. (2014). *Assessing Computer and Information Literacy across 20 Countries: The International Computer and Information Literacy (ICILS)*. Paper presented at the National Council on Measurement in Education (NCME), Philadelphia, PA.
- Friedman, T. L. (2007). *The World is Flat 3.0: A Brief History of the Twenty-first Century*. New York: Picador.
- Griffin, P., McGaw B. & Care, E. (Eds.) (2012). *Assessment and Teaching of 21st Century Skills*. Dordrecht, Germany: Springer.
- Griffiths, T. L., & Steyvers, M. (2007). Topics in Semantic Representation. *Psychological Review*, 114(2), 211-244.
- Kennedy, C. A., & Draney, K. (2006). Interpreting and using multidimensional performance data to improve learning. In X. Liu (Ed.), *Applications of Rasch Measurement to Science Education*. Chicago: JAM Press.
- Kozma, R. (Ed.). (2003). *Technology, innovation, and educational change. A global perspective. A report of the Second Information Technology in Education Study, Module 2*. Amsterdam: International Association for the Evaluation of Educational Achievement.

- Kozma, R. B. (2003). Technology and classroom practices: An international study. *Journal of Research on Technology in Education*, 36(1), 1-5.
- Lampe, C., Resnick, P., Forte, A., Yardi, S., Rotman, D., Marshall, T. & Lutters, W. (2010). Educational priorities for technology-mediated social participation. *IEEE Computer*, 43(11), 60-67.
- Lehrer, R., Kim, M-J., Ayers, E., & Wilson, M. (2014). Toward establishing a learning progression to support the development of statistical reasoning. In J. Confrey and A. Maloney (Eds.), *Learning over Time: Learning Trajectories in Mathematics Education*. Charlotte, NC: Information Age Publishers.
- Masters, G. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174.
- National Research Council (NRC). (2001). *Knowing what students know: The science and design of educational assessment*. Committee on the Foundations of Assessment. J. Pellegrino, N. Chudowsky, & R. Glaser (Eds). Washington, DC: National Academy Press.
- 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.
- Park, J. (2011). ICT in Education in Asia-Pacific. *Asia-Pacific Regional Forum on ICT Applications*. Retrieved from http://www.itu.int/ITU-D/asp/CMS/Events/2011/ict-apps/s5_JPark.pdf
- Pedersen, S. G., Malmberg, P., Christensen, A. J., Pedersen, M., Nipper, S., Græm, C. D., & Norrgård, J. (Eds.). (2006). *E-learning Nordic 2006: The impact of ICT on education*. Copenhagen, Denmark: Ramboll Management.
- Pirolli, P. (2007). Cognitive Models of Human-Information Interaction. In F. T. Durso (Ed.), *Handbook of Applied Cognition* (pp. 443-470). New York: John Wiley & Sons, Ltd.
- Pirolli, P. (2009). An elementary social information foraging model. *CHI 2009, ACM Conference on Human Factors in Computing Systems*, (pp. 605-614). Boston: Association for Computing Machinery.
- Pirolli, P., Preece, J., & Shneiderman, B. (2010). Cyberinfrastructure for social action on national priorities. *IEEE Computer*, 43(11), 20-21.
- Pirolli, P., & Wilson, M. (1998). A Theory of the Measurement of Knowledge Content, Access, and Learning. *Psychological Review*, 105(1), 58-82.
- Qualifications and Curriculum Authority. (2007). *About information and communication technology: Assessment guidance* Retrieved from <http://web.archive.org/web/20040815054520/http://www.qca.org.uk/7889.html>
- Rogers, W. A., Pak, R., & Fisk, A. D. (2007). Applied Cognitive Psychology in the Context of Everyday Living. In F. T. Durso (Ed.), *Handbook of Applied Cognition* (pp. 3-27). New York: John Wiley & Sons, Ltd.
- Russell, S., & Norvig, P. (2009). *Artificial Intelligence, A Modern Approach* (3rd Ed.). Upper Saddle River: Prentice Hall.

- Scalise, K. (2012). Using Technology to Assess Hard-to-Measure Constructs in the CCSS and to Expand Accessibility. *Invitational Research Symposium on Technology Enhanced Assessments, Educational Testing Service (ETS), Washington, DC*. Retrieved from http://www.k12center.org/events/research_meetings/tea.html
- Scalise, K., & Gifford, B. R. (2006). Computer-Based Assessment in E-Learning: A Framework for Constructing "Intermediate Constraint" Questions and Tasks for Technology Platforms. *Journal of Teaching, Learning and Assessment*, 4(6).
- Sclater, N. (2014). JISC: Code of practice for learning analytics: A literature review of the ethical and legal issues. Retrieved from http://repository.jisc.ac.uk/5661/1/Learning_Analytics_A-Literature_Review.pdf
- Stanton, J. M. (2012). *An Introduction to Data Science*. Retrieved from <http://surface.syr.edu/istpub/165/>
- U.S. Department of Education. (2010). *Transforming American Education: Learning Powered by Technology. National Education Technology Plan 2010*. Retrieved from <http://www.ed.gov/sites/default/files/netp2010.pdf>
- Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Erlbaum.
- Wilson, (2009a, December). *Assessment for Learning and for Accountability*. Paper presented at the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda, at ETS, Princeton, NJ. Retrieved from
- Wilson, (2009b, December). *Assessment for Learning and for Accountability*. Policy Brief from the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda, at ETS, Princeton, NJ. Retrieved from
- Wilson, M. (2012). Responding to a challenge that learning progressions pose to measurement practice: hypothesized links between dimensions of the outcome progression. In A. C. Alonzo & A. W. Gotwals, (Eds.), *Learning Progressions in Science*. Rotterdam, The Netherlands: Sense Publishers.
- Wilson, M., Bejar, I., Scalise, K., Templin, J., Wiliam, D., & Torres Iribarra, D. (2010). Assessment and Teaching of 21st Century Skills: Perspectives on Methodological Issues *Learning and Technology World Forum 2010*. White Paper presented at the Learning and Technology World Forum 2010, London.
- Wilson, M., Bejar, I., Scalise, K., Templin, J., Wiliam, D., & Torres-Iribarra, D. (2012). Perspectives on methodological issues. In Griffin, P., McGaw B. & Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills*. Dordrecht, Germany: Springer.
- Wilson, M. & Carstensen, C. (2007). Assessment to improve learning in mathematics: The BEAR Assessment system. In A. Schoenfeld (Ed.), *Assessing mathematical proficiency* (pp. 311-332). London: Cambridge University Press.
- Wilson, M., & Hoskens, M. (2005). Multidimensional item responses: Multimethod/multi-trait perspectives. In S. Alagumalai, D. D. Curtis, & N. Hungi (Eds.) *Applied Rasch measurement: A*

Book of Exemplars (Papers in honour of John Keeves) (pp. 287-308). Dordrecht, The Netherlands: Springer-Kluwer Academic Publishers.

Wilson, M., & Scalise, K. (2015). Assessment of Learning in Digital Networks. In Griffin, P., & Care, E. (Eds.), *Assessment and Teaching of 21st Century Skills Volume 2 - Methods & Approaches*. Dordrecht: Springer.

Wilson, M., Scalise, K., & Gochyyev, P. (2015). Rethinking ICT Literacy: From Computer Skills to Social Network Settings. *Thinking Skills and Creativity*, 18, 65-80.

Wilson, M. & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13(2), 181-208. Retrieved from <http://www.informaworld.com/smpp/content~content=a783685281~db=all>