Investigating Preservice Science Teacher Ethical Sensitivity through Computer Game and Video

Shawn Y. Holmes[1], Leonard Annetta[2], Loni Crumb[2]
[1]The You in You Consulting, Inc.
[2]East Carolina University

To cite this article:

JESEH

Investigating Preservice Science Teacher Ethical Sensitivity through Computer Game and Video

Shawn Y. Holmes, Leonard Annetta, Loni Crumb

Article Info	Abstract
	A computer-based simulation, Hazelton High at REST (HHR), with embedded performance-based assessments and Likert-type survey questions was created to assess preservice teacher recognition of racial and gender intolerant behaviors. The simulation was modeled after the Racial Ethical Sensitivity Test (REST), a reliable video-based assessment, measuring professionals' ethical sensitivity for cultural competence and recognition of intolerant behaviors in school settings. Ethical sensitivity skills, based on education professional ethics, are used to determine knowledge of ethical responsibility and professional attitudes and behaviors. A mixed-methods comparison study was conducted between HHR and the REST-video with 31 preservice science education students to investigate the effect of each on their ethical sensitivity. A post-simulation assessment was also conducted to determine and the usability and perception of HHR. It was hypothesized; HHR, an immersive simulation, would affect perspective taking of preservice science teachers and indirectly influence their ethical sensitivity. Data analysis determined ethical sensitivity score (EES) for HHR and REST-video using a survey and performance-based assessments. Post-simulation interview helped inform how a participant perceived the HHR. Results indicate the REST-video group had a significantly higher EES than the HHR group. Performance-based assessments within HHR showed a more comprehensive and complex picture of participant EES. The implication of this work for science education programs to use professional codes as a guide for evaluation and development of candidate disposition concerning cultural competency and the use of Serious Educational Games (SEGs) as ethical sensitivity assessments.

Introduction

Science education standards set expectations and principles for teachers to uphold the highest ethical standards and recommend teachers gain professional dispositions to support learning (Council for the Accreditation of Education Preparation (CAEP) 2013; National Science Teacher Association Standards (NSTA), 2012). Furthermore, the NSTA Standards for Science Teacher Preparation (NSTA-SSTP) (NSTA, 2012) was written to ensure science teacher education programs prepare preservice teachers with the appropriate knowledge and skills that are necessary to create learning environments that engage diverse learners. The Principles of Professionalism for Science Educators (PPSE) calls for science educators to conduct themselves in a moral and ethical manner in order to provide all learners with the best science education (NTSA, 2010). Essentially, the language in these documents speaks to the moral foundations and basis of teaching science with ethical sensitivity and competency in cultural and linguistic diversity. Despite the principles espoused in these standards, there is a continued need to connect science teacher education standards with professional ethics and cultural competence.

Too often, the professionalization of teaching places emphasis on knowledge and skill development and less attention is given to ethical and moral development. Fenstermacher (1990) reminds us why teaching is a moral activity:

> What makes teaching a moral endeavor is that it is, quite centrally, human action undertaken in regard to other human begins. Thus, matters of what is fair, right, just, and virtuous are present. Whenever a teacher asks a student to share something with another student, decides between combatants in a schoolyard dispute, sets procedures for who will go first, second third, and so on, or discusses the welfare of a student with another teacher, moral considerations are present. The teacher's conduct at all times and in all ways, is a moral matter. For that reason alone, teaching is a profoundly moral activity (p. 133).

Highlighting the essence of situating science education in a sociocultural perspective, Lemke (2001) argued, "The most sophisticated view of knowledge available to us today says that it is a falsification of the nature of science to teach concepts outside of their [students'] social, economic, historical, and technological contexts" (p. 300). Overall, science educators facilitate the development and construction of science concepts for learners by assuming the responsibility for enabling each learner to reach their potential. This includes empowering learners to think and respond with their cultural, ethnic, and linguistic differences, making science personal rather than abstract (NSTA, 2006). The Preservice Science Standards, Standard 6: Professional Knowledge and Skills speaks more specifically, "Effective teachers of science strive continuously to improve their knowledge and understanding of the ever changing knowledge base of both content, and science pedagogy, including approaches for addressing inequities and inclusion for all students in science" (NSTA, 2012).

Teachers in schools populated by underrepresented learners (e.g., racial/ethnic minorities, economically disadvantaged, immigrants) are guided by their personal beliefs, attitudes, and knowledge bases and can easily be unaware of their actions in the classroom (Fras-Abder, 2001; Welton & Martinez, 2014). Teaching is a "socially negotiated activity" that requires the refinement of moral sensibilities (Buzzelli & Johnston, 2002). If unfamiliarity and fear of teaching underrepresented learners are coupled with gaps in knowledge and skills, teachers' actions may be perceived as not enriching the curriculum with diverse cultural perspectives (Carter, Larke, Singleton-Taylor, & Santos, 2003; Villegas & Lucas, 2002), not recognizing acts of intolerant behavior, and being unaware of cultural influences on learning (Bryan & Atwater, 2002). Thus, it is appropriate to develop preservice teacher's ethical sensitivity skills to create equitable teaching and learning environments in science classrooms. Teacher preparation programs have a critical responsibility in this endeavor. The rationale for Standard 3: Candidate Quality, Recruitment, and Selectivity notes that educator preparation provider's responsibility begins with purposeful recruitment, continues through selectivity, and monitoring progress while providing support for proficient completion and being chosen for employment opportunities.

The Science Classroom

Science teachers as much as any other teacher will be confronted with dilemmas surrounding learner identity and unconscious biases. Learners often seek out teachers for advice concerning the social issues in their lives and these issues can arise during classroom discourse. Science teachers are not merely employed or expected to just teach science. The PPSE instructs science teachers to "reflect a positive professional image by being cognizant of the image portrayed to students, parents and the community through speech, attire, and action…uphold the highest standards of ethical behavior and be positive role models…protect, respect, and empower students" (NSTA, 2007, p. 1-3). This language calls for the attention to the development of professional ethics and professional responsibility in science teachers during their tenure in science education programs. CAEP acknowledges research that supports non-academic teacher qualities often called attributes, dispositions and abilities. Further, CAEP recognizes Interstate Teacher Assessment and Support Consortium (InTASC) standards as an area of growing development, research and interest in teacher knowledge, disposition and performance. The Professional Knowledge and Skills NSTA Preservice Science Standard, combines continuous improvement of content knowledge and science pedagogy with "approaches for addressing inequities and inclusion for all students in science." (2012).

Performance-based Assessments and Computer Games to Assess Professional Performance

Researchers and scholars in various professions have suggested that performance-based assessments and simulations provide insight to alternative pathways of thinking for professionals. For example, Hmelo et al. (2001) used case-based scenarios containing detailed patient histories to assist pre-medical professionals in learning to design clinical trials for oncology treatment. Performing medical consultations is an area where there was extensive use of cased-based simulations. Waldmann, Gulich and Zeitler (2008) utilized the case-based simulations and had medical students role-play as a physician. The researchers noted that simulations provide information beyond a paper-based assessment, mentors had a better idea about how participants approached a problem, and simulations reduced the number of training days (Hmelo et al., 2001; Waldmann et al., (2008).

Additionally, attributes for using digital technologies to study human behavior are, increased realism (Frenz, 2007), incorporating a sense of presence (Annetta & Holmes, 2006; Tung & Deng, 2006), presenting case-base scenarios (Annetta et al., 2009; Waldmann et al., 2008), and use of a risk-free, immersive setting that supports constructivist learning by causing cognitive dissonance (Dede, 2005; James Gee, 2003; Shaffer, 2007). Shaffer

suggested that simulations are effective for training because a person put into a simulated world learns to be, think, and do within the parameters of the rules implied in the simulation. This is especially evident for role-playing simulations. Participants involved in role-playing, immersive and case-based scenarios are actively involved with enhanced realism and perspective taking. Perspective taking provides the opportunity to induce cognitive conflict and consider others' viewpoints, an element of ethical sensitivity and moral development. This allows other abilities, skills, and knowledge to be assessed and since assessment drives learning, performance-based assessments can benefit teaching and research alike (Frenz, 2007; Waldmann et al., 2008). Ethical sensitivity skills make unconscious biases conscious, allow a professional to see the situation from the standpoint of the other and connects the professional regulations, codes, and norms and when they apply.

In respect the extant literature, the study discussed in this paper was a response to a call to action from the science education research community (Brown, 2006; Bryan & Atwater, 2002; Fraser-Abder, 2001) and science education professional organizations (AAAS, 1989; NSTA, 2007, 2012) for innovative methodologies that encourages ethical responsibility and creates learner-centric, culturally responsive learning environments in science classrooms.

Theoretical Perspective

This investigation was situated in the theoretical framework of the Four Component Model (FCM) (see Figure 1), which is grounded in ethical development and professional ethics. Rest (1986) used his investigations in moral development and generated the FCM for morality as an explanation for the effects of cognition, affect, and social dynamics on moral action. With the implementation of the Defining Issues Test (DIT) and the FCM, Rest changed how moral development was viewed and assessed (Rest, Narvaez, Bebeau & Thoma, 1999). This change moved traditional thinking about moral development from stage development theory to schema theory. According to Rest, cognition and affect interact in all areas of moral behavior (Bebeau & Monson, 2008). The FCM identifies the following four integrated abilities as needed conditions for effective ethical decision-making and behavior (a) moral sensitivity, recognize the situation as moral, empathize with the effect on others; (b) moral judgment, integrating and weighing various moral considerations; (c) moral motivation, prioritize moral values relative to other values; and (d) moral implementation, requiring moral character, conviction and courage (Brabeck et al., 2000; Rest & Narvaez, 1994). Following Rest, Bebeau (Bebeau, Rest, & Yamoor, 1985) developed the Dental Ethical Sensitivity Test (DEST) for dental students and Brabeck (Brabeck et al., 2000) developed the Racial Ethical Sensitivity Test (REST) for preservice teachers. Both instruments assess the first component of the FCM, moral sensitivity; in their respective professional areas (see Figure 1). A stimulation called *Hazelton High at REST* was developed for this study using the FCM framework; the content for the game is from the REST for preservice teachers.
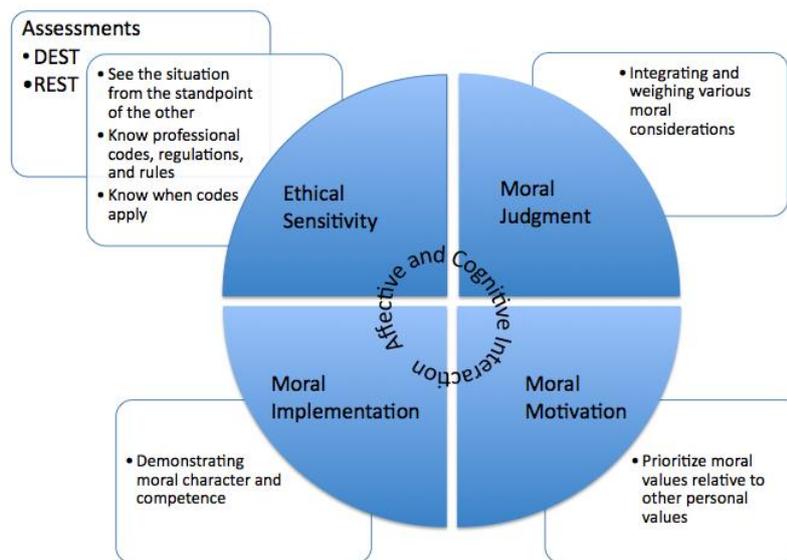
Figure 1. FCM professional development perspective

Research Purpose

In order for implementation of ethical sensitivity skills to occur, the professional must recognize the moral aspects in the situation (Bebeau & Monson, 2008). Therefore, the focus and rationale for this study was to facilitate recognition of racial and gender intolerant behaviors, moving unconscious bias to awareness as a way to initiate judgment, motivation and implementation of professional responsibilities; in other words to effect preservice science teachers' ethical sensitivity. This intent was achieved by creating a first-person perspective Serious Educational Games (SEG) (Annetta, 2008) *Hazelton High at REST,* which contains performance-based assessments and emulated after the REST for preservice teachers. The REST assesses educator ethical sensitivity to racial and gender intolerant issues in schools (Brabeck et al., 2000; Sirin, Rogers-Sirin, & Collins, 2010; Sirin, Brabeck, Satiani, & Rogers-Sirin, 2003). The rationale to use a SEG to assess ethical sensitivity focused on the immersive nature, enhanced realism, and the ability to cause cognitive dissonance in a risk free environment when using simulations (Bos, Shami, & Naab, 2006; Eck, 2006; Gee, 2005; Hmelo et al., 2001; Macedonia, 2002; Shaffer, Squire, Halverson, & Gee, 2005). The implementation of a computer simulation in the evaluation and development of professionalism in science teacher education might encourage more reflection of educator responsibility, awareness to implement professional standards, provide opportunities for technological instruction, and present a format to explore and increase ethically defensible responses to intolerant behaviors.

Hazelton High at REST

The current methodology involved the creation of the SEG *Hazelton High at REST* (created by the first author) with specific immersive characteristics and the implementation of an embedded performance-based assessment. The creation of *Hazelton High at REST* is a technological study and is out of the scope for this article. The embedded content in the SEG was grounded in professional ethics and ethical development and based on recommendations of NSTA professional standards and science education research. *Hazelton High at REST* was a response to issues associated with gaps in science teachers' ethical sensitivity and multicultural competence. Using a SEG with performance-based assessments of professional attitudes and behaviors were used as evidence for knowledge of ethical responsibility.

Hazelton High at REST was compared with the video REST to elucidate usability issues of the SEG as well as its use as a performance-based assessment for ethical sensitivity with science educators. The preliminary research question for this study was: how does an immersive 3D SEG affect ethical sensitivity of science educators when compared to the video case assessment? In addition to the embedded assessment, three open-ended questions were asked immediately following the SEG experience to gain more information concerning how the participants perceived the SEG. This research is part of a larger quasi-experimental comparison study between preservice and in-service science teachers in addition to comparing the SEG and video experiences.

Assessments of Professional Ethics

The DEST assesses a dental student's ability to identify and respond to moral dilemmas embedded in dentist-patient interactions via an audiotape. Bebeau rephrased moral sensitivity to ethical sensitivity because the ability of professionals or training professionals to interpret factors directly related to obligations in their professional code of ethics was being assessed.

Similar to the DEST, the REST assesses an educator's ability to recognize racial and gender ethical violations embedded in interactions between education professionals and learners using video recordings instead of audio recordings. Brabeck et al. (2000) depicted overt and covert behaviors in the REST videos, that violated ethical principles of professional codes of ethics from the teaching, social work, school administration, and school psychology professions, as well as groups who worked specifically with learners from culturally and linguistically diverse populations. Six ethical principles common to all professional codes reviewed were incorporated in the REST. They were; 1) professional competence, 2) integrity, 3) professional and scientific responsibility, 4) respect for others' rights and dignity, 5) concern for others' welfare, and 6) social responsibility. Focusing on ethical sensitivity skills for racial and gender intolerant behaviors in schools, Brabeck et al., (2000) specified the need for educators to (a) recognize verbal and nonverbal behavior; (b) identify the needs or wants of others in the situation; (c) predict others reactions to help; and (d) to respond with appropriate concern.

The DEST and REST assessments are similar in that both ask the training professional to assume the role of the professional in the scenarios thereby initiating a role-playing assessment. The scenarios present clues about ethical issues without signaling what the issue is or what professional responsibility is needed (Bebeau & Monson, 2008). Both assessments ask open-ended interview questions to determine ethical sensitivity on a scale of 1 to 3, thus providing researchable variables and authentic measures of professional performance (Bebeau & Monson, 2008).

Hawthorne Effect

The Hawthorne Effect as defined by Payne and Payne (2004) became an additional lens for this study. Initial participant reaction and responses to *Hazelton High at REST* led to a re-examination of the post-SEG responses and opened consideration of the Hawthorne Effect. According to Payne and Payne, "The Hawthorne Effect is the tendency, particularly in social experiments, for people to modify their behavior because they know they are being studied, and so to distort (usually unwittingly) the research findings," (p. 107). The following question was investigated during a re-examination of responses after participants completed *Hazelton High at REST*; did the immersive 3D SEG interfere with participant recognition of being studied?

The Hawthorne studies contributed to a deeper appreciation of factors affecting participants and encouraged social science and education researchers to focus more on the interactions between the participant and the study environment (Holden, 2000; Peacock, 2005). Investigations in this realm ask about the effectiveness of interventions used with professionals and the transfer of skills to work environments. Separating which parts of an intervention improved the performance of a professional is desirable yet complex and the Hawthorne Effect shed light on the "artificial nature" of these types of studies.

Methods

Hazelton High at REST was an immersive first person SEG comprised of two interactive REST scenarios (explanation in the next section) utilizing original audio tracks, five interspersed performance-based prompts, 36 Likert-type survey questions, and three post-SEG prompts, all of which unfold at a fictitious virtual high school that provided context for the REST SEG scenarios. An NSF grant provided funding to create the high school SEG containing two REST scenarios, Basketball Practice (BP) (see Figure 2) and Faculty Lounge (FL) (see Figure 3).
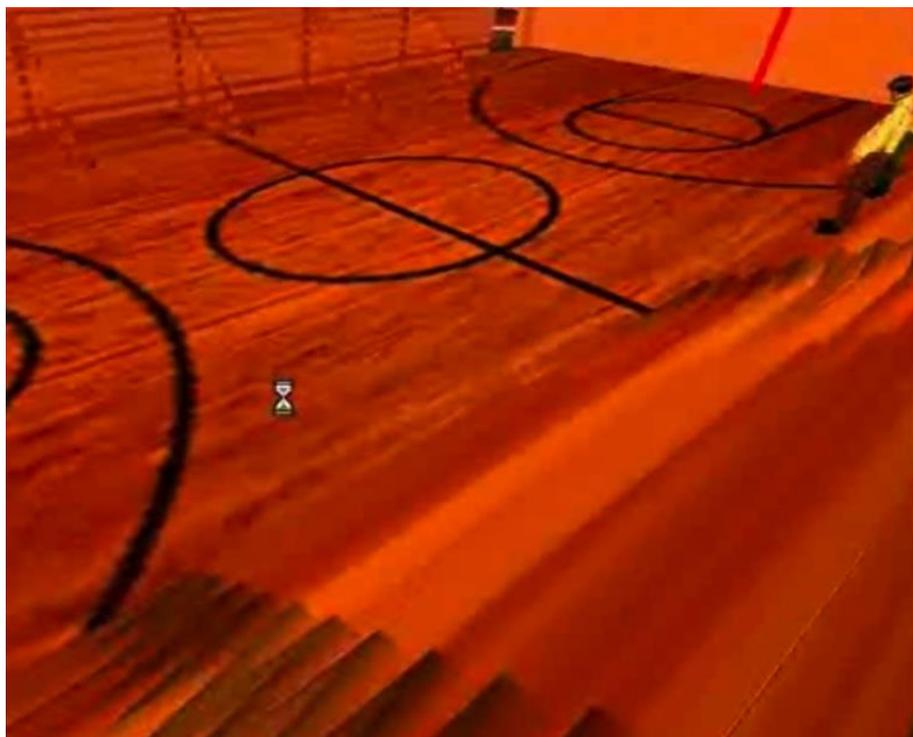
Figure 2. Screen shot of basketball practice scenario

Figure 3. Screen shot of student in the school counselor's office

A drag-and-drop authoring software was used to transform the original video format into an interactive first person SEG. *Hazelton High at REST* contained virtual characters, an audio track, and school settings similar to the REST video scenarios to facilitate a comparison study and gain knowledge about the usability of the SEG. The BP and FL scenarios occurred in a gymnasium, student counselor's office, and faculty lounge of a high school. Hazelton High SEG contained these virtual environments as well as a science classroom, laboratory prep room with desks and office space for five teachers, hallways, and a locker room. The additional rooms were designed to influence immersion during the SEG experience. Participant responses about the effect of the SEG on their responses initiated an additional exploration in the Hawthorne Effect (Adair, 1984; Holden, 2000; Payne & Payne, 2004).

Video REST instrument

The REST was constructed around identifying intolerant racial and gender related behaviors that signify ethnically insensitive conduct in US schools. The instrument comprised of five video scenarios. Embedded in each scenario were behaviors depicting five to nine ethical issues; eight were in the BP and nine in the FL. The quantitative version of the REST and Quick-REST (Sirin et al., 2010), uses both BP and FL scenarios, each containing a Likert-type survey of 18 questions. The REST qualitative version uses an open-ended interview protocol by Brabeck et al. (2000). Response categories for the interview protocol were scored 1-3, with 1 being the lowest level of recognition: (a) no recognition of intolerant behavior; (b) recognition of intolerant behavior; and (c) recognition of intolerant behavior, elaboration of the implications, sensitivity, and complexity of issues within the context depicted in the scenario (Brabeck et al., 2000). Both quantitative and qualitative assessment strategies were used in this study. The following is a brief synopsis of the BP and FL scenarios:

> Basketball Practice: A Black student enters the gymnasium a few minutes late for practice and is chewed out by the coach in front of the team. Being made to run extra laps then punishes him. Meanwhile, a White student is sitting on the bleachers making out with his girlfriend and ends up being even later that the first student, in addition to being out of uniform. The coach chides him for being a "stud" and does not make him do laps, and then tells the White student that Black students keep "guys like you" out of school. Later, the Black student complains to his guidance counselor about the racist behavior of the coach. His counselor minimizes the problem and tells him to stick it out, stating that he needs basketball to get into college. This is despite the fact that the student is on the honor roll.

> Faculty Lounge: Two teachers are discussing a student in front of a new faculty member in the faculty lounge. The two veteran teachers discuss the student's academic and private life in stereotypical and derogatory ways. They show no concern for the student's privacy and a complete disregard for her

rights to confidentiality. In addition, it is clear that they have no understanding of her culture. When the new faculty member tries to share her thoughts and stand up for the student, she is met with hostility and ridicule (Brabeck et al., 2000, p. 125 - 126).

Simulation Hazelton High at REST

The *Hazelton High at REST* SEG began by Mr. Delaney, a school administrator, greeting a participant outside the science classroom and saying, "So you think you can cut it at Hazelton High? Well, we have a couple of science teacher positions open. I can't make any promises now; we'll see how you do in the game scenarios." This introduction was intended to pique interest and the realism of the SEG. Afterward, participants were free to walk around the science classroom and laboratory preparation areas allowing them to become familiar with Hazelton High. Mr. Delaney reappears and notes, "The game will start when you click on the red button by the door." Upon clicking the red button, the participant was teleported into the counselor's office where they are briefed by the school counselor about the BP scenario and told to "Assume you are me…." This instruction was intended to have participants role-play and respond as the professional in the scenario. Mr. Delaney returned twice during the BP scenario and asked what the participant, as the professional, would say to the student in the scenario.

Participants typed their responses to two prompts into a text box on the screen, submitted and the scenario resumed automatically. At the conclusion, the participant answered 18 survey questions and "walked" back through the laboratory preparation area, clicked the black button to begin the FL scenario. Similar to the BP scenario, a new teacher greeted the participant, introduced the FL scenario and said, "Assume you are me the new teacher." Participants typed their responses to three prompts during the FL scenario. At the conclusion, participants answered another set of 18 survey questions followed by three open-ended prompts about their SEG experience.

Participant Sample and Setting of Study

Recruitment efforts utilized electronic mail to contact secondary science education preservice teachers from eight universities in the mid-Atlantic region for this posttest only research design (Campbell & Stanley, 1963). Thirty-two science teacher candidates responded from three universities located in two states. Testing occurred at participant school sites. Contacts from the remaining universities either did not have secondary science education preservice teachers or did not respond to recruitment requests. The preservice sample contained 22 females and 10 males and the self-reported ethnic composition was: 91% Caucasian, 6% African American and 3% West Indian. Participant ages ranged from 22 – 27 years. Fifty-six percent reported taking classes or workshops for ethics, and/or multicultural issues as part of their professional training, however it was not known whether attendance was voluntary.

Participants were assigned to one of two groups, "V" for participating in the REST video and "S" for participating in the Hazelton High at REST SEG. *Hazelton High at REST* was operational on 15 laptops and accessed through the Internet. Since testing occurred at participant school sites, Internet access became the determining factor for group placement, V or S. Consequently, there were 18 in the V group and 14 in the S group. It is important to note that many students referred to the SEG as a simulation and that is the term we used when quoting student responses.

Data Collection

The procedural sequence for comparison groups V and S are in Table 1. At a decision-making point for the education professional in the scenarios, usually during conversations between virtual characters, the performance-based prompts appeared in a question text box on the screen. Assuming the role of the education professional, participants wrote what they would say in the situation. The scenario proceeded after responses were submitted. All embedded prompt responses, survey responses, and post-game responses were sent electronically and stored on a server then aligned by a unique login created by each participant. A server upgrade during the study resulted in the loss of a small amount of data. Responses were cross-referenced with each other and participant demographic information and analyzed.

Table 1. An overview of the procedural sequence

BP scenario	Scenario	Performance-based assessment	Quick-REST assessment	Post-game prompts
V Group	Watch REST video	N/A	18 BP Likert-type questions	N/A
S Group	Participant in Hazelton High at REST SEG	1. James is very upset and accuses you and administrators of knowing about Coach Nichols, how would you respond to James? 2. How would you counsel James? Write what you would say to him.	18 BP Likert-type questions	1. What was this experience like for you? 2. How might this experience be different from viewing a video? 3. How might these differences affect your responses?
FL scenario	Scenario	Performance-based assessment	Quick-REST assessment	Post-game prompts
V Group	Watch REST video	N/A	18 FL Likert-type questions	N/A
S Group	Participant in Hazelton High at REST SEG	3. How would you respond to the students who made the stereotypic remarks? 4. As Ms. Highland the new teacher, write what you would say (if anything) in response to Mr. Lynch's comment about Latinas? 5. As Ms. Highland the new teacher, what if anything, disturbs you about this conversation?	18 FL Likert-type questions	1. What was this experience like for you? 2. How might this experience be different from viewing a video? 3. How might these differences affect your responses?

Data Analysis

Non-parametric analysis on the validated 36 Likert-type Quick-REST survey was conducted. Non-parametric tests are suitable since we cannot assume a normal distribution with the lower N in this study. The survey was a 5-point scale proceeding from, "I strongly disagree" to "I strongly agree." Ten of 18 questions were reverse-coded for the BP survey and eight of 18 were reversed-coded for the FL survey. Responses from the 36 questions created a score that reflects ethical sensitivity. A high score indicates a more ethically sensitive participant towards issues of racial intolerance in schools (Sirin et al., 2010). Scores were determined using a SPSS syntax provided by the developers of the Quick-REST quantitative measure. The syntax reverse coded the appropriate responses and provided a score for each scenario, BP and FL, then averaged for a final score. The *Mann-Whitney U* determined a difference in Quick-REST scores between the comparison groups, V and S.

Responses to the performance-based assessments were analyzed as recommended by Creswell (2003), notes were written in the margins to "chunk" the topics then coded to distinguish the categories. Responses were coded and grouped utilizing the scoring categories from the REST as "prior themes" (Miles & Huberman, 1994) and were scored a 1 – 3 Ethical Sensitivity Score (EES), 1 being the lowest EES recognition score. Scoring categories of the REST were contingent on recognizing behaviors that violated ethical principles based on professional codes. Table 2 provides six ethical principles based on professional codes of ethics used to develop the REST and have implications of working with individuals from diverse populations. A wide range of behaviors represented the violations. For example in the FL, a teacher portrayed sentiments showing a lack of cultural awareness of his students thus depicting a violation of the ethical principles; professional competence and concern for other's welfare. Additional behaviors revealed the teacher also lacked self-awareness which, violated ethical principles; integrity, respect for other's rights, and dignity (Brabeck et al., 2000). Naming the violated ethical principle was not the goal of the assessment. Recognition of the behavior and consideration of the associated complexities were the basis of the scoring categories.

Table 2. Ethical principles based on professional codes

Ethical Principle	Description
Professional Competence	Conduct which brings credit to one's profession, including cultural competence
Integrity	Self-knowledge of professional values, needs, limitations, and the effect of these on one's work
Professional and Scientific Responsibility	Ethical collaboration with other professionals and holding each to ethical and professional standards
Respect for Other's Rights and Dignity	Treating others with respect, guarding confidentiality, awareness of individual, cultural and role differences
Concern for Other's Welfare	Welfare of student is paramount, knowing impact of adverse social, political, and environmental factors
Social Responsibility	Helping others to understand the extent racism can cause suffering and working to improve social policy

Three post-game prompts were analyzed for emergent themes related to the influence of the SEG on participant responses. All responses were read in a recursive manner and organized for further analysis and interpretation. The responses were transformed into frequency counts, called "quantizing" by Miles and Huberman (1994). This facilitated the recognition of patterns and allowed for statistical display of the data.

Results

Survey responses. Participants of the V group had a mean rank score that was significantly higher, $p = .01$, than the S comparison group, suggesting the V group was more sensitive to racial and gender related ethical violations. A comparison determined if a pattern emerged among the types of questions the V group showed a higher sensitivity for than the S group participants did. A cross-analysis between Quick-REST scores and demographics of the participants showed no discernible pattern emerged in either case. Means, standard deviations, and mean ranks from the *Mann-Whitney U* displayed in Table 3.

Table 3. *Mann-Whitney U* for Quick-REST survey

	N	M	SD	Mean Rank	Sum of Ranks	U	P
V	18	4.27	.39	20.14	362.50		
S	14			11.82	165.50	60.50	.01*

*p<.05, two-tailed.

Performance-based assessments. Scoring categories in Tables 4A-D were contingent on recognizing and understanding the complexity of behaviors in the BP and FL scenarios that violated ethical principles based on professional codes. Table 4E was based on recognizing a major ethical violation. The number of participants were 13 instead of 14 due to a lost of data during a server upgrade.

Table 4A. First BP question

Performance-based assessment	N	EES 1 No recognition	EES 2 Recognition	EES 3 Recognition & Complexity
James is very upset and accuses you and administrators of knowing about Coach Nichols, how would you respond?	13	85%	15%	0

The majority of the participant responses did not demonstrate recognition of James' feelings. These responses requested further information without acknowledging James' anger. It is important to note the question told participants how James was feeling. This was also apparent in his tone of voice, his words, and actions depicted in the SEG. The remainder of the responses, 15%, acknowledged James' feelings and requested more information, as evident by this quote, "I am not sure I do know James. Can you share with me what is upsetting you about Coach Nichols?" No responses demonstrated understanding the complexity of biases inherent in the relationship between James, Coach Nicolas and the school administrators nor that the biases might be related to James' feelings.

Table 4B. Second BP question

Performance-based assessment	N	EES 1 No recognition	EES 2 Recognition	EES 3 Recognition & Complexity
How would you counsel James? Write what you would say to him.	13	38%	0	62%

This question presented an opportunity to state a plan of action and/or demonstrate an understanding of the complexity of the violations. Approximately two-thirds articulated a plan for counseling James and recognized a need to confront the Coach's behavior. Yet the remaining third justified the Coach's actions and comments by saying, "I understand that Coach Nichols may come off strong at time[s] but I assure you he has your best interest in mind," and "James, Coach sees a lot of potential in you and knows you are a great player. He is just harder on you because he see[s] how great of a player you are." There was a disparity in the proposed plans, 50% of the plans advised James to talk with the Coach "about insensitive remarks" or "point out your feelings to the Coach and tell him how unfairly he is treating you." Whereas 38% of those who offered a plan that involved the assumed role of the education professional to talk to the Coach on James' behalf, "I would tell James that he is valid in feeling the way he does. It is obvious that Coach Nichols show preferences to other players and makes un[in]appropriate comments. I would reassure James of his abilities and his talent. I will call a meeting." A minority of the plans 13%, suggested, "life lessons" as a way to handle Coach's behavior, "Keep perspective on what is important…don't allow actions of Coach Nichols to interfere with potential. Don't quit; rise above."

Table 4C. First FL question

Performance-based assessment	N	EES 1 No recognition	EES 2 Recognition	EES 3 Recognition & Complexity
How would you respond to the students who made stereotypic remarks?	13	0	31%	69%

The majority, 69%, of the participants classified the remarks as "biased," "prejudiced," "stereotypical," or "racist," and suggested a plan. A typical response was, "There is no place that is appropriate for that type of statement. Using stereotypical remarks like that will not be tolerated in this classroom (the student would be referred to the counselor.)" The remainder of the responses recognized the remarks as "inappropriate" and "those remarks have no place in my school" however they did not go further and suggest a plan or demonstrate understanding of the complexity in the situation.

Table 4D. Second FL question

Performance-based assessment	N	EES 1 No recognition	EES 2 Recognition	EES 3 Recognition & Complexity
As Ms. Highland the new teacher, write what you would say (if anything) in response to Mr. Lynch's comment about Latinas?	13	16%*	84%	0

*All responded similarly, "I wouldn't say anything unless spoken to."

From the perspective of a new teacher, 16% would not respond to the comments about Latinas. The position to "not say anything" taken by several participants needs attention. Note that unknown political climates in school cultures can be a hindrance in recognizing a need for professional judgment. The majority labeled the remarks as "generalizations" or "assumptions," one stated, "I would say that you make an assumption about all Latin American girls based on the actions of one?" However no one acknowledged the complexity or possible impact of Mr. Lynch's comments about Latinas.

Table 4E. Third FL question

Performance-based assessment	N	Recognition of major violations			
As Ms. Highland the new teacher, what, if anything, disturbs you about this conversation?	13	Confidentiality 8%	Cultural awareness 62%	Professional treatment 15%	Self awareness 15%

This question was used to unearth the major violation in the FL scenario, confidentiality (Brabeck et al., 2000). A majority, 62%, of the participants said, teachers in the FL were "…talking so horribly about different races," and they "…have many preconceived notions about Latinas" suggest that cultural awareness was their major disturbance in the scenario. One or more of the characters in the FL scenario violated all six ethical principles (Table 2), however all violations occurred in a public forum thereby violating the student's right to privacy as noted by this quote, "It is disturbing to me that confidential information about a student is being made public."
Post-SEG prompts. Analysis focused on the influence of *Hazelton High at REST* on participant responses. When asked, *what was this experience like for you?* Over 64% described a reflective and enjoyable component in their experience similar to this quote,

> This was a good experience for me. I feel as if cultural sensitivity issues do not come up in education as often as they should because people feel it is a difficult topic to discuss. I was happy to be exposed to the topic through the simulations.

The remainders were a critique of the SEG scenarios however the majority was positive, for example, "I enjoyed the simulation because it was very technological based." Interaction was the focus when asked how the SEG might differ from watching a video. Seventy-one percent thought the interaction was helpful, as identified by these quotes, "…I can walk through the rooms and explore things unlike if I was to sit and watch a video…this was a more active experience for me," and "…I feel like I as at the school and I directly impacted what was happening at the school." For 7% of the participants, the amount of interaction were a distraction or not enough to be helpful, "sometimes you were wondering what to do next." In the final question, *how might these differences affect your responses?* Comments again pointed to reflection. These quotes represent 66% of the responses, "I thought more about what happened and examined my own reactions more closely," and "I might empathize better because I can feel what it's like to be in the room." Twenty percent of the responses reflected on how they reacted during the SEG, it "Made me feel like I was in the room, more a part of the experience," and "Perhaps if I were to watch a video, my responses would be characterized as ones of a "judge", and not necessary as my personal ones."

Watching a video could make your responses bias because you have not been confronted with an actual scenario. By participating in a simulation, I felt more motivated to give the response I think I would say, not just what I know I should say.

Re-evaluation

In response to the above comments, there was a re-examination of participant responses. The total number of responses was analyzed for emergent themes versus separating them by question as in the earlier analysis. In addition, the above comments enticed the researchers to consider the SEG as providing more realistic performance outcomes possibly because they were immersed in the SEG. Two researchers independently read all responses and organized them for coding. An initial face-to-face meeting to compare results and determine interrater agreement showed coding themes was almost identical and rater agreement was 88% for 42 responses. Discrepancies were discussed and final interrater agreement was 100%.

Emergent themes were: (a) no difference, (b) technology or user issues, (c) like video better, (d) immersion, realistic and/or interaction (e) enjoyed, (f) reflective, and (g) Hawthorne. The Hawthorne theme was responses that noted being a part of an assessment, for example, "…less involvement might skew my extreme responses (strongly agree/strongly disagree) to a milder response." Out of 42 responses the following four themes received the highest frequencies; (a) interaction 36%, (b) reflection 33%, (c) enjoyed 21%, and (d) Hawthorne 19%.

Discussion and Implications

The V group had a significantly higher mean rank Quick-REST score than the S group suggesting the V group was more ethically sensitive at the time of the investigation. Overall, analysis of performance-based assessment questions indicated no recognition of ethical violations was the most frequent response given by the S group. Recognition of ethical violations and showing complexity of the issue was the second most frequent response and recognition of ethical violations was the least frequent response. These findings suggest the SEG provides valuable information concerning candidate disposition when creating equitable learning environments for learners from diverse multicultural and multilinguistic backgrounds.

It was hypothesized that an immersive SEG might affect perspective taking of the participant and indirectly influence ethical sensitivity. Evidence from this study indicates some perspective taking occurred, however little evidence shows an effect on ethical sensitivity. What does seem evident from this study is the use of the SEG as a diagnostic or training tool during the professional education program. Studies have shown ethical sensitivity skills are not improved through conventional professional education (Bebeau & Monson, 2008), yet can be with specific ethical skills education related to the profession (Bebeau & Monson, 2008; Bebeau et al., 1985; Clarkeburn, 2002). The SEG required participants to produce a response instead of choosing a response. This allows analysis of problem-solving skills and reveals individual areas of needed remediation.

Reactions to the SEG indicated an enjoyable, reflective environment that was interactive and allowed some participants to respond more realistically. Fifty-seven percent of the participants made both reflective and immersive comments as apparent by this quote, "it proposes interesting questions and situations that are complicated and thought provoking" and "the interaction makes you feel part of the situation...makes you feel there in the school itself." "Giving responses within the game allows for real time responses and feelings in the moment," suggests there is interplay between the immersive qualities and the reflective nature of the virtual environment affecting the ability to provide responses that are more realistic. As these quotes indicate, the SEG was more than intermittent questions strategically placed. It was the SEG in total, the context, immersive qualities, interaction, and producing responses in the moment to which the participants were responding. The relationship between reflection and immersion suggests the virtual environment revealed realistic performance outcomes, in effect mitigating the Hawthorne effect for some participants. However, seven percent of the responses stated "frustration" or an inability to "know what to do next." Lee, Plass, and Homer (2006) cautions designers and remind them to consider the cognitive load requirements necessary for processing images and concepts not visual to the human eye (Reigeluth & Schwartz, 1989) especially for the novice gamer. This type of inefficacy and possible deterrent to performance-assessment is an area of future exploration. Strides in determining how simulated environments affect learning are difficult and further research in this area is needed (Baek, 2009).

The use of Serious Educational Games (Annetta, 2008) for professional performance-based assessments is a new frontier in education following in the footsteps of the military and medical fields. Science education programs and researchers have the ability to track participant actions and note patterns which maybe instrumental in determining skill level, areas of needed remediation or refinement of skills. This investigation merges specific characteristics designed into *Hazleton High at REST* while pursuing realistic performance outcomes. Equally important, this study explores science teacher ethical sensitivity for racial and gender intolerant behavior in schools where unconscious biases can be examined. The aim was to gain insight into the problem solving skills of preservice science teachers when confronted with racial and gender intolerant behavior in a classroom. The scenarios were a close approximation of what the participant might encounter in real life in a risk-free computer simulated environment. The responses in the role-playing situations allowed participants to reflect and apply their knowledge in an appropriate manner to authentic information. These attributes are several of the basic design features for simulations noted for effectiveness by Magee (2006).

Recommendations for Future Research and Conclusion

Future researchers should analyze preservice teacher behavior using a mixture of surveys and performance-based assessments in an immersive and engaging environment that mirrors real life teaching settings. Furthermore, researchers should also examine challenges in user-interface, usability, and cognitive-load experienced by those who utilize SEGs and other performance based assessments. Even with known challenges, a well-designed SEG has the potential to improve participant skills and critical thinking in a complex authentic context while moving beyond multiple-choice assessments.

Set forth by science education professional standards, science educators are to develop and demonstrate knowledge of diverse cultures, relate science to the stakeholders in the community, value cultural heritage, and respect learner worldview when selecting and developing lessons (AAAS, 1989; CAEP 2012; NSTA 2012). Demonstrating these responsibilities may appear daunting in today's multicultural and multilinguistic learning environments. Expressing cultural knowledge of K-12 learners and respecting their cultural heritage and identity can be challenging due to unfamiliarity with cultural backgrounds, unsubstantiated beliefs about the other, and the rapid change in learner demographics (Welton & Martinez, 2014). A risk-free SEG or simulation may allow novice teachers to feel a part of the environment and perform a desired skill that closely matches their true capabilities and promote culturally-responsive science education.

Acknowledgements and Notes

References

AAAS. (1989). *Science for all Americans*. New York: Oxford University Press. Retrieved from http://www.project2061.org/publications/sfaa/online/sfaatoc.htm

Abrahamson, S., Denson, J. S., & Wolf, R. M. (2004). Effectiveness of a simulator in training anesthesiology residents. *Quality and Safety in Health Care*, *13*(5), 395-397. doi:10.1136/qhc.13.5.395

Adair, J. G. (1984). The Hawthorne effect: A reconsideration of the methodological artifact. *Journal of Applied Psychology*, *69*(2), 334.

Annetta, L., & Holmes, S. (2006). Creating presence and community in a synchronous virtual learning environment using avatars. *International Journal of Instructional Technology and Distance Learning*.

Annetta, L., Minogue, J., Holmes, S., Cheng, M., Folta, E., & Klesath, M. (2009). Digital simulations for Improving Education: Learning Through Artificial Teaching Environments. In *Using Case Studies as the Narrative to Game Design and Development*. Hershey, PA: IGI Global.

Annetta, L. (2008). *Serious educational games: From theory to practice.* The Netherlands: Sense Publishers.

Baek, Y. (2009). Digital simulation in teaching and learning. In *Digital simulations for improving education: Learning through artificial teaching environments*. Information Science Reference.

Bebeau, M. J., & Monson, V. E. (2008). Guided by theory, grounded in evidence: A way forward for professional ethics education. In *Handbook of moral and character education* (pp. 557-582). New

York: Routledge.

Bebeau, M. J., Rest, J. R., & Yamoor, C. M. (1985). Measuring dental students' ethical sensitivity. *Journal of Dental Education*, *49*(4), 225-235.

Bos, N. D., Shami, N. S., & Naab, S. (2006). A globalization simulation to teach corporate social responsibility: Design features and analysis of student reasoning. *Simulation & Gaming*, *37*(1), 56-72.

Brabeck, M. M., Rogers, L. A., Sirin, S., Henderson, J., Benvenuto, M., Weaver, M., & Ting, K. (2000). Increasing ethical sensitivity to racial and gender intolerance in schools: Development of the racial ethical sensitivity test. *Ethics & Behavior*, *10*(2), 119.

Brown, B. A. (2006). "It isn't no slang that can be said about this stuff": Language, identity, and appropriating science discourse. *Journal of Research in Science Teaching*, *43*(1), 96-126. doi:10.1002/tea.20096

Bryan, L. A., & Atwater, M. M. (2002). Teacher Beliefs and Cultural Models: A Challenge for Science Teacher Preparation Programs. *Science Teacher Education*, *86*(6), 821-839.

Buzzelli, C., & Johnston, B. (2002). *The moral dimensions of teaching: Language, power, and culture in classroom interaction*. London: RoutledgeFalmer.

Carter, N. P., Larke, P. J., Singleton-Taylor, G., & Santos, E. (2003). Multicultural science education: Moving beyond tradition. *Counterpoints, 120*, 1-19.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company.

Census, U. S. (2008). *An Older and More Diverse Nation by Midcentury*. Washington, DC. Retrieved from http://www.census.gov/Press-Release/www/releases/archives/population/012496.htm

Chiesa, M., & Hobbs, S. (2008). Making sense of social research: How useful is the Hawthorne Effect? *European Journal of Social Psychology*, *38*(1), 67-74.

Clarkeburn, H. (2002). A test for ethical sensitivity in science. *Journal of Moral Education*, *31*(4), 439-453.

Council for the Accreditation of Educator Preparation. (2012). *Standard 3: Candidate Quality, Recruitment, and Selectivity*. Retrieved from www.ncate.org/standards/standard-3

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed-methods approaches* (Vol. 2). Thousand Oaks: Sage.

Dede, C. (2005). Planning for neomillennial learning styles. *Educause*. Retrieved from www.educause.edu/apps/eq/eqm05/eqmo511

Dotger, B. H., Harris, S., & Hansel, A. (2008). Emerging authenticity: the crafting of simulated parent–teacher candidate conferences. *Teaching Education*, *19*(4), 337. doi:10.1080/10476210802438324

Eck, R. V. (2006). Digital game-based learning. *EDUCAUSE Review*, *41*, 17-30.

Fenstermacher, G. (1990). Some moral considerations on teaching as a profession. In J. Goodlad, R. Soder, & K. Sirotnik (Eds.), *The Moral Dimension of Teaching* (pp. 130 - 151). San Francisco: Jossey-Bass.

Fraser-Abder, P. (2001). Preparing Science Teachers for Culturally Diverse Classrooms. *Journal of Science Teacher Education*, *12*(2), 123.

Frenz, C. (2007). Possibilities and limitations of computer simulation. *IEEE Potentials*, (March/April), 30 - 33.

Gee, J. P. (2005). Learning by design: good video games as learning machines. *E-Learning*, *2*(1), 5-16.

Gee, J. (2003). *What video games have to teach us about learning*. New York: Palgrave.

Hmelo, C. E., Ramakrishnan, S., Day, R. S., Shirey, W. E., Brufsky, A., Johnson, C., Baar, J., et al. (2001). Oncology Thinking Cap: Scaffolded Use of a Simulation to Learn Clinical Trial Design. *Teaching & Learning in Medicine*, *13*(3), 183-191.

Holden, J. (2000). Hawthorne effects and research into professional practice. *Journal of Evaluation on Clinical Practice*, *7*(1), 65-70.

Jones, S. R. G. (1992). Was there a Hawthorne effect? *American Journal of Sociology*, *98*(3), 451-468.

Lee, H., Plass, J. L., & Homer, B. D. (2006). Optimizing Cognitive Load for Learning From Computer-Based Science Simulations. *Journal of educational psychology*, *98*(4), 902-913.

Lemke, J. L. (2001). Articulating Communities: Sociocultural Perspectives on Science Education. *Journal of Research in Science Teaching*, *38*(3), 296-316.

Macedonia, M. (2002). Games soldiers play. *IEEE Spectrum*, *39*(3), 32-37.

Magee, M. (2006). State of field review: Simulation in Education. Retrieved from www.ccl-cca.ca/NR/.../SFRSimulationinEducationJul06REV.pdf

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (Vol. 2). Thousand Oaks, CA: Sage Publications.

Montero-Sieburth, M. (1996). Teachers', administrators', and staff's implicit thinking about "at-risk" urban high school Latino students. In *Teacher thinking in cultural contexts*. Albany, NY: State University of New York Press.

National Science Teachers Association. (2010). *NSTA Position statement: Principles of professionalism for science educators*. Retrieved from http://www.nsta.org/about/positions/professionalism.aspx

Payne, G., & Payne, J. (2004). *Key concepts in social research*. London: Sage Publications. Retrieved from

http://search.trln.org/search?id=UNCb4466094

Peacock, B. (2005). Laws and rules: Remember Hawthorne. *Ergonomics in Design*, *13*(1), 6.

Reigeluth, C., & Schwartz, E. (1989). An Instructional Theory for the Design of Computer-based Simulations. *Journal of Computer-Based Instruction*, *16*(1), 1-10.

Rest, J. R., & Narvaez, D. (1994). *Moral development in the professions: psychology and applied ethics*. Hillsdale, NJ: Lawrence Erlbaum.

Rest, J. (1986). *Moral development: Advances in research and theory.* New York: Praeger.

Selman, R. L. (1977). A structural-developmental model of social cognition: Implications for intervention research. *Counseling Psychologist*, *6*(4), 3-6.

Selman, R. L. (2003). *The promotion of social awareness: Powerful lessons from the partnership of developmental theory and classroom practice*. New York: Russell Sage Foundation.

Shaffer, D. (2007). Epistemic frames for epistemic games. *Computers & Education*, *46*(3), 223.

Shaffer, D. W., Squire, K. R., Halverson, R., & Gee, J. P. (2005). Video games and the future of learning. *Phi Delta Kappan*, *87*(2), 105-111.

Sirin, S., Rogers-Sirin, L., & Collins, B. A. (2010). A measure of cultural competence as an ethical responsibility: Quick-Racial and ethical sensitivity test. *Journal of Moral Education*, *39*(1), 49-64.

Sirin, S. R., Brabeck, M. M., Satiani, A., & Rogers-Sirin, L. (2003). Validation of a measure of ethical sensitivity and examination of the effects of previous multicultural and ethics courses on ethical sensitivity. *Ethics & Behavior*, *13*(3), 221.

Soderberg, P., & Price, F. (2003). An Examination of Problem-Based Teaching and Learning in Population Genetics and Evolution Using EVOLVE, a Computer Simulation. *International Journal of Science Education*, *25*(1), 35-55.

Tung, F., & Deng, Y. (2006). Designing social presence in e-learning environments: Testing the effect of interactivity on children. *Interactive Learning Environments*, *14*(3), 251-264.

Villegas, A. M., & Lucas, T. (2002). Preparing Culturally Responsive Teachers: Rethinking the Curriculum. *Journal of Teacher Education*, *53*(1), 20-32.

Waldmann, U., Gulich, M. S., & Zeitler, H. (2008). Virtual patients for assessing medical students-important aspects when considering the introduction of a new assessment format. *Medical teacher*, *30*(1), 17-24.

Welton, A. D., & Martinez, M. A. (2014). Coloring the college pathway: A more culturally responsive approach to college readiness and access for students of color in secondary schools. *The Urban Review, 46*(2), 197-223. doi:10.1007/s11256-013-0252-7

Yeh, Y. C. (2004). Nurturing reflective teaching during critical-thinking instruction in a computer simulation program. *Computers & education*, *42*(2), 181.

Author Information

Shawn Y. Holmes
The You in You Consulting, Inc.
1-2321 Island View Road
Victoria, BC V8Z 0C9, Canada.
Contact e-mail: *shawn@theuiun.com*

Loni Crumb
East Carolina University
213 Ragsdale Hall, MS 121
Greenville, NC 27858, U.S.A.

Leonard Annetta
East Carolina University
Flanagan 307, MS 566
Greenville, NC 27858, U.S.A.