Exploring Opportunities for Caring Assessments

Diego Zapata-Rivera [1] and Julita Vassileva [2]

[1] Educational Testing Service, [2] University of Saskatchewan

Preface

The notion of intelligent systems that "care" about students is at the center of ITS research [1-3]. A variety of adaptive learning systems that "care" have been developed in the past [4, 5]. These systems make use of student/user models to adapt their interactions to a particular student (e.g., amount and type of feedback, content sequencing, scaffolding, and access to visualization tools and other materials). Student model variables include cognitive abilities, metacognitive skills, affective states, and other variables such as personality traits, learner styles, social skills, and perceptual skills [5].

Caring assessment systems are defined as systems that provide students with a positive assessment experience while improving the quality of evidence collected about the student's knowledge, skills and abilities (KSAs) [6]. Taking a test is typically a stressful situation, and many people underperform due the stress. Caring assessment systems take into account assessment information from both traditional and non-traditional sources (e.g., student emotions, prior knowledge, and opportunities to learn) to create situations that students find engaging, and to collect valid and reliable evidence of students' KSAs.

Taking a test is not just a passive mechanism for assessing how much people know. It actually helps people learn, and it works better than a number of other studying techniques [7]. Caring formative assessment can be done by a computer system or by peer-learners. Learners testing each other in a friendly, collegial and constructive way, can be an engaging and effective form of collaborative learning and preparation for assessment that also helps establish peer-mentorship relationships among learners. Developing systems or approaches (e.g. games) that support learners test each other in a friendly, collegial and constructive way, is a new and promising direction of research.

This workshop provides a great opportunity for ITS and assessment researchers to share information about the potential of applying ITS techniques and approaches in the development of a new generation of caring assessments. Examples of ITS technologies that have been successfully used for assessment purposes include automatic scoring of essays and short responses [8]. The use of dialogue systems for assessment is being explored [9, 10]. This workshop is a timely and relevant event for the ITS and assessment communities. New assessments for skills such as problem-solving, collaboration, and scientific inquiry include the use of highly interactive simulations and collaboration with artificial agents. Advances in ITSs will play an important role in the development of the next generation of assessment systems.

Eight recognized members of the research community were invited to serve as members of the program committee. Each member reviewed up to two submissions. The program committee members are: Ivon Arroyo, *Worcester Polytechnic Institute*; Ricardo Conejo, *University of Malaga*; Vania Dimitrova, *University of Leeds*; Sidney D'Mello, *University of Colorado Boulder*; Art Graesser, *University of Memphis*; G. Tanner Jackson, *Educational Testing Service*; Irvin R. Katz, *Educational Testing Service*; and Steve Ritter, *Carnegie Learning*.

Seven papers were submitted and all of them were accepted for presentation at the workshop. Each paper received feedback from at least two reviewers. The accepted papers include: *When Should an Adaptive Assessment Care?* (Blair Lehman, Jesse R. Sparks, and Diego Zapata-Rivera); *Incorporating Emotional Intelligence into*

Assessment Systems (Han-Hui Por and Aoife Cahill); *Diagnostic Assessment of Adults' Reading Deficiencies in an Intelligent Tutoring System* (Genghu Shi, Anne M. Lippert, Andrew J. Hampton, Su Chen, Ying Fang, and Arthur C. Graesser); *Tower of Questions: Gamified Testing to Engage Students in Peer Evaluation* (Nafisul Islam Kiron and Julita Vassileva); *Exploring Gritty Students' Behavior in an Intelligent Tutoring System* (Erik Erickson, Ivon Arroyo, Beverly Woolf), *Disengagement Detection Within an Intelligent Tutoring System* (Su Chen, Anne Lippert, Genghu Shi,Ying Fang, and Arthur C. Graesser); and *Assessments That Care About Student Learning* (Stephen E. Fancsali and Steven Ritter).

These papers offer different perspectives and current research toward the goal of making "caring" assessments part of the educational milieu.

The workshop included a thought-provoking discussion section that covered topics such as:

- The need for educating the public on the characteristics of different types of assessments and their appropriate use.
- Alternate criteria for adaptive testing that not only take into account the difficulty and the sequencing of questions but also other aspects of the student and the learning context and way of interaction.
- Assessments that provide additional feedback/guidance on content related issues and testing strategies (e.g., time management warnings).
- Using student model information from formative learning environments to inform the assessment systems.
- Possible approaches for integrating emotion data into assessment.
- Strategies for engaging students in peer assessment gaming activities.
- Exploring connections with other research areas (e.g., persuasive technologies).
- Evaluating the effects of additional features on test reliability, validity, and fairness.

We thank the authors for submitting relevant papers to the topic of the workshop, the program committee members for their time reviewing and providing constructive feedback to the authors, and the ITS workshop organizers, Nathalie Guin and Amruth Kumar, for providing us with this great opportunity to convene and address this topic.

Best regards,

Diego Zapata-Rivera and Julita Vassileva

References

1. Self, J.A. 1999. The distinctive characteristics of intelligent tutoring systems research: ITSs care, precisely, *International Journal of Artificial Intelligence in Education*, 10, 350–364
2. DuBoulay, B., Avramides, K., Luckin, R., Martinuz-Miron, E., Rebolledo Mendez, G., & Carr, A. 2010. Towards systems that care: a conceptual framework based on Motivation, Metacognition and Affect. *International Journal of Artificial Intelligence in Education*, 20, 197–229

3. Kay, J., & McCalla, G. 2003. The careful double vision of self. *International Journal of Artificial Intelligence and Education*, 13, 1–18
4. Brusilovsky, P., & Milan, E. 2007. User models for adaptive hypermedia and adaptive educational systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web. Methods and strategies of web personalization*. LNCS 4321, Berlin Heidelberg: Springer-Verlag. 3–53
5. Shute, V. J., & Zapata-Rivera, D. 2012. Adaptive educational systems. In P. Durlach (Ed.), *Adaptive technologies for training and education*. New York, NY: Cambridge University Press. 7–27
6. Zapata-Rivera, D. 2017. Toward Caring Assessment Systems. *In Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP '17)*, ACM, New York, NY, USA, 97–100. DOI: https://doi.org/10.1145/3099023.3099106
7. Karpicke, J., & Blunt, J. R. 2011. Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. Science 20 Jan 2011: 1199327 DOI: 10.1126/science.1199327
   http://science.sciencemag.org/content/early/2011/01/19/science.1199327
8. Shermis, M.D., & Burstein, J. 2013. *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. Routledge Chapman & Hall.
9. Zapata-Rivera, D., Jackson, T., Liu, L., Bertling, M., Vezzu, M., and Katz, I. R. 2014. Science Inquiry Skills using Trialogues. *12th International conference on Intelligence Tutoring Systems*. 625–626.
10. Graesser, A.C., Dowell, N., & Clewley, D. 2017. Assessing Collaborative Problem Solving Through Conversational Agents. In: von Davier A., Zhu M., Kyllonen P. (eds) *Innovative Assessment of Collaboration. Methodology of Educational Measurement and Assessment*. Springer, Cham. 65–80