Research Article

Zoo U: A Stealth Approach to Social Skills Assessment in Schools

Melissa E. DeRosier,[1] Ashley B. Craig,[2] and Rebecca P. Sanchez[3]

[1] *3-C Institute for Social Development, 1901 North Harrison Avenue, Suite 200, Cary, NC 27513, USA*
[2] *Games Research Department, 3-C Institute for Social Development, 1901 North Harrison Avenue, Suite 200, Cary, NC 27513, USA*
[3] *Games Research Department, Center for Research in Emotional and Social Health, 1901 North Harrison Avenue, Suite 200, Cary, NC 27513, USA*

Correspondence should be addressed to Rebecca P. Sanchez, sanchez@cresh.org

This paper describes the design and evaluation of *Zoo U*, a novel computer game to assess children's social skills development. *Zoo U* is an innovative product that combines theory-driven content and customized game mechanics. The game-like play creates the opportunity for *stealth assessment*, in which dynamic evidence of social skills is collected in real time and players' choices during gameplay provide the needed data. To ensure the development of an engaging and valid game, we utilized an iterative data-driven validation process in which the game was created, tested, revised based on student performance and feedback, and retested until game play was statistically matched to independent ratings of social skills. We first investigated whether the data collected through extensive logging of student actions provided information that could be used to improve the assessment. We found that detailed game logs of socially relevant player behavior combined with external measures of player social skills provided an efficient vector to incrementally improve the accuracy of the embedded assessments. Next, we investigated whether the game performance correlated with teachers' assessments of students' social skills competencies. An evaluation of the final game showed (a) significant correlations between in-game social skills assessments and independently obtained standard psychological assessments of the same students and (b) high levels of engagement and likeability for students. These findings support the use of the interactive and engaging computer game format for the stealth assessment of children's social skills. The created innovative design methodologies should prove useful in the design and improvement of computer games in education.

1. Introduction

Social skills comprise a group of behaviors and knowledge that help children create and maintain friendships and navigate a multitude of situations involving other people. A child's social skills competence can strongly influence his or her sense of well being and adjustment [1, 2]. The importance of social skills and peer relationships increases through the elementary school years and into adolescence, with peers becoming key providers of support, advice, companionship, and affirmation [3]. Children with strong social skills and the relationships built on these skills tend to have more positive emotional, behavioral, and academic functioning. Positive peer relationships also function as a protective factor against negative outcomes in the face of stressful life events such as poverty [2, 4, 5]. In contrast, a lack of social skills competence can increase children's risk for poor adjustment across many areas of functioning. Children who experience problems in social interactions with their peers are more likely to exhibit depression [6], anxiety disorders [7], suicide [8], delinquency and antisocial behavior [9, 10], substance abuse [11, 12], educational underachievement [13, 14], and other mental health difficulties [2]. Children's risk for negative outcomes increases as peer problems become more chronic or severe [1].

A key to preventing the development of more serious maladjustment is intervening with problematic social behaviors before they become chronic and intractable [2, 15–17]. The first step in intervention is identifying children in need of social skills help with an effective assessment.

Recognition of these needs has led to inclusion of social goals in many Individualized Education Plans (IEPs), Student Support Team strategies, and overall school improvement plans. To meet these goals, schools must have access to assessment tools that are valid, easy to use, and able to identify students struggling with specific social skills as well as track students' response to intervention (RTI).

The gold standards of social skills assessments are naturalistic behavioral observation and behavior rating scales [18]. These methods, however, require extensive time and cost commitment and often have psychometric challenges such as unreliable or biased observers, lack of social comparison data, situational specificity, observer reactivity (i.e., students modify their behavior because they are being observed), and inappropriate recording techniques [19, 20].

A desire for alternative tools that provide rich accurate data and increase student engagement has led to a transdisciplinary interest in computer games for assessment and learning [21]. Computer games provide a promising avenue for the assessment of social skills and have several advantages over traditional methods. Once the assessment is developed, professional time and training to use the system is minimal. Subjective bias and reliability issues, as well as recording errors, are eliminated because behaviors are automatically scored rather than being coded by observers. Social comparison data can be collected efficiently from a large group of students. Situations that are important for assessment but unlikely to be observed in naturalistic observation because of low frequency can be engineered into the assessment. The issue of observer reactivity can be overcome with *stealth assessment* techniques, in which assessments are "woven directly and invisibly into the fabric" of the game [21] and players' choices during gameplay provide the data needed for assessment. This greatly reduces the likelihood that students will alter their behavior to please an observer.

Further, traditional measures of students' social skills generally lack sensitivity of measurement and have limited utility for informing identification of students for social intervention or for tracking a student's response to that intervention. Computer-based systems offer the potential for more effective, sensitive, and reliable social skills assessment compared to traditional methods. Technology also offers a more engaging platform for students, an affordable method for broad scale everyday use by schools, and a seamless means of integrating data-driven decision making into school-based social interventions. This type of comprehensive in-game modeling of an individual player's knowledge or skills is becoming more common as Intelligent Tutoring Systems research and technology are increasingly applied to computer games [18, 22–24].

In the present study, we used theory-driven content and customized game mechanics to design and evaluate a stealth assessment computer game to determine children's social skills levels. We used extensive game log data and user feedback to revise the game during an iterative testing procedure. When the game was finalized, we used aggregate log data to create performance indices in order to assess students' social skills aptitude and examined correlations between in-game performance and independent measures of social skills to establish the validity of our computer game for social skills assessment.

Figure 1: Dialogue selection menu.

2. Method

2.1. Development of Zoo U. The primary design goal of *Zoo U* was to create virtual situations analogous to those commonly experienced by children in school settings in order to assess social problem solving strategies and aptitude. To accomplish this, we created *Zoo U* as a single-player point-and-click problem solving game situated in a school-like setting that, in addition to teacher and student NPCs (nonplayable characters), also contains zoo animals. The animals provide a bit of fun and fantasy to the school setting as well as opportunities for novel social problem solving tasks (e.g., working with other students to feed and care for the animals).

Game Mechanics. Each scene employs the same set of point-and-click mechanics and is rendered in 2.5 D, allowing for the appearance that the avatar can move in front of or behind objects in the scene via perceptual depth. For example, as depicted in Figure 1, the player is allowed to move to a location adjacent to one of the desks, which obscures the lower part of the avatar. The player clicks to move within a geometrically defined "walkable area" by using a modified A* algorithm to construct the walk path so that collisions between the avatar and objects in the scene are avoided.

Zoo U allows the player to initiate dialogue with NPCs by clicking on them. Players are informed of clickable objects and NPCs by the change in their standard cursor to one that glows blue. A subset of other objects in each scene also can be manipulated by clicking; some of these objects are integral for scene completion, and others are included as distractors similar to those in real world settings. The avatar initiates and/or responds to NPCs' dialogue by selecting from scripted dialogue choices (as shown in Figure 1) that are presented in random order. As the player moves the mouse over a choice, the corresponding actor-recorded audio is heard and the player selects a speech choice by clicking on it. In accordance

with the principles of Intelligent Systems [25], in nearly all cases a player's choice in one dialogue bubble impacts the available dialogue choices in subsequent interactions with a particular NPC in each scene.

Scene Content. Subject matter experts (SMEs) used educational and developmental psychology theories and empirical research to develop six independent social scenes to assess distinct social competencies (emotion regulation, impulse control, communication, empathy, cooperation, and initiation).

In each scene, players are presented with a social problem that needs to be solved. Content was developed to elicit measurable player behaviors and dialogue choices known to be integral to the social skill competency being assessed in each scene. For example, in the impulse control scene the player needs to feed the elephant before the class can go to recess. Players have access to two NPCs (the teacher and a peer) with which the student can interact and seek assistance in solving the presented task, as well as a number of clickable objects (e.g., food crates). Dialogue choices were written to assess particular components of the measured social skill (e.g., asking the teacher a specific question about the animal's name rather than asking her a more impulsive question about recess). Some of the objects are problem relevant (e.g., a clipboard on the wall that provides instructions for feeding the elephant), and others are included as distractors (e.g., the food crates, none of which contain food the elephant can eat). For both distractor NPCs and objects, students are allowed some exploration (e.g., the student is allowed to talk with a less-than-useful peer once). Students' scores decline; however, if they diverge too far from on-task behavior.

In order to receive a high impulse control score, the player needs to control his or her impulses to click on the distractor objects and direct full attention to appropriately determine how they should feed the elephant. Of course, each social skill assessed in *Zoo U* has unique requirements, and each scene varies with regard to how much emphasis is placed on interaction with NPCs versus interaction with the environment to solve a problem. Whereas scenes like impulse control, cooperation, and initiation are largely scored on students' behaviors in the scene (e.g., clicking on objects and time spent between behaviors), and social skills like emotion regulation and communication place greater emphasis on the individual and sequence of dialogue choices made when talking to NPCs.

2.2. Iterative Testing

Student Testing. In order to optimize both the gameplay and assessment quality of *Zoo U*, two independent iterative evaluations were conducted with third- and fourth-grade students and their teachers in two schools in central North Carolina. Parental consent forms were sent to the homes of all third- and fourth-grade students ($n = 254$) in regular classrooms. More than 90% of students received parental consent to participate. The first iterative test included students from one third- ($n = 16$) and one fourth-grade

($n = 23$) classroom and their teachers. Data collected from this initial sample were used to revise the content and gameplay options, as well as verify the difficulty of in-game challenges within each of the six social problem solving scenes. The revised scenes then were tested by a group of 187 students from 14 third- and fourth-grade classrooms and their teachers. Across both tests, students were approximately evenly divided across grades and genders and represented the full range of socioeconomic status with a racial distribution of 55% White, 32% African American, 7% Asian American, and 6% multiracial. Twenty-seven percent of students were of Hispanic/Latino ethnicity.

In both iterative test groups, testing took place during a regularly scheduled one-hour computer lab class at the school. Each child was assigned a computer on which to play *Zoo U* while research staff observed. Following the project introduction and computer orientation, students were given 30 minutes to complete the six *Zoo U* scenes. Teachers were asked to observe without providing assistance. Trained observers monitored and recorded students' level of attention, areas of difficulty, and reactions to *Zoo U* mechanics and content. Computer logs tracked the location and time of each mouse click so that students' interactivity could be mapped. The resulting data from the first iterative test group were used to refine any dialogue choices that did not relate in the expected directions with teacher ratings of students' social skills, as well as provide insights about particular game mechanics that gave students difficulty. The log data from the second iterative test were used to verify the refinements from the first iterative test as well as to calculate algorithms for students' performance in each *Zoo U* social problem solving task.

Student Ratings of Zoo U. After playing *Zoo U*, students completed a brief survey evaluating their experience with the software. We asked students to rate how fun/interesting the game was, how easy it was to use and understand, whether they liked the characters/graphics, whether they wanted to play more, and how much they liked it overall. Students rated these items on a four-point scale from 1 = "Not at all true for me" to 4 = "Very true for me." Research staff then led a group discussion to gather comments from students and suggestions for how to make the navigation or user experience better. In creating the second iteration of the game, we paid special attention to any ratings that were relatively low and refined the game accordingly.

2.3. Zoo U as a Valid Measure of Social Skills

Game Log and Scoring. At the time of the initial design of each scene, the SMEs created the content with the intention that children with varying levels of proficiency in the relevant social skill for that scene would perform differently. For instance, as shown in Figure 2, the impulse control scene contains three crates showing different possible foods for the elephant and a clipboard nearby labeled "Feeding Instructions," but the food crates are red herrings. Each time a player clicks a crate and attempts to feed the elephant that
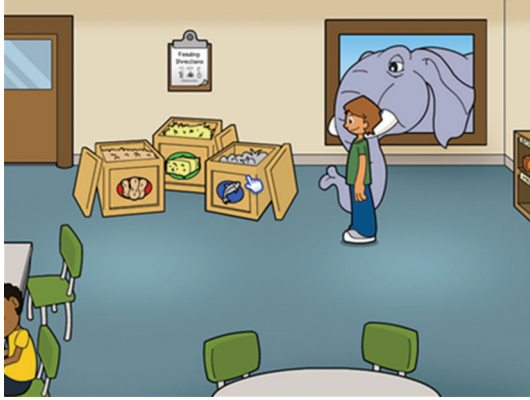
Figure 2: Impulse control scene.

kind of food, the elephant sticks out its tongue and the avatar says, "I don't think he likes it."

We expected impulsivity to vary directly with the number of clicks on the food crates, but the optimal number of clicks that would indicate varying levels of impulsivity was not evident at the time of initial design (i.e., a priori). To establish these optimal thresholds, gameplay was captured to provide minute annotated logging of the times and locations of every player click event. We parsed these logs to provide aggregate data, such as how many times a player clicked on a particular in-scene object, average response time to a dialogue choice menu versus response time on a particular choice, and the sequence of problem solving choices made by a particular student.

We then calculated a composite performance index (i.e., an algorithm) for each of the six scenes based on the aggregate information generated by the parsed log files for each scene (e.g., the number of times a player clicked on an individual NPC). Performance indices consisted of three core game-based components. The first component was scored based on dialogue choices to determine the quality of the student's attempts to solve the problem by communicating with the NPCs. For example, when asking the peer NPC for information, credit was given when an on-task menu choice was selected, whereas credit was not given when an off-task dialogue choice was selected.

In contrast to this menu-driven scoring method, the other two components assessed the quality of the student's behaviors while he or she interacted with the scene. The second component measured the amount of time spent engaged in appropriate problem solving activities versus inappropriate ones. For impulse control, this component was reflected via percentage of total time spent engaged in appropriate problem solving behaviors and reading the provided instructions on the clipboard versus time spent off-task (e.g., choosing off-task dialogue options and clicking on distractor objects). The third component measured the ratio of on-task versus off-task behaviors while completing the scene. In the impulse control scene, this was accomplished by calculating the number of impulsive clicks, including clicks on unrelated objects, and objects that the student had already learned were not useful for solving the problem.

Table 1: Student usability ratings of *Zoo U*.

Area rated	Mean (std)
Fun/interesting	3.90 (.31)
Easy to use and understand	3.95 (.22)
Liked characters and graphics	3.80 (.45)
Want to play more	3.95 (.22)
Overall liked	3.90 (.31)

Teacher Ratings of Students' Social Skills. An important aspect of this study was to examine the degree to which *Zoo U* performance was related to an independent external assessment of students' social skills (i.e., external validation). To this end, teachers completed an online survey rating the social skills of each of their students prior to students interacting with *Zoo U* (Table 1). This survey presented behavioral descriptors (e.g., "gets distracted easily," "good to have in a group"), and teachers rated the degree to which each descriptor was true of each student (from a low of "never true" to a high of "almost always true"). In order to target the teacher assessment more directly to the social skills and behaviors assessed through *Zoo U*, items for this survey were drawn from previously validated measures of social skills in this age group, including the Teacher Checklist [26], the Social Skills Rating Scale [27], and the Social Competence Scale-Teacher Version [28]. The resulting Social Skills Behavior Inventory (SSBI) [29] included 34 items across the six social skills scales (communication, cooperation, empathy, initiation, impulse control, and emotion regulation). Internal consistency for each subscale was acceptable (*mean Cronbach* α = .86) and was statistically similar for ratings of third- and fourth-grade students.

3. Results

3.1. Student Ratings of Zoo U. Overall, students rated *Zoo U* very positively with every area rated ≥3.8 on a four-point scale. In addition, researcher observations revealed high levels of student engagement with 96% on-task behavior and numerous positive comments about the game (e.g., "This is awesome!"; "When can I play more?"). Students easily understood how to navigate *Zoo U*, and analysis of usage data indicated a low rate of errors (e.g., misclicks). Researchers observed very few misunderstandings or requests for technical help. Desire for replay was also strong; after students had completed all six scenes of *Zoo U*, they were told that they could replay *Zoo U* a second time or go online to play other games for the remainder of the session. We were pleased that 89% of students elected to replay *Zoo U*.

3.2. Zoo U as a Valid Measure of Social Skills. To assess external validity, we conducted correlational analyses to examine the relations between *Zoo U* performance indices and the teacher ratings of students' social skills for each subscale. Table 2 displays these results. For all six scenes, the *Zoo U* composite performance index was significantly correlated with the teacher rating on the analogous SSBI

Table 2: Correlations between *Zoo U* composite performance and teacher ratings of student social skills.

SSBI teacher rating scales	*Zoo U* composite performance index					
	Imp con	Emo reg	Comm	Coop	Init	Emp
Impulse control	**.29*****	.31***	.15*	.16*	.17*	.16*
Emotion regulation	.15*	**.30*****	.08	.17*	.25**	.17*
Communication	.26***	.28***	**.20****	.20**	.27***	.19*
Cooperation	.23**	.31***	.12	**.21****	.16*	.17*
Initiation	.17*	.14*	.08	.18**	**.27*****	.15*
Empathy	.26***	.22**	.10	.17*	.24**	**.23****

Note: ***$P < .0001$, **$P < .01$, and *$P < .05$.

Table 3: Intercorrelations amongst *Zoo U* subscales.

	Emo reg	Comm	Coop	Init	Emp
Impulse control	.23**	.07	.21**	.15*	.24**
Emotion regulation		.20**	.09	.26**	.16*
Communication			.07	.19*	.10
Cooperation				.21**	.16*
Initiation					.14

Note: **$P < .01$ and *$P < .05$.

subscale and correlated in expected ways with other SSBI subscales.

To ensure that *Zoo U* composite indices were discriminating amongst the targeted social skills, intercorrelations between *Zoo U* subscales were calculated to determine the degree to which *Zoo U* subscales were independent. Table 3 displays the results of those correlations. Although *Zoo U* subscales are related in a number of cases, the composition of those associations is expected given the overlap of social skills competencies in children's real world behaviors (e.g., it is not surprising that impulse control is correlated with emotion regulation).

4. Discussion

Computer games provide a promising avenue for the assessment of student skills. It is likely that computer-based assessments will be utilized heavily in the future, as the student assessment paradigm shifts from a single-source (e.g., teacher) time-intensive approach to a more multifaceted and interactive methodology that will require a new way of thinking about student identification. In the present study, experts in educational/developmental psychology and computer science collaborated to create and test an innovative stealth assessment of children's social skills. This study contributes to the improvement of computer games in education by describing a novel design methodology for developing these kinds of assessments.

This study underscores the potential cumulative value of assessing multiple dimensions of player behavior when formulating student performance indices. In-game data logs captured students' dialogue choices, behavioral choices, and time on task and off task. *Zoo U*'s performance indices were developed, tested, and refined based on both theory and these data logs. These performance indices then were validated by measuring their association with a standard teacher report of students' social skills.

A key challenge in designing *Zoo U* was determining the appropriate level of difficulty in order to garner enough student variability in responses to develop useful performance indices. The starting point of our solution was to employ an iterative design process. The results of this process demonstrated that stealth assessment principals can be used in the context of identifying children's levels of social skills in ways that are commensurate with standard identification practices (i.e., teacher report). We currently are testing these performance-derived indices with a larger, more nationally representative sample of students to further calibrate the in-game assessment of students' social skills.

The results of this assessment study are linked to the promise of educational games' potential to offer opportunities for instruction and intervention to build children's social and emotional skills. Merrill [18] notes that "the present and future challenge in assessment is to find meaningful ways to make assessment results functional, in the sense of tying specific results to important social outcomes and to the development of effective instructional and therapeutic programs." We concur with the importance of this challenge and currently are developing a number of computer-based intelligent social tutoring systems that will utilize *Zoo U* assessment for the identification of students' social skills aptitude. The assessment capabilities of *Zoo U* offer flexibility for use as both a prepost assessment of intervention effects and as a way of modulating the difficulty so that students are continuously challenged but not frustrated (i.e., scaffolded learning). Utilizing this capability, we are developing tutorials for both universal and indicated populations with the intent of improving children's social skills strategies across a range of functioning.

In this study, we established the validity of *Zoo U* with one of the gold standards of social skills assessment, a behavior rating scale. Future research should further validate *Zoo U* by using the other gold standard, naturalistic observation. We believe that because *Zoo U* was developed specifically to provide game-based social problem solving scenarios analogous to authentic situations children encounter, ecological validity will be maintained and that many of the limitations (e.g., observer bias and reactivity, recording error) of naturalistic observation will be minimized.

Zoo U leverages innovative technologies to provide an engaging and powerful social skills assessment tool with real-time reporting functions for educators to easily track students' progress toward social goals. Online access makes *Zoo U* cost effective and enables easy access for students and schools across the nation, making it an appealing option for social skills assessment. Data derived through this new form of assessment can be used to inform decisions regarding implementation of social interventions by schools, to identify children in need of social skills interventions, and to track progress over the course of an intervention. Compared to standard measures, the engaging nature of computer games and the efficiency with which assessments can be conducted and scored make them less arduous for both students and teachers and more accessible, informative, and effective on a broad scale.

Disclosure

Acknowledgment

References

[1] J. Kupersmidt and M. DeRosier, "How peer problems lead to negative outcomes: an integrative mediational model," in *Children's Peer Relations: From Development to Intervention*, pp. 119–138, American Psychological Association, 2004.

[2] J. Parker, K. Rubin, S. Erath, J. Wojslawowicz, and A. Buskirk, "Peer relationships, child development, and adjustment: a developmental psychopathology perspective," *Developmental Psychopathology*, vol. 1, pp. 419–493, 2006.

[3] W. Furman and D. Buhrmester, "Age and sex differences in perceptions of networks of personal relationships," *Child Development*, vol. 63, no. 1, pp. 103–115, 1992.

[4] A. E. Kazdin, "Effectiveness of psychotherapy with children and adolescents," *Journal of Consulting and Clinical Psychology*, vol. 59, no. 6, pp. 785–798, 1991.

[5] S. Luthar, *Resilience and Vulnerability: Adaptation in the Context of Childhood Adversities*, Cambridge University Press, New York, NY, USA, 2003.

[6] M. Boivin and S. Hymel, "Peer experiences and social self-perceptions: a sequential model," *Developmental Psychology*, vol. 33, no. 1, pp. 135–145, 1997.

[7] S. A. Erath, K. S. Flanagan, and K. L. Bierman, "Social anxiety and peer relations in early adolescence: behavioral and cognitive factors," *Journal of Abnormal Child Psychology*, vol. 35, no. 3, pp. 405–416, 2007.

[8] J. V. Carney, "Bullied to death: perceptions of peer abuse and suicidal behaviour during adolescence," *School Psychology International*, vol. 21, no. 2, pp. 213–223, 2000.

[9] M. Brendgen, F. Vitaro, and W. M. Bukowski, "Affiliation with delinquent friends: contributions of parents, self-esteem, delinquent behavior, and rejection by peers," *Journal of Early Adolescence*, vol. 18, no. 3, pp. 244–265, 1998.

[10] M. E. Feinberg, T. A. Ridenour, and M. T. Greenberg, "Aggregating indices of risk and protection for adolescent behavior problems: the Communities That Care Youth Survey," *Journal of Adolescent Health*, vol. 40, no. 6, pp. 506–513, 2007.

[11] J. D. Hawkins, R. F. Catalano, and J. Y. Miller, "Risk and protective factors for alcohol and other drug problems in adolescence and early adulthood: implications for substance abuse prevention," *Psychological Bulletin*, vol. 112, no. 1, pp. 64–105, 1992.

[12] C. Spooner, "Causes and correlates of adolescent drug abuse and implications for treatment," *Drug and Alcohol Review*, vol. 18, no. 4, pp. 453–475, 1999.

[13] C. B. Fleming, K. P. Haggerty, R. F. Catalano, T. W. Harachi, J. J. Mazza, and D. H. Gruman, "Do social and behavioral characteristics targeted by preventive interventions predict standardized test scores and grades?" *Journal of School Health*, vol. 75, no. 9, pp. 342–349, 2005.

[14] L. J. Woodward and D. M. Fergusson, "Childhood peer relationship problems and later risks of educational under-achievement and unemployment," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 41, no. 2, pp. 191–201, 2000.

[15] J. D. Coie, K. A. Dodge, R. Terry, and V. Wright, "The role of aggression in peer relations: an analysis of aggression episodes in boys' play groups," *Child development*, vol. 62, no. 4, pp. 812–826, 1991.

[16] M. E. DeRosier, J. B. Kupersmidt, and C. J. Patterson, "Children's academic and behavioral adjustment as a function of the chronicity and proximity of peer rejection," *Child Development*, vol. 65, no. 6, pp. 1799–1813, 1994.

[17] M. Greenberg, C. Domitrovich, and B. Bumbarger, "Prevention of mental disorders in school-aged children: current state of the field," *Prevention & Treatment*, vol. 4, no. 1, pp. 1–52, 2001.

[18] K. W. Merrell, "Assessment of children's social skills: recent developments, best practices, and new directions," *Exceptionality*, vol. 9, no. 1-2, pp. 3–18, 2001.

[19] K. W. Merrell, *Behavioral, Social, and Emotional Assessment of Children and Adolescents*, Laurence Erlbaum Associates, Mahwah, NJ, USA, 1999.

[20] K. W. Merrell and G. A. Gimpel, *Social Skills of Children and Adolescents: Conceptualization, Assessment, Treatment*, Laurence Erlbaum Associates, Mahwah, NJ, USA, 1998.

[21] V. J. Shute and F. Ke, "Games, learning, and assessment," in *Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives*, D. Ifenthaler, D. Eseryel, and X. Ge, Eds., Springer, New York, NY, USA, 2012.

[22] Y. Cheong, A. Jhala, B. Bae, and R. M. Young, "Automatically generating summaries from game logs," in *Proceedings of the 4th Artificial Intelligence and Interactive Digital Entertainment International Conference (AIIDE '08)*, pp. 167–172, 2008.

[23] J. Rowe and J. Lester, "Modeling user knowledge with dynamic bayesian networks in interactive narrative environments," in *Proceedings of the 6th Annual AI and Interactive Digital Entertainment Conference*, pp. 57–62, Palo Alto, Calif, USA, 2010.

[24] A. Tveit and G. B. Tveit, "Game usage mining: information gathering for knowledge discovery in massive multiplayer games," in *Proceedings of the International Conference on Internet Computing (IC '02)*, Session on Web Mining, 2002.

[25] V. J. Shute and D. Zapata-Rivera, "Educational assessment using intelligent systems," Educational Testing Service Report RR-08-68, ETS, Princeton, NJ, USA, 2008.

[26] M. E. DeRosier and S. H. Mercer, "Improving student behavior: the effectiveness of a school-based character education program," *Journal of Research and Character Education*, vol. 5, pp. 131–148, 2007.

[27] F. Gresham and S. Elliott, *Social Skills Rating System*, American Guidance Service, Circle Pines, Minn, USA, 1990.

[28] Conduct Problems Prevention Research Group (CPPRG), Social Competence Scale (Teacher Version), 1990, http://www.fasttrackproject.org/.

[29] M. E. DeRosier, "Using computer-based social tasks to assess students' social skills: findings from the Zoo U pilot evaluation," Final Report for the U.S. Department of Education, Washington, DC, USA, 2011.