Jl. of Technology and Teacher Education (2018) **26**(2), 299-326

Early Career Teacher Professional Development: Bridging Simulation Technology with Evidence-Based Behavior Management

ELISA SHERNOFF
Rutgers University, USA
elisa.shernoff@gsapp.rutgers.edu

STACY FRAZIER
Florida International University, USA
slfrazi@fiu.edu

CHRISTINE LISETTI
Florida International University, USA
christine.lisetti@gmail.com

CEDRIC BUCHE
Centre Européen de Réalité Virtuelle, France
buche@enib.fr

STEPHANIE LUNN
Florida International University, USA
slunn002@fiu.edu

CLAIRE BROWN
Rutgers University, USA
claire.l.brown@rutgers.edu

ALBAN DELMARRE
Florida International University, USA
adela177@fiu.edu

TOMMY CHOU
Florida International University, USA
pchou006@fiu.edu

JOSEPH GABBARD
Virginia Tech, USA
jgabbard@vt.edu

EMILY MORGAN
Rutgers University, USA
emily.morgan@rutgers.edu

Early career teachers working in high poverty schools face overwhelming challenges navigating disruptive behaviors with studies highlighting behavior problems as one of the strongest predictors of turnover (Ingersoll & Smith, 2003). Simulation-based technology leverages important pedagogical strengths (e.g., realistic training context, opportunities for practice and feedback, low-stakes training) that may confer advantage over traditional training for teachers in behavior management. This manuscript illustrates the development and refinement of Interactive Virtual Training for Teachers (IVT-T), a simulation-training model to help early career teachers hone their behavior management skills by interacting with virtual students in a 3D classroom. Advisory board members (*N*=5) with extensive experience in high poverty schools provided iterative feedback on the authenticity and realism of IVT-T students, classrooms, and vignettes. Advisory board members rated authenticity/realism of the students, classrooms, and vignettes favorably and ratings also increased over time as new prototypes were shared. In addition, prototypes introduced earlier underwent more revisions than prototypes introduced later suggesting advisory board feedback helped the computer science team improve upon their development approach. Open-ended feedback revealed important shortcomings and the need for further refinements to improve the training model. Implications for teacher educators and researchers involved in the development of instructional technologies are discussed.

Early career teachers working in high poverty schools face significant challenges related to preventing and managing challenging behaviors. Studies document that early career teachers (with four or fewer years of experience) are “shocked” by persistent non-compliance and report a troubling discrepancy between their expectations regarding teaching and the stark reality of a complex classroom (Shernoff, Lakind, Frazier, & Jakobsons, 2015; Shernoff et al., 2016; Veenman, 1984). Teacher effectiveness in behavior management can mitigate student behavior problems by improving instruction and increasing opportunities for learning (Creemers, 1994; Crone & Teddlie, 1995; Oliver & Reschly, 2007). Thus, training and support for early career teachers working in high poverty schools around managing disruptive behaviors represents a critical step toward facilitating successful school trajectories for students.

Limitations to Traditional Professional Development in Behavior Management

Behavior management encompasses a complex set of skills related to preventing and responding to challenging behaviors. Prevention includes proactive monitoring, establishing clear rules and expectations, and delivering engaging instruction. Responding to challenging behaviors involves non-punitive, non-reactive approaches and helping teachers develop skills in differential attention -- knowing when to attend and when to ignore speci c behaviors using principles of operant conditioning (Kazdin, 2005). Developing early career teachers' knowledge and skills in evidence-based behavior management strategies is a critical professional development goal given the empirical link between effective behavior management and student academic outcomes (Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016).

Despite the critical role that effective behavior management plays in student engagement and learning, traditional professional development (e.g., workshops, whole group inservice training) has failed to produce substantial improvements in teachers' work-related knowledge and skills (Penuel, Fishman, Yamaguchi & Gallagher, 2007). This situation is exacerbated for early career teachers who already receive limited mentoring in behavior management (Grossman & McDonald, 2008) and then must acquire these skills on-the-job, with real students while delivering instruction. Fast-paced, high-stakes, live instruction leaves little time for practice with feedback, which can be costly to teachers and learners (Henry, Farrell, Schoeny, Tolan, & Dymnicki, 2011; Schussler, Frank, Lee & Mahfouz, 2017).

It comes as no surprise that early career teachers rank disruptive behavior as a signi cant stressor and their top ranked reason for initiating transfer requests or leaving the eld of teaching (Gonzalez, Brown & Slate, 2008; Ingersoll & Smith, 2003). The inadequate impact of behavior management training on teacher knowledge and skill has been linked to the time-limited nature of training that leaves few opportunities for re ection and feedback to enhance future performance (Penuel et al., 2007). In this manuscript, we introduce Interactive Virtual Training for Teachers (IVT-T), a professional development model designed to help early career teachers working in high poverty schools to hone their behavior management skills by interacting with virtual disruptive students in an interactive virtual classroom (Magnenat-Thalmann & Thalmann, 2005). The goals of the manuscript are twofold. First, we describe the process of developing and re ning IVT-T and illustrate the interdisciplinary collaborations required to integrate technology with teacher training in evidence-based practices (Girard et al., 2012; Hartson & Pyla, 2012; Sitzman, 2011; Vogel et al., 2006). Second, we evaluate key design components such as logic, authenticity, and realism of IVT-T students, classrooms, and vignettes and explore how much end users could improve upon these key design components by providing iterative feedback.

Supporting Teachers' Professional Development via Simulation Training

Technology has long been used to train a range of professionals (e.g., surgeons, pilots, rst responders) to make high-stakes decisions for which errors are costly. Simulation training delivered via personal computers or mobile devices immerse trainees in decision-making exercises in a 3D environment to explore their options and consider the consequences of their decisions (Regalla, Hutchinson, Nutta & Ashtari, 2016; Sitzman, 2011). These training models leverage the interactive nature of video games to improve work-related skills and knowledge (Schrader, Archambault & Oh-Young, 2011; Sitzman, 2011; Vogel et al., 2006). Simulation training models, including intelligent tutoring systems (designed to provide guidance and support to learn new skills where thinking is made visible) and serious games (designed to immerse learners in a playful environment that also fosters learning and targeted skills) are newer to teacher education (Regalla et al., 2016; Schussler, Frank, Lee, & Mahfouz, 2017) but hold promise for supporting teachers' professional development (Shernoff et al. 2016). These models are, however, highly individualized and emerging at a rate that outpaces research on the methodology, procedures, and tools required to evaluate them (Alvarez & Michaud, 2008; Arnab et al., 2015; Jeremic, Jovanovic & Gaševic, 2009; Schrader et al., 2011). Lack of standardized measures to evaluate instructional design features leaves gaps in the literature regarding how to assess authenticity, realism, and immersion (Hartson & Pyla, 2012; Girard, Escalle, & Magnan, 2012; Pallavincini, Toniazzi, Argenton, Aceti,

& Mantovani, 2015). Four elements of simulation training described below may confer advantages over traditional training models, in particular for supporting early career teachers working in high poverty schools.

Promotes realistic work scenarios. Simulation training builds on studies documenting that learning and transfer are maximized when users confront realistic work scenarios and make decisions that are relevant to their job performance (Bellotti, Berta & De Gloria, 2010; Sitzman, 2011; Tawadrous, Kevan, Kapralos & Hogue, 2012). A recent meta-analysis comparing the instructional effectiveness of simulation games to alternative training methods (Sitzmann, 2011) revealed that user confidence in performing training-related tasks was 20% higher, knowledge regarding how to perform specific actions was 14% higher, immediate memory for facts and principles was 11% higher, and recall was 9% higher in the simulation training condition versus comparison (alternative instructional method or no-training control). For early career teachers, access to an authentic, realistic virtual training experience may enhance their knowledge of how to respond to disruptive behaviors by observing the consequences of their decisions (Amory, 2007; Belloti et al., 2010; Gee, 2005).

Maximizes practice and feedback. Simulation training can also provide extensive opportunities for practice with feedback to develop users' knowledge and skills. Using technology to support learning can help users differentiate between correct and incorrect decisions and reconcile misconceptions regarding the best course of action (Heeren & Jeuring, 2014). For early career teachers, learning how to effectively prevent and respond to disruptive behaviors is a complex skill that improves over time (Denton & Hasbrouck, 2009; Sabers, Cushing & Berliner, 1991). Traditional professional development for new teachers may also underestimate the time required to develop and hone these skills. For instance, a recent study of professional development to support new teachers highlighted that 50% of the early career teacher sample required nearly two years to cultivate these skills (Shernoff et al., 2016).

Provides low-stakes training. When the stakes associated with making errors are low, deliberate decision-making may require less time, practice, and cognitive effort (Dunbar et al., 2014). However, when the cost associated with making errors is high, simulations may be uniquely advantageous due to the practical and ethical advantage of encouraging freedom to explore and fail in a low-stakes setting. Examples include training pilots to navigate challenging flying conditions (Fletcher, 2009), training novice clinicians to conduct assessments for individuals at risk for suicide (Beutler & Harwood, 2004; Horswill & Lisetti, 2011), helping teacher candidates identify and respond to student bullying (Schussler et al., 2017), helping children to respond to bullies (Vannini, et al., 2011), and helping first responders make decisions during emergencies (Mantovani, Castelnuovo, Gaggioli, & Riva, 2003). As the achievement gap widens, there is a pressing need for early career teachers working in high poverty schools - where students are at greatest risk for academic failure - to develop their behavior management skills in a compelling but, low-consequence-for-failure environment. This is in contrast to developing these skills on-the-job, with real students, for whom cumulative teacher errors may have long-term consequences.

Enhances access. Improving teachers' behavior management skills is among the most important predictors of student learning (Atkins et al., 2015). Simulation, unlike individual consultation or coaching with early career teachers, provides an unprecedented opportunity to reach large numbers of teachers in need of support (Kazdin & Blasé, 2011; Kazdin, 2015; Schussler et al., 2017). Recently, avatars have been used to develop preservice teachers' classroom management skills (Dieker, Hynes, Hughes, & Smith, 2008; Dieker, Hynes, Stapleton, & Hughes, 2007) in a mixed-reality environment called TeachLivE. Dieker and colleagues' groundbreaking work exposes teachers to disruptive avatars via a projection screen, with urban teachers in particular reporting a realistic and compelling training experience (Dieker et al., 2007). However, avatars are controlled by actors using high-end laboratory equipment which may limit its reach. IVT-T, on the other hand, will be programmed such that avatars will be endowed with behaviors that can be activated by the system itself, eliminating the need for actors or equipment and widening potential reach. IVT-T will also be web-based, which allows users to access the system from a range of locations using a variety of computing systems. The web-based support may also circumvent significant time barriers associated with teachers participating in live coaching during the instructional day (Shernoff et al., 2015; Shernoff, Lekwa, Reddy, & Coccaro 2017).

Bridging Simulation Technology with Research on Evidence-Based Behavior Management

Given the urgent needs of early career teachers working in high poverty schools, IVT-T is being developed, via a four-year IES-funded Goal 2 award, to provide new teachers with extended opportunities to improve their behavior management skills through virtual learning opportunities (ies.ed.gov/ncer/projects/grant.asp?ProgID=21&grantid=1725&NameID=258).

Years 1 and 2 are devoted to developing and iteratively re ning IVT-T while Years 3 and 4 focus on examining whether the virtual training improves student engagement and achievement (Creemers, 1994; Crone & Teddlie, 1995; Oliver & Reschly, 2007).

To date, the instructional potential of merging simulation technology with evidence-based behavior management strategies has been limited by the absence of standardized methods and measures to evaluate and improve 3D systems with regard to physical appearance and instructional content (Alvarez & Michaud, 2008; Arnab et al., 2015; Jeremic, Jovanovic & Gaševic, 2009; Schrader et al., 2011). Therefore, the focus of the current manuscript is twofold. First, we describe the process of developing and re- ning IVT-T to highlight opportunities for interdisciplinary teams of scientists to advance the science on integrating technology with teacher education (Girard et al., 2012; Hartson & Pyla, 2012; Sitzman, 2011; Vogel et al., 2006). Second, we present the iterative process and outcome by which feedback improved key design features of IVT-T (Hartson & Pyla, 2012). Research questions included: (1) Did the characters, classrooms, and vignette logic, authenticity, and realism change over time to meet a minimum standard on these design features? and, (2) Did feedback from experienced educators improve the 3D characters, 3D classrooms, and vignettes?

METHOD

Advisory Board Members

Although IVT-T end users are early career teachers, their limited experience in high poverty schools restricted the type of feedback they could provide regarding authenticity, realism, and representativeness. Therefore, ve retired educators, who have collaborated with the research team in prior grants, served on an advisory board during Year 1. Advisory board members had extensive teaching experience in high poverty schools (Mean = 31.8 years, SD = 13, Range = 11 to 35). Four were female, all had attained Master's Degrees in Education, four self-identi ed as European American and one self-identi ed as African American.

IVT-T Design Components to Support Teacher Professional Development

Students. The computer science team developed and re ned thirty 3D virtual students who were racially and ethnically diverse re ecting demographics in high poverty schools. Students were assigned a range of skin coloring, hairstyles, body shapes and sizes, and were dressed in standard uniforms commonly worn in K-8 schools in high poverty communities. Table 1 illustrates the age, gender, and racial/ethnic demographic characteristics of the final set of virtual students. Four primary virtual students are the main protagonists whose behaviors will be programmed (during Years 2 and 3) to interact with the early career teacher as they traverse various vignette paths: two off-task students (i.e., 1st and 6th grade) and two aggressive/non-compliant students (i.e., 1st and 6th grade). The behaviors of the remaining twenty-six non-disruptive students (13 per classroom) will be programmed to cycle through a set of pre-defined behaviors (e.g., reading, turning a page in a book).

Table 1

Demographic Characteristics of the Virtual Students

1st Grade Students (N= 15)	
Gender	7 Female 8 Male
Race/Ethnicity	7 African American 8 Latino/a
Primary Virtual Students	1 Off-task (African American Male) 1 Aggressive/Non-compliant (Latina)
6th Grade Students (N=15)	
Gender	9 Female 6 Male
Race/Ethnicity	6 African American 9 Latino/a
Primary Virtual Students	1 Off-task (African American Female) 1 Aggressive/Non-compliant (8 Latino)

Note. 1st grade students (13 non-disruptive, 2 disruptive) and 6th grade students (13 non-disruptive, 2 disruptive).

Classrooms. Two classrooms (i.e., one 1st grade and one 6th grade) were developed and refined with key considerations given to object placement within each classroom, size, organization, textures, and lighting. Given the high poverty context, close attention was also paid to depicting a resource-limited classroom environment, including desks and books that appeared worn, floors that appeared scratched, and limited supplies.

Vignettes. Four vignettes (one for each primary student) were developed in teams led by the rst and second author. These hypothetical scenarios were designed to simulate provocative teacher-student interactions common to K-8 classrooms (Junod, DuPaul, Jitendra, Volpe, & Cleary, 2006; Paschall, Fishbein, Hubal & Eldreth, 2005). Teacher response options were guided by evidence-based strategies (i.e., praise, ignore, redirect, use of proximity, instructions, empathy, if/then statements) designed to mitigate disruptive behaviors and improve attention, compliance, and engagement (Evertson & Weinstein, 2006; Junod et al., 2006; Kazdin, 2005; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008). Vignettes progress such that vir tual students escalate (become more off-task and aggressive) or de-escalate (become more engaged and compliant) contingent on teachers' responses. This simulated interaction mimics the antecedent-behavior-consequence cycle of disruptive behaviors in real classrooms and re ects the in uence of the environment in maintaining them (Kazdin, 2005).

Figure 1 depicts a sequence of interactions between one virtual student and an early career teacher using Lucidchart (www.lucidchart.com). In this example, "Jordan" (aggressive/non-compliant 6th grader) walks into the classroom late, icks another student on the arm, and drops his books loudly on his desk. At this point, the early career teacher makes his or her rst decision.Option 1: Ignore Jordan, Option 2: Call attention to the fact that Jordan is late again, and Option 3: Welcome Jordan to the classroom and direct him to complete the assignment. Based on each teacher's initial decision, the vignette proceeds, and at each subsequent decision point, teachers receive three options for how to respond.

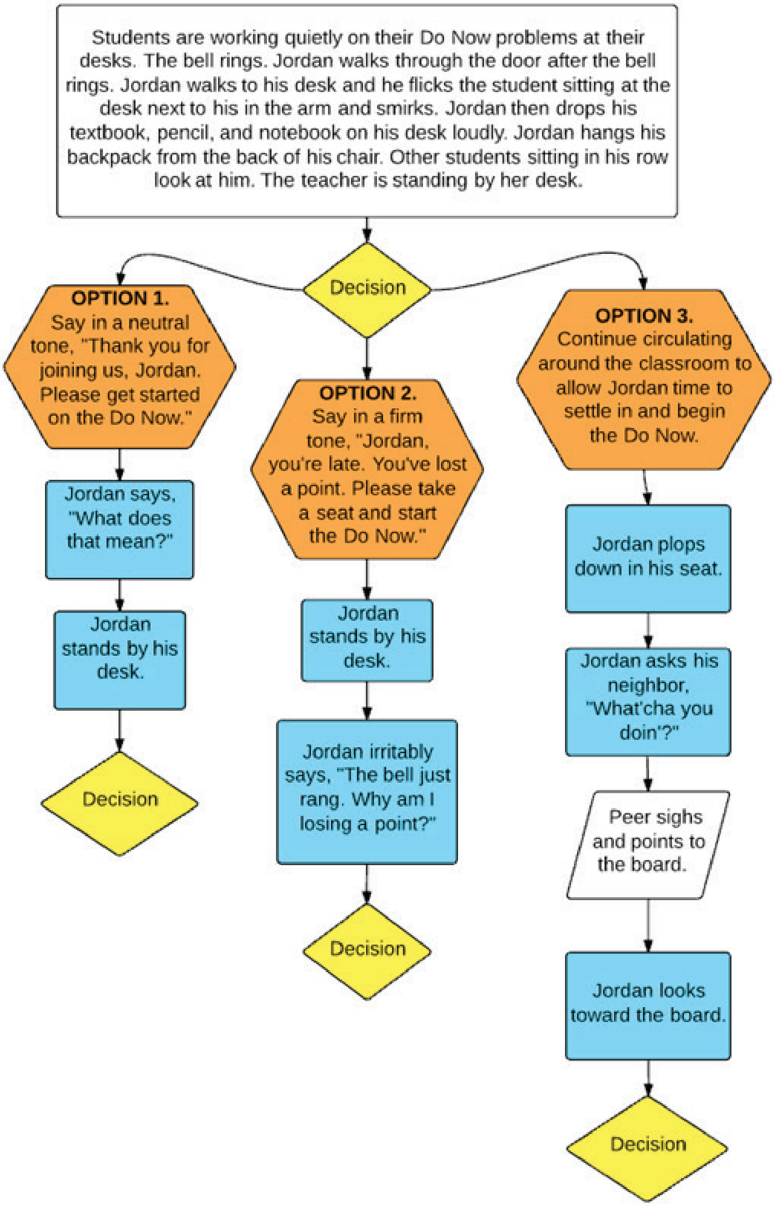

Figure 1. Illustration of a vignette in Lucidchart.

Measures

Three measures were developed to obtain feedback on authenticity/realism of the 3D students and classrooms, and the logic and realism of the vignettes.

Authenticity/realism of students. While viewing student prototypes on a secure website, advisory board members completed a Qualtrics survey in which they provided quantitative and open-ended ratings on the visual quality of the 3D student prototypes. Advisory board members provided feedback on whether the prototype re ected the age and developmental nuances of a 1st grader or 6th grader on four different dimensions: face, body, clothing, and hair. Each dimension was rated on a 4-point scale (1 = Poor, 2 = Fair, 3 = Good, and 4 = Outstanding). Advisory Board Members also provided open-ended feedback regarding features they liked along with recommended changes.

Authenticity/realism of classrooms. While viewing classroom prototypes on a secure website, advisory board members completed a Qualtrics survey in which they provided quantitative and open-ended ratings focused on the visual quality of the 1st and 6th grade classrooms. Advisory board members provided feedback on four different classroom dimensions: (1) Physical Arrangement (e.g., room size, desk arrangement, and placement of furniture, windows, and doors), (2) Wall Décor (e.g., visual quality of bulletin boards, student work, and classroom rules), (3) Materials/Supplies (e.g., student and teacher supplies, books, computers), and (4) Overall Appearance (e.g., lighting, shadows, color). Each dimension was rated on a 4-point scale (1 = Poor, 2 = Fair, 3 = Good, and 4 = Outstanding). In addition, advisory board members provided open-ended feedback regarding what they liked and recommended changes to improve upon the classroom prototype.

Authenticity/realism of vignettes. Advisory board members rated the degree to which they agreed with statements regarding the logic and realism of the storylines. Items focused on the logic and realism of the students' behavior/dialogue based on grade (1st or 6th grade) and presenting problem (off-task or aggressive/non-compliant) on a 4-point scale (1 = Strongly Disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly Agree). Advisory board members also rated the logic and realism of teachers' responses to students based on grade and presenting problem. Advisory board members also rated how much they agreed with the statement that the storylines were engaging. Finally, advisory board members provided open-ended feedback regarding what they liked regarding the vignettes and recommended changes to improve upon the storylines.

Procedures and Analyses

Prototypes of the 3D students and classrooms were first shared internally with the interdisciplinary research team, who provided iterative feedback and made joint decisions regarding changes to the students, classrooms, and vignettes. As more developed prototypes were available, advisory board members received a link to a secure website where screen shots of the students, classrooms, and vignette prototypes were shared.

Quantitative ratings illustrated *how much* advisory board members liked specific design features and alerted the team to unfavorable ratings (e.g., fair or poor ratings on student and classroom dimensions, low ratings of engagement on vignettes) requiring additional revisions. Open-ended feedback highlighted *why* advisory board members liked and disliked particular features along with *specific design changes* to enhance logic, authenticity, and realism. Iterative revisions concluded when all advisory board members rated the students and classrooms as *Good* or *Outstanding* on all dimensions and *Agreed* or *Strongly Agreed* that the students' behavior, actions, and dialogue were logical and realistic based on age and presenting problem. Descriptive analyses compared mean ratings of the student, classroom, and vignette prototypes over time to evaluate the extent to which authenticity/realism iteratively improved to achieve the minimum standard on these design features.

RESULTS

Did Student Authenticity/Realism Change Over Time?

Three groups of virtual students were developed and introduced to the advisory board. Group 1 was introduced in November 2015 and included three 1st grade boys, one 6th grade boy, and one 6th grade girl. Group 2 was introduced in December 2015 and included one 1st grade boy, four 1st grade girls, one 6th grade boy, and two 6th grade girls. Group 3 was introduced in January 2016 and included four 1st grade boys, three 1st grade girls, four 6th grade boys, and six 6th grade girls. The number of iterations by character ranged from one to three (Mean = 1.47; SD =. 54). When evaluating authenticity/realism ratings for students' face, body, clothing, and hair, only two of the 30 characters were rated in the *Fair* range. Advisory board ratings were generally high and also increased from Prototype 1.0 to Prototype 2.0 when comparing students' face (Mean Prototype 1.0 = 3.59, SD = .59; Mean Prototype 2.0 = 3.62, SD = .54), body (Mean Prototype 1.0 = 3.72, SD = .48;

Mean Prototype 2.0 = 3.85, SD = .37), clothing (Mean Prototype 1.0 = 3.66, SD = .48; Mean Prototype 2.0 = 3.67, SD = .48), and hair (Mean Prototype 1.0 = 3.51, SD = .60; Mean Prototype 2.0 = 3.69, SD = .52).

Figure 2 illustrates the physical transformation to clothing, face, hair, and body for a 6th grade student. For example, feedback for Prototype 1.0 (see Figure 2) included: "*His facial features should be softened so he looks more like a preteen. His clothes should be loose fitting and he should wear a uniform.*" This feedback informed revisions by the computer science team which were incorporated into Prototype 2.0, including reducing the overall size of the student and softening his facial features so that he looked more like a 6th grader. In Prototype 3.0, the student was featured wearing a uniform and advisory board feedback was overall very positive, "*This boy looks age appropriate. Face looks like an 11-year-old.*"

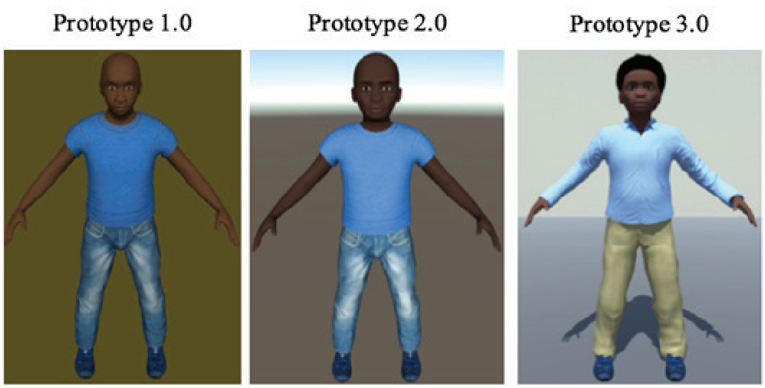

Figure 2. Evolution of a 6th grade male character.

Did Classroom Authenticity/Realism Change Over Time?

The original 1st grade classroom prototypes (developed in September 2015) and 6th grade prototypes (developed in May 2016) were rated, on average, as *Poor* (Mean 1st grade = 1.33; Range = N/A; Mean 6th grade =

2.96, SD = .60). This can be compared to the revised 1st grade prototypes (developed in March 2016) and 6th grade prototypes (developed in November 2016), which were rated, on average, as *Good* (Mean 1st grade = 3.50, SD = .71; Mean 6th grade = 3.56, SD = .59). The original 1st grade classroom prototypes (developed in December 2015 to March 2016) were rated, on average, as *Fair* (Mean Physical Arrangement = 2.20, SD = 1.30; Mean Wall Décor = 2.20, SD = 1.30; Mean Overall Appearance = 2.60, SD = .89) when compared to the revised prototypes (developed in March to November 2016), which were rated, on average, as *Good* (Mean Physical Arrangement = 3.25, SD = .50; Mean Wall Décor = 3.25, SD = .50; Mean Overall Appearance = 3.50, SD = .58). Advisory board ratings for the 6th grade prototype remained, on average, in the *Good* range when comparing the original (developed in May 2016) and revised prototypes (developed in November 2016) for all dimensions except Wall Décor, which increased from an overall rating of *Fair* (Mean = 2.14, SD = .90) to *Good* (Mean = 3.75, SD = .50).

Figure 3 illustrates the evolution of the 1st grade classroom, including the original classroom design which was rated as "*fairly sterile for a 1st grade classroom*." One advisory board member also noted that the teacher's desk was "*too modern*" and out of character for classrooms situated in high poverty communities. The advisory board member noted, "*usually you see wood or metal but beat up*." The revised prototype was described by an advisory board member as mixed, including the need for "*more wall hangings*" but that the "*flooring looks better, more beat up*." Subsequent prototypes were rated as "*much more realistic*" although concerns regarding lighting (i.e., "*not much natural light coming in*") informed subsequent adjustments so that light emanated through the windows as well as the ceiling.

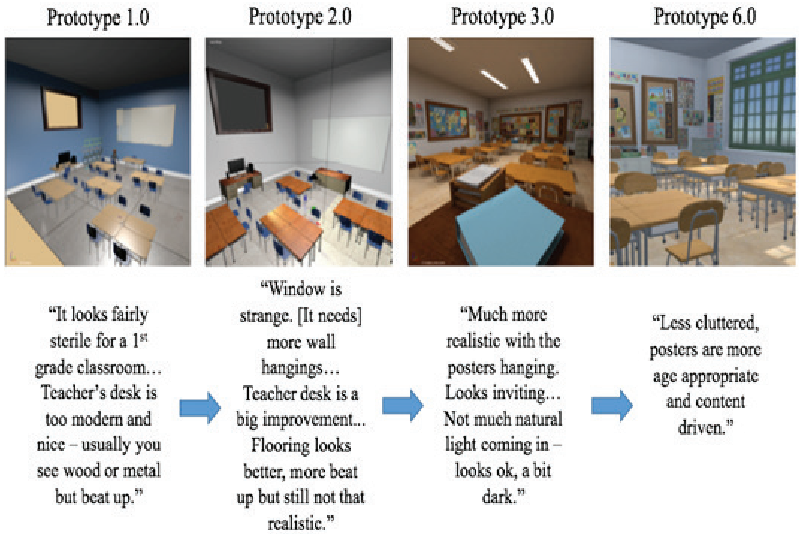

Figure 3. Evolution of the 1st grade classroom.

Did Vignette Authenticity/Realism Change Over Time?

Advisory board members provided feedback on the vignettes five times (see Figure 4) with each advisory board member contributing feedback on each vignette at least once. Original vignettes (developed in March 2016) had slightly lower ratings for logic and realism of the disruptive (Mean = 3.71, SD = .51) and non-disruptive students' behavior and dialogue (Mean = 3.60, SD = .53) when compared to ratings of revised vignettes (November 2016) for logic and realism of the disruptive (Mean = 3.81, SD = .28) and non-disruptive students' behavior and dialogue (Mean = 3.67, SD = .50). Original vignettes (developed in March 2016) also had slightly lower ratings for logic and realism of teacher responses to students (Mean = 3.58, SD = .72) when compared to ratings of revised vignettes (November 2016) for teacher responses to students (Mean = 3.62, SD = .60). When comparing vignettes by presenting problem there were higher ratings for logic and realism of the non-disruptive students' behavior and dialogue (Mean = 3.70, SD = .43) when compared to the off-task vignettes (Mean = 3.6, SD = .35). Teacher response options were also rated higher for the aggressive/non-compliant vignettes (Mean = 3.63, SD =.492) when compared to the off-task vignettes (Mean = 3.4, SD = .53). In addition, the aggressive/non-compliant vignettes were rated as more engaging (Mean = 3.76, SD = .36) than the off-task vignettes (Mean = 3.44, SD = .39).

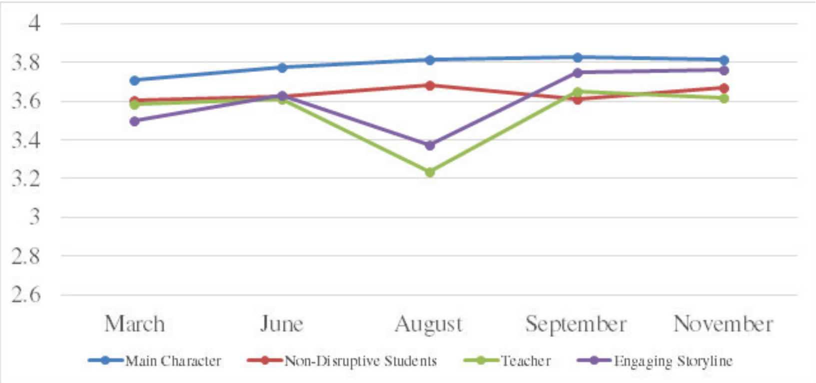

Figure 4. Vignette ratings over time.

Descriptive feedback from advisory board members was generally positive and converged with high numeric ratings related to logic and realism of the vignettes. Descriptive feedback also highlighted additional areas for which storylines could be more realistic. As an example, one advisory board member shared, *"This vignette was very hard to follow... It's not realistic to use class time to tell student to clean his backpack as he is already falling behind in reading and the goal is to get him on task*." Descriptive feedback also suggested the need to improve the logic and realism of some of the dialogue, interactions, and responses of the teachers and ensuring decision point options were predictable across vignettes. For instance, one advisory board member noted, *"I really liked the paths and I thought that the students' reactions and behaviors were pretty typical for 1st graders. However, I thought some of the teacher options were too harsh for a 1st grade teacher. I also thought that some decision point options were inconsistent…why was the teacher ignoring Sofia when she starts drawing on her desk?"*

DISCUSSION

This study illustrated the development and iterative refinement of IVT-T, a simulation-based training model for early career teachers to enhance their behavior management skills with virtual students. The IVT-T interdisciplinary development team capitalized on the expertise of advisory board members who had extensive experience teaching in high poverty schools and provided feedback on authenticity/realism of IVT-T vignettes, virtual students, and classrooms in an effort to iteratively improve upon IVT-T design features.

Main Findings

Students and classroom authenticity/realism ratings were generally high and increased across prototypes. We compared mean ratings over time for student, classroom, and vignette prototypes to evaluate the extent to which authenticity/realism achieved a minimum standard on these design features. a minimum standard on these design features. Iterative revisions concluded when student and classroom ratings fell in the *Good* or *Outstanding* range and when advisory board members *Agreed* or *Strongly Agreed* that the students' behavior, actions, and dialogue were logical and realistic based on age and presenting problem. The iterative nature of developing and refining students and classrooms proved to be useful across time. For example, early student prototypes in Group 1 underwent the most changes, while later student prototypes in Group 3 underwent the least, pointing to the value of feedback by the advisory board to ultimately generate a set of virtual students (28 of 30) judged authentic with regard to body, face, hair, and clothing (Bannan-Ritland, 2012; Hix & Hartson, 1993). Similarly, the 1st grade classroom (developed first) required three times as many revisions as the 6th grade classroom (developed second), again suggesting that advisory feedback was usable and generalizable, improved design revisions, and reduced the need for extensive changes to prototypes introduced later.

Advisory board members rated the aggressive/non-compliant vignettes as more realistic and engaging in terms of student and teacher behaviors and storylines when compared to vignettes featuring off-task students. This finding may reflect differences between the types of interactions that occur between teachers and students who struggle with hyperactivity and inattention as opposed to oppositionality and defiance, which often are characterized as social in nature and occurring in a relational context (Egger & Angold, 2004; Gray et al., 2012). Open-ended feedback supported positive numeric ratings but also highlighted the need to improve teacher tone, developmental appropriateness of student and teacher responses, and consistency within storylines.

Advisory Board Contributions

There is a rich literature pointing to the advantages of inviting end users to contribute early and often to the development of interventions, technologies, and professional development models that ultimately are intended to serve and support them (Hartson & Pyla 2012). Several considerations of research-practice partnerships, and discussions with school personnel participating in prior studies with the research team, led to forming and relying on an advisory board as part of the iterative process for building IVT-T. Increasing priority on research-practice partnerships in health services and education in uenced the team's conceptualization of the role and goals of the advisory board. Speci cally, active stakeholder voice was expected to improve the process of iterative development and the product that would result, yielding a higher quality, authentic, engaging and pedagogically rich training model.

Although end users for this study are early career teachers, their limited classroom experience minimized the extent to which they could comprehensively assess the realism and authenticity of high poverty classrooms and student-teacher interactions. Thus, we invited retired teachers from high poverty schools to serve on the board, thereby capitalizing on the views and experiences of highly experienced educators during the early stages of development (Morin, Maiorana, Koester, Sheon, & Richards, 2003; Santos & Chess, 2003). Their participation involved frequent and systematic opportunities to provide feedback and recommendations on the evolving students, classrooms, and vignettes. In this way, the advisory board served as a bridge between the technical team and content experts, always informing and often changing the dialogue and decisions in ways that better re ected the nuances of high poverty schools. Interestingly, numeric ratings from advisory board members were generally high, even during early iterations. Therefore, the team relied even more heavily on open-ended feedback to improve the technology and to ultimately enhance teacher engagement and learning.

Next Steps in IVT-T Development

After the three IVT-T components described herein are fully developed, the remaining funding years are devoted to assessing usability (Year 2), eld testing the model (Year 3), and conducting an underpowered randomized controlled trial (Year 4). Completion of the IVT-T website in Year 2 will begin with casting vignettes into storyboards or pictorial translations of what the classrooms will look like, the sequence of interactions between the teachers and virtual students, and visual representations of the 3D characters using stick figures. Storyboarding will also facilitate decisions regarding the functional and non-functional requirements for IVT-T and the website user interface. Storyboards will inform a 2D prototype that will undergo expert heuristic evaluation by the usability engineer to assess the quality of the user interface and to identify usability problems (Gabbard, Hix & Swan, 1999). A formative usability evaluation will follow, during which representative users (e.g., education students) will use a 2D prototype to identify potential usability problems that may not have been identified during the expert evaluation (Gabbard et al., 1999; Hix & Hartson, 1993). These additional opportunities for feedback will help ensure that the sequence of behaviors correspond to the vignette and that interactions are authentic before extensive animations are programmed. Year 3 will include initial field testing of IVT-T with a sample of early career teachers in high poverty schools to assess knowledge and skill transfer from the virtual to the live classrooms. Year 4 is a planned randomized controlled trial to assess the promise of the model to improve teacher effectiveness and student behavior.

Our emphasis in the current set of planned studies includes supporting early career teachers in high poverty schools. However, because we are developing IVT-T as a web-based application for use with existing computing systems widely available to teachers, IVT-T, if effective, could accommodate broader professional development needs, including experienced teachers struggling with behavior management and teachers in other geographic areas (e.g., rural and suburban school districts) with minimal cost associated with system maintenance after the initial investment in development is made. Future lines of inquiry will focus on evaluating which teachers derive the most benefits from IVT-T, for instance related to years of experience; geographic location; experience with technology; or quality of preservice training. Future research may also indicate that IVT-T is necessary, but insufficient to meet all early career teachers' training needs, and there may be a subset of teachers who require more intensive support in the form of live coaching and/or additional support to help them effectively prevent and manage disruptive behaviors.

Interdisciplinary Collaborations: Challenges and Opportunities

The existing literature highlights two common approaches to integrating technology into traditional training models (Moreno & Mayer, 2007; Rooney, O'Rourke, Burke, MacNamee, & Igbrude, 2009). The first approach includes adapting a commercial off-the-shelf system for new users and the second approach includes developing a new training system to address unique pedagogical or curricular content needs (Dalgarno & Lee, 2010; Jacobson, June Lee, Hong Lim, & Hua Low, 2008). Each approach requires interdisciplinary collaboration and has corresponding advantages and disadvantages (Rooney et al., 2009). Repurposing a commercial training system to support early career teachers to improve their behavior management skills (e.g., using TeachLivE developed by Dieker et al., 2008; 2014) could have been more ef cient given the cost in time, equipment, and computational power required for 3D applications such as IVT-T (Dalgarno & Lee, 2010). Deadlines associated with Year 1 activities required the interdisciplinary team to make dif cult decisions regarding how many times to iterate, particularly the 3D elements of IVT-T given the time-consuming nature of graphic design and the goal of creating highly realistic students and classrooms.

Despite the signi cant time and resource challenges, development and iterative re nement from the ground-up allowed the team to align the model with the unique needs of early career teachers; incorporate important pedagogical elements of altering the antecedents and consequences of disruptive behavior (Evertson & Weinstein, 2006; Simonsen et al., 2008); portray a relatively large number of students in each classroom environment; and create autonomous and automatic behaviors for the virtual students (Kelley, 1984). A ground-up approach also facilitated a web-based and open access training model that is compatible with existing computing systems available to teachers.

The interdisciplinary collaborations required to build IVT-T also provided unique opportunities to expand our mutual understanding of the focus, roles, contributions, and expertise of each discipline. These collaborations also revealed additional lines of inquiry focused on lifecycle activities to develop 3D training systems that involve high-quality graphics, advanced user interface and user experience design, and software engineering.

Limitations

A number of limitations warrant mention. First, designing a simulation training model from the ground-up is both dif cult and costly in time and resources, limiting the number of iterations that are possible in the context of development. The advisory board was small and the total number of iterations and data points was limited. This study also relied exclusively on self-report measures from a single type of reporter (i.e., advisory board members), which are subject to bias. Ceiling effects related to advisory board ratings for the vignettes, students, and classrooms also made comparisons between prototypes over time more challenging. In addition, mean increases and decreases do not reflect statistically or clinically significant differences. Nevertheless, this work represents a step forward with regards to using standardized measures and methods for evaluating and refining simulation systems.

Implications for Teacher Education

Perhaps one of the most interesting findings was the inconsistency between quantitative ratings and open-ended responses from advisory board members regarding the classroom, character, and vignette prototypes. Quantitative ratings were generally high and increased across prototypes, whereas open-ended feedback revealed important shortcomings and the need for further refinements to enhance the authenticity and realism of IVT-T. These inconsistencies have important implications for teacher educators involved in the development of instructional technologies, particularly during the early phases of development and iterative refinement. Inviting both types of feedback can bring a more comprehensive lens by which to evaluate instructional technology programs, and yield a more nuanced, thorough, and complete understanding of training components than either method alone would provide (Creswell & Plano Clark, 2011; Shernoff et al., 2016; Tashakkori & Teddlie, 2008). Soliciting both numeric and open-ended feedback also may help to mitigate socially desirable evaluations of a product when users are aware that their opinions are being shared directly with developers. In the current study, open-ended responses pulled for constructive criticism and yielded specific recommendations to help improve upon IVT-T components.

Findings related to the number and timing of character and classroom prototypes introduced also have implications for teacher educators and researchers involved in the development and evaluation of instructional technologies. Specifically, early student and classroom prototypes underwent the most iterations and later prototypes underwent the fewest. Together, these findings point to the value and efficiency of investing time early in the development process to obtaining feedback and making refinements to the labor-intensive graphics component of 3D training programs. Specifically, the computer science team worked closely with the education team to establish requirements for the virtual students and classrooms, and iterated multiple times to create prototypes increasingly corresponding to team specifications and advisory board feedback. Understanding, for example, the basics of the body sizes and proportions based on early advisory board feedback informed design decisions for the remaining characters of the same age group.

Given new instructional technologies are expected to be developed, tested, and deployed to improve teacher education at exponential rates (Battista & Boone, 2015; Schrader, 2008; Schrader et al., 2011), school districts, administrators, and teacher educators are likely to be flooded with and required to make important decisions regarding which technologies to integrate. Historically, a high premium has been placed on selecting programs that have been rigorously evaluated via traditional and gold standard empirical approaches, including random assignment of participants to conditions, statistically significant differences between conditions, and replication by independent investigators (What Works Clearinghouse, retrieved from https://ies.ed.gov/ncee/wwc/). Findings from the current study illustrate the unique opportunity for seasoned educators to improve technology-based programs, highlighting the value of blending scholarly with local knowledge particularly early in the development of instructional technologies. In particular, contributions by the expert advisory board helped to maximize alignment and fit of IVT-T with the local context of urban, high poverty schools, in part by ensuring that characters, classrooms, and vignettes adequately reflected the realities of teaching in those contexts. Veteran educators may also play a key role in the installment of educational technologies within school districts, particularly if they are indigenous to schools and influential in their role as instructional experts, which also has implications for the sustainability of these models over time (Atkins et al., 2015). Altogether, findings from the current study provide teacher educators with additional evidence by which to evaluate and select instructional technologies, including the extent to which local knowledge informed early development to improve user experience.

Conclusions

Studies document that student disruptive behavior is one of the strongest drivers of turnover among early career teachers (Gonzalez et al., 2008; Ingersoll & Smith, 2003; Shernoff, Mehta, Atkins, Torf, & Spencer, 2011) and that currently certified teachers are not adequately prepared to prevent or manage the disruptive behaviors that interrupt classroom instruction on a regular basis (Coalition for Psychology in the Schools and Education, 2006). As the achievement gap widens, there is a pressing need for early career teachers to build these skills particularly in high poverty schools where students are at increased risk for both behavior problems and academic failure. Simulation technology provides an unprecedented opportunity to in uence teacher education by providing novices with access to realistic, low-stakes interactions with disruptive students. This includes extended opportunities to practice responding to disruptive behaviors that affords signi cant practical and ethical advantages over traditional high-stakes on-the-job training.

Simulation training models are emerging quickly and hold promise for enhancing the professional development of early career teachers but without a strong evidence-base regarding the methodology, procedures, and tools required to evaluate and improve upon these training systems (Alvarez & Michaud, 2008; Arnab et al., 2015; Jeremic, Jovanovic & Gaševic, 2009; Schrader et al., 2011). As a result, there are a limited number of uniform measures to assess important design features of the instructional platform, which leaves a critical gap in the literature (Hartson & Pyla, 2012; Girard et al., 2012; Pallavincini, et al. 2015). This impedes developers from systematically evaluating the training system and its potential bene ts for users in terms of skills and competencies. As technology-based interventions become more prevalent in education, there is demand for systematic methodologies and tools to evaluate simulation training models in carefully controlled, larger studies (Alvarez & Michaud, 2008; Arnab et al., 2015; Jeremic, Jovanovic & Gaševic, 2009; Schrader et al., 2011). The procedures described herein highlight the promise of using an advisory board model and an iterative process of soliciting quantitative and open-ended feedback from experienced users to enhance the authenticity, realism, and representativeness of a simulation training model for teachers.

References

Amory, A. (2007). Game object model version II: A theoretical framework for educational game development. Educational Technology Research and Development, 55(1), 51-77. doi:10.1007/s11423-006-9001-x

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., Freitas, S., Louchart, S., ... & De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis. British Journal of Educational Technology, 46(2), 391-411. doi:10.1111/bjet.12113

Alvarez, J. & Michaud, L. (2008). Serious Games—Advergaming, edugaming, training and more. IDATE Consulting and Research. Retrieved July 15, 2013, from http://www.scribd.com/doc/28528416/2008-Serious-Games-Report#

Atkins, M.S., Shernoff, E.S., Frazier, S.L., Schoenwald, S.K., Cappella, E., Marinez-Lora, A., Mehta, T.G., & Bhaumik, D. (2015). Re-designing community mental health services for urban children: Supporting schooling to promote mental health. *Journal of Consulting and Clinical Psychology, 83*, 839-852.

Bannan-Ritland, B. (2009). The integrative learning design framework: An illustrated example from the domain of instructional technology. In T. Plomp & N. Nieveen (Eds.), *An Introduction to Educational Design Research*. Enschede, Netherlands: Institute for Curriculum Development.

Battista, N. U., & Boone, W. J. (2015). Exploring the impact of TeachME Lab virtual classroom teaching simulation on early childhood education majors' self-efficacy beliefs. *Journal of Science Teacher Education*, *26*, 237–262.

Bellotti, F., Berta, R., & De Gloria, A. (2010). Designing effective serious games: Opportunities and challenges for research. *iJET*, *5*(SI3), 22-35. doi:10.3991/ijet.v5s3.1500

Beutler, L. E., & Harwood, T. M. (2004). Virtual reality in psychotherapy training. *Journal of Clinical Psychology*, *60*, 317-330. doi:10.1002/jclp.10266

Coalition for Psychology in Schools and Education (2006, August). Report on the Teacher Needs Survey. Washington, D.C.: American Psychological Association, Center for Psychology in Schools and Education.

Creemers, B. P. M. (1994). Effective instruction as a basis for effective education in schools. *Tijdschrift Voor Onderwijsresearch*, *19*, 3-16.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed-methods research* (2nd Ed.). Thousand Oaks, CA: Sage.

Crone, L. J., & Teddlie, C. (1995). Further examination of teacher behavior in differentially effective schools: Selection and socialization process. *The Journal of Classroom Interaction, 30*(1), 1-9.

Dalgarno, B., & Lee, M. J. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10-32.

Denton, C. A., & Hasbrouck, J. A. N. (2009). A description of instructional coaching and its relationship to consultation. *Journal of Educational and Psychological Consultation*, *19*(2), 150-175. doi: 10.1080/10474410802463296

Dieker, L. A., Hynes, M., Hughes, C., & Smith, E. (2008). Implications of mixed reality simulation technologies on special education and teacher preparation. *Focus on Exceptional Children*, *40*(6), 1-20.

Dieker, L., Hynes, M., Stapleton, C., & Hughes, C. (2007). Virtual classrooms: STAR simulator. *New Learning Technology SALT*, *4*, 1-22. doi:10.1.1.76.5686

Dieker, L. A., Rodriguez, J. A., Lignugaris, Kraft, B., Hynes, M. C., & Hughes, C. E. (2014). The potential of simulated environments in teacher education: Current and future possibilities. *Teacher Education and Special Education*, *37*(1), 21-33.

Dunbar, N. E., Jensen, M. L., Bessarabova, E., Burgoon, J. K., Bernard, D. R., Harrison,

K. J., Kelley, K. M., … & Eckstein, J. M. (2014). Empowered by persuasive deception: The effects of power and deception on dominance, credibility, and decision making. *Communication Research*, *41*(6), 852-876.

Egger, H. L., & Angold, A. (2004). The Preschool Age Psychiatric Assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. *Handbook of infant, toddler, and preschool mental health assessment*, 223-243. New York, NY: Oxford University Press.

Evertson, C. M., & Weinstein, C. S. (Eds.). (2006). *Handbook of classroom management: Research, practice, and contemporary issues*. Mahwah, NJ: Lawrence Erlbaum.

Fletcher, J. D. (2009). Education and training technology in the military. *Science 323*(5910), 72–75. doi:10.1126/science.1167778

Gabbard, J. L., Hix, D., & Swan, J. E. (1999). User-centered design and evaluation of virtual environments. *IEEE computer Graphics and Applications*, *19*(6), 51-59.

Gee, J. P. (2005). Learning by design: Good video games as learning machines. *E-Learning and Digital Media*, *2*(1), 5-16. doi:10.2304/elea.2005.2.1.5

Girard, C., Ecalle, J., & Magnan, A. (2012). Serious games as new educational tools: how effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *29*(3), 207-219. doi: 10.1111/j.1365-2729.2012.00489.x

Gonzalez, L., Brown, M. S., & Slate, J. R. (2008). Teachers who left the teaching profession: A qualitative understanding. *The Qualitative Report*, *13*(1), 1-11.

Gray, S. A. O., Carter, A. S., Briggs-Gowan, M. J., Hill, C., Danis, B., Keenan, K., & Wakschlag, L. S. (2012). Preschool children's observed disruptive behavior:Variations across sex, interactional context, and disruptive psychopathology, *Journal of Clinical Child and Adolescent Psychology, 41*(4), 499–507.

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184-205. doi: 10.3102/0002831207312906

Hartson, R., & Pyla, P. S. (2012). *The UX Book: Process and guidelines for ensuring a quality user experience*. New York, NY: Elsevier.

Heeren, B., & Jeuring, J. (2014). Feedback services for stepwise exercises. *Science of Computer Programming*, *88*(1), 110-129. doi:10.10166/j.scico.2014.02.021

Henry, D. B., Farrell, A. D., Schoeny, M. E., Tolan, P. H., & Dymnicki, A. B. (2011). Influence of school-level variables on aggression and associated attitudes of middle school students. *Journal of School Psychology*, *49*(5), 481-503. doi:10.1016/j.jsp.2011.04.007

Hix, D., & Hartson, H.R. (1993). *Developing user interfaces: Ensuring usability through product & process.* New York: John Wiley and Sons.

Horswill, I., & Lisetti, C. L. (2011, November). On the Simulation of Human Frailty. In A.V., Samsonovich, & K.R., Jóhannsdóttir (Eds.), Biologically Inspired Cognitive Architectures 2011. Paper presented at the Second Annual Meeting of the BICA Society, Arlington, USA.

Ingersoll, R. M., & Smith, T. M. (2003). The wrong solution to the teacher shortage. Educational Leadership, 60(8), 30-33.

Jacobson, M.J., June Lee, B.K., Hong Lim, S., & Hua Low, S. (2008). An intelligent agent augmented multi-user virtual environment for learning science inquiry: Preliminary research findings. Paper presented at the 2008 American Educational Association Conference. New York, NY.

Jeremic, Z., Jovanovic J., & Gasevic, D. (2009). Evaluating an intelligent tutoring system for design patterns: The DEPTHS experience. Educational Technology & Society, 12(2), 111-130.

Junod, R. E. V., DuPaul, G. J., Jitendra, A. K., Volpe, R. J., & Cleary, K. S. (2006). Classroom observations of students with and without ADHD: Differences across types of engagement. Journal of School Psychology, 44(2), 87-104.

Kazdin, A. E. (2015). Technology-Based interventions and reducing the burdens of mental illness: Perspectives and comments on the special series. Cognitive and Behavioral Practice, 22(3), 359-366. doi:10.1016/j.cbpra.2015.04.004

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. Perspectives on Psychological Science, 6(1), 21-37.

Kazdin, A.E. (2005). Parent management training (pp. 35-64). New York: Oxford University Press.

Kelley, J. F. (1984). An iterative design methodology for user-friendly natural language office information applications. ACM Transactions on Information Systems, 2(1), 26–41.

Kelly, H., Howell, K., Glinert, E., Holding, L., Swain, C., & Burrowbridge, A. (2007). How to build serious games. Communications of the ACM, 50(7), 44-49. doi:10.1145/1272516.1272538

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. Review of Educational Research, 86(3), 643-680. doi: 10.3102/0034654315626799

Magnenat-Thalmann, N., & Thalmann, D. (Eds.). (2005). Handbook of virtual humans. Chichester, UK: John Wiley & Sons.

Mantovani, F., Castelnuovo, G., Gaggioli, A., & Riva, G. (2003). Virtual reality training for healthcare professionals. CyberPsychology & Behavior, 6(4), 389-395.

Moreno, R., & Mayer, R (2007). Interactive multimodal learning environments. Educational Psychology Review, 19(3), 309-326. doi:10.1007/s10648-007-9047-2

Morin, S. F., Maiorana, A., Koester, K. A., Sheon, N. M., & Richards, T. A. (2003). Community consultation in HIV prevention research: A study of community advisory boards at 6 research sites. Journal of Acquired Immune De ciency Syndromes, 33(4), 513-520.

Oliver, R. M., & Reschly, D. J. (2007). Effective Classroom Management: Teacher Preparation and Professional Development. TQ Connection Issue Paper. National Comprehensive Center for Teacher Quality.

Pallavincini, F., Toniazzi, N., Argenton, L., Aceti, L., & Mantovani, F. (2015). Developing effective virtual reality training for military forces and emer-gency operators: from technology to human factors. In International Con-ference on Modeling and Applied Simulation, MAS 2015 (pp. 206-210). Dime University of Genoa

Paschall, M., Fishbein, D., Hubal, R., Eldreth, D. (2005) Psychometric proper-ties of virtual reality vignette performance measures: a novel approach for assessing adolescents' social competency skills, Health Education Research, 20(1) 61-70 https://doi.org/10.1093/her/cyg103

Penuel, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. American Educational Research Journal, 44(4), 921-958.

Regalla, M., Hutchinson, C., Nutta, J., & Ashtari, N. (2016). Examining the im-pact of a simulation classroom experience on teacher candidates' sense of ef cacy in communicating with English learners. Journal of Technology and Teacher Education, 24(3), 337-367.

Rooney, P., O'Rourke, K. C., Burke, G., MacNamee, B., & Igbrude, C. (2009, March). Cross-disciplinary approaches for developing serious games in Higher Education. In Games and Virtual Worlds for Serious Applications, 2009. VS GAMES'09 (pp. 161-165). IEEE.

Sabers D., Cushing, K., & Berliner, D. (1991). Differences among teach-ers in a task characterized by simultaneity, multidimensionality, and immediacy. American Educational Research Journal, 28(1), 63-88. doi:10.3102/00028312028001063

Santos, S. L., & Chess, C. (2003). Evaluating citizen advisory boards: The im-portance of theory and participant-based criteria and practical implications. Risk Analysis, 23(2), 269-279.

Schrader, P. G., Archambault, L. M., & Oh-Young, C. (2011). Training by gam-ing: Preparing teachers of today for tomorrow's learning environments. Journal of Technology and Teacher Education, 19(3), 261-286.

Schussler, D., Frank, J., Lee, T., & Mahfouz, J. (2017). Using virtual role-play to enhance teacher candidates' skills in responding to bullying. Journal of Technology and Teacher Education, 25(1), 91-120.

Shernoff, E.S., Frazier, S.L. Marinez-Lora, A., Lakind, D., Atkins, M.S., Jakob-sons, L., & Smylie, M. (2016). Expanding the role of school psychologists to support early career teachers: A mixed-method study. School Psychology Review, 45(2), 226-249.

Shernoff, E.S., Lakind, D., Frazier, S.L. & Jakobsons, L. (2015). Coaching early career teachers in urban elementary schools: A mixed-method study. *School Mental Health, 7*(1), 6-20. doi:10.1007/s12310-014-9136-6

Shernoff, E.S., Lekwa, A.J., Reddy, L.A. & Coccaro, C. (2017). Examining teachers' attitudes and experiences with coaching to inform research-based practice: An iterative developmental design study. *Journal of Educational and Psychological Consultation, 27*(4) 459-485, DOI: 10.1080/10474412.2016.1255850

Shernoff, E. S., Mehta, T. G., Atkins, M. S., Torf, R., & Spencer, J. (2011). A qualitative study of the sources and impact of stress among urban teachers. *School Mental Health*, *3*(2), 59-69.

Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children*, *31*(3), 351-380. doi:0.1353/etc.0.0007

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology*, *64*(2), 489-528.

Tawadrous, M., Kevan, S. D., Kapralos, B., & Hogue, A. (2012). A serious game for incidence response education and training. *International Journal of Technology, Knowledge & Society*, *8*(4), 177-184, doi:10.18848/18323669/CGP/v08i04/55528

Tashakkori, A., & Teddlie, C. (2008). Quality of inferences in mixed methods research: Calling for an integrative framework. In M.M. Bergman (Ed., pp. 101-119), *Advances in Mixed Methods Research: Theories and Applications*. Thousand Oaks, CA: Sage.

Vannini, N., Enz, S., Sapouna, M., Wolke, D., Watson, S., Woods, S., ... & Aylett, R. (2011). "Fear Not:" A computer-based anti-bullying-programme designed to foster peer intervention. *European Journal of Psychology of Education*, *26*(1), 21-44.

Veenman, S. (1984). Perceived problems of beginning teachers. *Review of Educational Research*, *54*(2), 143-178. doi:10.3102/00346543054002143

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research*, *34*(3), 229-243. doi: 10.2190/FLHV-K4WA-WPVQ-H0YM