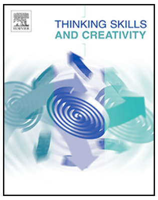
Thinking Skills and Creativity 18 (2015) 81–90

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: http://www.elsevier.com/locate/tsc



Using educational data mining to assess students' skills at designing and conducting experiments within a complex systems microworld

Janice D. Gobert[a,*], Yoon Jeon Kim[a,1], Michael A. Sao Pedro[a,1], Michael Kennedy[b,2], Cameron G. Betts[c,3]

[a] *Social Sciences & Policy Studies Department, WPI, 100 Institute Rd., Worcester, MA 01609, USA*
[b] *WPI, 100 Institute Rd., Worcester, MA 01609, USA*
[c] *Apprendis, LLC, USA*

ARTICLE INFO

Article history:
Received 4 October 2014
Received in revised form 25 April 2015
Accepted 30 April 2015
Available online 2 May 2015

Keywords:
Complex systems
Inquiry assessment
Performance assessment
Educational data mining
21st century skills

ABSTRACT

Many national policy documents underscore the importance of 21st century skills, including critical thinking. In parallel, recent American frameworks for K-12 science education call for the development of critical thinking skills in science, also referred to as science inquiry skills/practices. Assessment of these skills is necessary, as indicated in policy documents; however, this has posed a great challenge for assessment researchers. Recently, some science learning environments seek to assess these science skills. These systems log all students' interactions within the given system, and if fully leveraged, these logs provide rich assessments of inquiry skills. Here, we describe our environment Inq-ITS (inquiry intelligent tutoring system), that uses educational data mining to assess science inquiry skills, as described as 21st century skills. Additionally, here, we describe how we measure students' skills at designing controlled experiments, a lynchpin skill of inquiry, in the context of complex systems. In doing so, our work addresses 21st century skill assessment in two ways, namely of inquiry (designing and conducting experiments), and in the context of complex systems, a key topic area of 21st century skills. We use educational data mining to develop our assessment of this skill for complex systems.

1. Introduction

1.1. Background

Following the launching of Sputnik in October of 1957, policy makers in the United States began to question the quality of science instruction in schools, which, in turn, instantiated a call for change in all science curricula. Post-Sputnik, educators

---

[*] Corresponding author. Tel.: +1 508 831 5619; fax.: +1 508 831 5296.
E-mail addresses: jgobert@wpi.edu (J.D. Gobert), yjkim@wpi.edu (Y.J. Kim), mikesp@wpi.edu (M.A. Sao Pedro), kennedym3@gmail.com (M. Kennedy), cam@apprendis.com (C.G. Betts).
[1] Tel.: +1 508 831 6618.
[2] Tel.: +1 774 249 5327.
[3] Tel.: +1 617 538 6868.

http://dx.doi.org/10.1016/j.tsc.2015.04.008

and policy makers sought that science literacy should include science content knowledge, inquiry skills, and understanding of the nature of science (Perkins, 1986). Secondly, post-Sputnik reform efforts also called for educating the broad populace rather than the top 10% of high achieving students. Taken together, the goal was and continues to be to develop a citizenry with knowledge and skills so that they can participate fully in a democracy (Stokes, 1997).

In more recent reports, policy makers continue to emphasize the need for 21st century skills (National Research Council, 2010; Partnership for 21st Century Skills, 2007). In brief, 21st century skills broadly include: *cognitive knowledge/skills* (e.g., critical thinking), *interpersonal skills* (e.g., communication and teamwork skills), and *intrapersonal skills* (e.g., metacognitive/motivational, self-regulated learning (Partnership for 21st Century Skills, 2007). Twenty-first century skills predict both college grades and future employment success, and as technological advancements continue, people will be increasingly expected to think in creative and divergent ways (Lai & Vierling, 2012). Lastly, 21st century skills are acknowledged as important for developing innovative thinkers (Sawyer, 2006; Sternberg, 2006; Sternberg & Lubart, 1991, 1995), necessary for a knowledge-based economy Bereiter, 2002; Resnick, 2008).

In the present work, we focus on the *cognitive components* of 21st century skills, which include: critical thinking, non-routine problem-solving, and systems-thinking. Specifically, here we assess inquiry skills, critical thinking in science, in the context of complex systems (cf.,Hmelo-Silver & Azevedo, 2006; Jacobson et al., 2006; Yoon, 2008). In other work, we address intrapersonal skills, namely engagement (Gobert, Baker, & Wixon, 2015).

1.2. Traditional educational assessments

The purpose of educational assessments, broadly described, is to make inferences about students' knowledge and skills. Traditionally, as in the case of science, formal assessment is done on the basis of standardized tests, which use multiple-choice items to determine the level of proficiency a student has achieved. Items are developed using standards, for example, state content standards; these tests are criterion-referenced in that they are intended to measure students in terms of their level of mastery on grade-appropriate knowledge and skills. These tests are also norm-referenced in that they compare students relative to their peers. These tests are typically implemented using paper and pencil format and multiple-choice items (Anastasi & Urbina, 2009).

However, given the richness of critical thinking involved in science inquiry, it has been acknowledged that typical science achievement tests do not adequately reflect the complex science knowledge and inquiry process skills that are important components of scientific literacy or of 21st century skills (Clarke-Midura, Dede, & Norton, 2011; Haertel, Lash, Javitz, & Quellmalz, 2006; Leighton & Gierl, 2011; National Committee on Science Education Standards and Assessment, 1996; Quellmalz & Haertel, 2004; Quellmalz, Kreikmeier, DeBarger, & Haertel, 2007). As discussed elsewhere (Gobert, Sao Pedro, Raziuddin, & Baker, 2013), the limitations of these tests are partly due tothe simplified conceptions of the nature of science understanding at the time that the tests were designed (DiCerbo & Behrens, 2012; Mislevy, Behrens, Dicerbo, & Levy, 2012). Thus, more recently, it has been widely acknowledged that multiple choice items are not suitable means to assess rich inquiry skills, and instead, tasks need to be designed to elicit data that can address what students know and how they use their knowledge, rather than elicit data that we can easily collect and analyze (Pellegrino, 2009). In doing so, one can assess both the products and processes of inquiry (Rupp, Gushta, Mislevy, & Shaffer, 2010).

In short, the problem becomes: how do we use policy documents about critical thinking in science (National Research Council, 2010; Partnership for 21st Century Skills, 2007) use to inform the design and development of valid, reliable assessments of rich inquiry skills? (Leighton & Gierl, 2011). Furthermore, specific to this paper, we address how to do this type of assessment in the context of complex systems, a key topic area of 21st century thinking.

1.3. Inq-ITS (inquiry intelligent tutoring system)

Our design work started with the specifications for what knowledge and skills students should possess (NGAA, 2013) in order to develop a system that could provide fine-grained assessment data on students' science inquiry skills. Our environment, Inq-ITS (http://slinq.org) is a rigorous, technology-based learning environment that assesses and scaffolds middle school students in earth, life, and physical science during learning. Our work recognizes that these environments can provide a more fertile basis upon which to develop performance-based assessments by leveraging computational techniques to analyze students' log files of their inquiry processes (Gobert, Sao Pedro, Baker, Toto, & Montalvo, 2012; Gobert et al., 2013).

Inq-ITS uses microworlds (Papert, 1980) to engage students in inquiry. Microworlds are computerized representations of real-world phenomena whose properties can be inspected and changed (Pea & Kurland, 1984; Resnick, 1997). Since microworlds share many features with real apparati (Gobert, 2005; in press), they provide greater authenticity for "doing science". In turn, microworlds afford authentic performance assessment of inquiry skills because with a microworld in Inq-ITS, students can generate a hypothesis, test it, interpret data, warrant their claims with data, and communicate findings with regard to what they discover. These inquiry tasks reflect the national frameworks for inquiry (National Committee on Science Education Standards and Assessment, 1996; National Research Council, 2010), and represent the critical thinking skills used to reason logically about scientific concepts as reflected in 21st century skills documents (Partnership for 21st Century Skills, 2007).

In terms of assessment techniques, we employ techniques that originate from educational data mining (EDM henceforth; cf., Baker & Yacef, 2009; Romero & Ventura, 2010), which grew from computer science, human-computer interaction, and