

Using Virtual Role-Play to Enhance Teacher Candidates' Skills in Responding to Bullying

DEBORAH SCHUSSLER, JENNIFER FRANK,
TSAN-KUANG LEE, AND JULIA MAHFOUZ

Pennsylvania State University, USA

dls93@psu.edu

jfrank@psu.edu

tklee@psu.edu

jxm1093@psu.edu

Nearly one in three students in the United States today is negatively impacted by bullying. Teachers can play a critical role in stopping bullying-related violence, but many struggle with how to engage students in difficult conversations. Traditional classroom-based pedagogy used to teach communication skills (e.g., modeling & role-play) is inefficient and limited in scope. Recent advances in Artificial Intelligence have resulted in the development of virtual pedagogical agents designed to simulate authentic conversations and can be leveraged to teach vital communication skills. This exploratory study examines the effectiveness of a “virtual role-play” (VRP) tool developed to help teacher candidates effectively respond to classroom bullying by providing them with opportunities to engage in repeated, authentic practice conversations. We hypothesized that practice in simulated conversations provided by VRP would improve teacher candidates’ communication skills by improving their fluency in responding to classroom bullying. Our findings revealed that the repeated practice afforded by VRP improved candidates’ fluency by helping them focus on more refined word choice, reducing their reliance on written notes, and improving their nonverbal skills. Implications regarding use of VRP to prepare candidates to communicate in bullying situations are discussed.

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A150391 awarded to The Pennsylvania State University (Jennifer L. Frank). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Keywords: teacher preparation, role-play, artificial intelligence, instructional technology

Nearly one in three students in the United States today is negatively impacted by bullying (U.S. Department of Education, 2013). Left untreated, victims and perpetrators are at increased risk for a variety of psychosocial problems resulting from repeated exposure to violence (Pepler, Jiang, Craig, & Connolly, 2008). Teachers can play a critical role in stopping bullying-related violence. However, a majority of teachers report they are reluctant to become involved in peer conflicts because they do not know what to do and fear making the situation worse (Olweus, 1993; Rigby & Griffiths, 2010), so frequently, they do nothing.

In teacher certification programs, traditional classroom-based pedagogy used to teach communication skills frequently consists of modeling and role-play. Typically, only a select few volunteers have an opportunity to engage in role-play practice and receive substantive feedback, both which are critical for student learning (Jeffries, 2005). Furthermore, students frequently claim this pedagogical method is ineffective (Stevenson & Sander, 2002). Role-play-like activities, such as “micro-teaching” and “performance tasks,” which grew in popularity because they attempted to replicate authentic tasks of teachers (Girod & Girod, 2008), lack the contextual elements of interacting with others and repetitive practice. Thus, many teachers enter the field less confident, less prepared, and therefore less likely to intervene when bullying actually occurs (American Educational Research Association, 2013).

Recent advances in Artificial Intelligence (A.I.) have resulted in the development of virtual pedagogical agents, or “chatbots,” designed to simulate authentic conversations (Kinnebrew & Biswas, 2011; Segedy, Kinnebrew, & Biswas, 2013). Capitalizing on this technology, we developed a “virtual role-play” (VRP) tool to help teacher candidates effectively respond to bullying by providing them with opportunities to engage in repeated, authentic practice conversations. In this exploratory study to investigate the effectiveness of VRP, we asked the following research questions:

1. In a hypothetical bullying situation, how does the communication fluency of teacher candidates who practiced their skills using VRP compare to candidates who practiced their skills through in-depth reflection of their role-play dialogues?
2. How do teacher candidates compare their experiences using VRP with traditional role-play?

To explore these questions, we employed a content analysis of candidates' role-plays both before and after they received instruction on bullying across six communication domains. It was our belief that practice in simulated conversations provided by VRP would improve teacher candidates' communication skills by improving their fluency in responding to classroom bullying.

REVIEW OF LITERATURE

Bullying

Bullying is commonly defined as unwanted, intentional, aggressive behavior that involves a real or perceived power imbalance that is often repeated over time (Olweus, 1993). It differs from teasing or physical play in that the bully is unilaterally exhibiting power over and intends to harm the target (Olweus, 1993). Bullying is frequently classified as verbal (e.g., teasing, name calling, making threats, taunting), social/relational (e.g., marginalizing, starting rumors, embarrassing someone), or physical (Espelage & Swearer, 2003). In the last decade with the proliferation of social media outlets, "cyber-bullying" has also become more prevalent (Wang, Iannotti, & Nansel, 2009).

The prevalence of bullying is difficult to pinpoint, both because definitions of bullying are not used consistently in the research and because bullying that occurs out of school frequently spills over into bullying in school. According to the National Center for Education Statistics (2013), 28% of children between 6th and 12th grade have been bullied. Wang et al. (2009) report anywhere from 13% to 53% of students have either bullied or been the target of bullying, depending on the type of bullying (cyber: 13%; verbal: 53%). Some students are more likely to be the target of bullying behaviors. Students with disabilities as well as students who identify or are perceived as lesbian, gay, bisexual, or transgender tend to be more targeted by bullying (Rose & Espelage, 2012). It should also be noted that when an individual's peers perceive him or her to have low social skills or difficulties communicating, that individual tends to be a target for bullying behaviors more than an individual who is higher functioning (Rose, Monda-Amaya, & Espelage, 2011). Eli, the potentially bullied chatbot we created, displayed some of these qualities.

The effects of bullying are significant, for the target of bullying as well as the bystanders who witness bullying. The targets of bullying experience

higher rates of anxiety, depression, physical health problems, and social adjustment problems which can persist into adulthood (Espelage, Low, & De La Rue, 2012). They are less engaged in school, experience lower grades and declining test scores (Mehta, Cornell, Fan, & Gregory, 2013). They also are likely to skip, miss, or drop out of school (Cornell, Gregory, Huang, & Fan, 2013). Bystanders also experience mental health problems like anxiety and depression, are at risk for increased substance use, and are at increased risk of dropping out of school or missing classes. In fact, researchers have concluded that the psychological distress experienced by targets of bullying and bystanders was similar or greater than the psychological distress experienced by those exposed to rape, earthquakes, avalanches, sexual abuse, community violence, combat, and terrorism (Janson, Carney, Hazler, & Oh, 2009).

Teachers play a critical role in curtailing bullying (Rigby & Griffiths, 2010). However, many teachers fail to intervene when bullying occurs. The reasons for this are inter-related. First, there is ample evidence that teachers grossly underestimate bullying that occurs (Holt & Keys, 2004). Although teachers are more adept at identifying bullying that is physical in nature, both practicing teachers and teacher candidates struggle to identify relational bullying as bullying and are therefore, less apt to intervene (Voulgaridou & Kokkinos, 2015). In addition, there are students who are perceived as non-aggressive, based on teachers' observations of their behaviors in the classroom, who exhibit bullying behaviors in less conspicuous places, especially on the school playground (Craig, Pepler, & Atlas, 2000). In fact, in a rare observational study Craig and Pepler (1997) found that adult supervisors reported playground bullying only 4% of the time. Because relational bullying is less likely to be identified by teachers (Voulgaridou & Kokkinos, 2015), especially when it occurs on the playground (Craig, Pepler, et al., 2000), we purposefully created a case incorporating these elements.

Another reason teachers often fail to intervene in bullying is they report they do not know what to do or say, especially when it comes to less overt bullying, including relational bullying, like social exclusion (Rigby & Griffiths, 2010). In a study specifically examining teacher candidates' attitudes toward bullying, researchers found that physical interactions were more likely to be labeled as bullying and were viewed as more serious, and the likelihood of intervening was greater if the bullying was witnessed by the adult (Craig, Henderson, & Murphy, 2000). Social exclusion was the least likely to be labeled as bullying or to warrant intervention. However, the researchers found empathy was the only significant predictor of teacher candidates labeling witnessed social exclusion as bullying and one of the predictors in labeling and intervening in verbal aggression (Craig, Hender-

son, et al., 2000). Similarly, Yoon (2004) found that perceived seriousness of the bullying incident, empathy of the teacher, and teacher efficacy were major predictors for determining the extent to which elementary teachers would intervene. Not only do teachers need to know how to identify different kinds of bullying, they also need the empathy and efficacy to know how to communicate effectively to students about these incidents.

Although researchers of bullying have made a plea for the topic receiving more attention in teacher education, a lack of information and especially a lack of practice in handling such situations prevails (American Educational Research Association, 2013; Craig, Henderson, et al., 2000). A research report produced by the American Educational Research Association (2013) cites the “serious disconnect” between universities and K-12 schools regarding bullying preparation and notes the “unmet need for preservice university training” (p. 60). Most bully prevention programs (e.g., Steps to Respect, Method of Shared Concern, Olweus) are introduced only to in-service teachers leaving teacher candidates without any training.

In addition, we observed that bullying prevention programs focus more on general principles and less on the specifics of what to say and when to say it. Exacerbating matters, traditional pedagogies, like in person role-play, often fail to provide students with adequate “think-time” to make connections between content learning and applied practice (Jeffries, 2005). Traditional role-play is unable to slow down normal conversations so that teacher candidates have time to think deeply about how they ideally would like to respond. Also, in most classrooms, only a select few volunteers have an opportunity to engage in role-play practice and receive substantive feedback. Without adequate practice, it is difficult to develop the fluency and confidence to know what to do and what to say when critical incidents, like bullying, arise (Yoon, 2004).

Research on role-play in the medical field consistently shows that realistic scenarios and quality feedback are critical indicators of student learning (Jeffries, 2005). Providing teacher candidates with opportunities to practice their skills has been a continual challenge for teacher educators. In fact, NCATE’s Blue Ribbon Panel on Clinical Preparation demands “programs that are fully grounded in clinical practice” (National Council for Accreditation of Teacher Education, 2010, p. ii). The ten design principles include the use of “technology applications to foster high-impact preparation,” clinical preparation that is “integrated throughout every aspect of teacher education,” and “use of data” to evaluate whether candidates are learning skills. The move toward more performance assessments, including edTPA, which has been adopted by institutions in 34 states, is indicative of this increased emphasis on authentic, performance-based tasks (Newton, 2010). However,

there exists a tension in providing teacher candidates with authentic experiences while not risking the well-being of real children in the process.

Virtual Role-play Technology

One means of providing teacher candidates with authentic, meaningful practice is engaging in simulated conversations with an animated pedagogical agent, also called a “chatbot.” A chatbot is an A.I. computer program designed to simulate an intelligent conversation with human users. Classical chatbots use pattern matching algorithms and natural language processors to comprehend user input and select appropriate responses. Among many representational systems to define and describe this process, an extensible markup language (XML) called Artificial Intelligence Markup Language (AIML) is widely used. Open-source chatbot packages come with pre-built AIML knowledge bases and AIML interpreters. For example, A.L.I.C.E (Artificial Linguistic Internet Computer Entity) contains a knowledge base of over 41,000 categories that provides users a highly realistic experience. Developers in turn customize this knowledge base to simulate desired patterns of responses or “personalities,” by modifying its knowledge categories and conditional triggers. Such chatbots can be hosted on commercial sites or dedicated servers running the requisite AIML interpreters. The goal is to create a highly realistic chat experience. In fact, in 2011 the chatbot “cleverbot” participated in a formal Turing Test and was judged to be 59.3% human, while actual humans were judged only 63.3% human (<http://www.cleverbot.com/human>).

There is some research examining the use of chatbots in a variety of settings. They have become an increasingly popular tool to support training in skills that require sustained practice and a self-paced approach (Kinnebrew & Biswas, 2011; Kinnebrew, Biswas, Sulcer, & Taylor, 2012; Segedy et al., 2013). For example, agent technology has been used to help students master foreign languages (Jia & Ruan, 2008) and improve job interviewing skills, and help children with autism develop important social skills (Hoque, Curgeon, Martin, Mutlu, & Picard, 2013).

Given the need for teacher candidates to practice communication skills during critical moments, like when they suspect a bullying incident occurred, yet being mindful of avoiding causing harm to actual children in the process, programming a chatbot became an attractive solution for developing fluency in important communication skills. Our purpose was to create a virtual role-play whereby teacher candidates could communicate with a programmed chatbot and feel as though they were conversing with an ac-

tual child. The authenticity of the VRP would rely on the appropriateness of responses the chatbot gave to the user, such as becoming “upset” when the user failed to respond appropriately. The project involved two main phases: 1) Developing the technology for the VRP, 2) Evaluating its effectiveness. The purpose of this exploratory study is to examine whether practice in simulated conversations provided by VRP improves teacher candidates' communication skills namely by improving their fluency in responding to bullying.

METHODS

Participants

Our participants included 27 undergraduate teacher candidates enrolled in two classroom management courses for elementary education majors at a large, public university in the northeast. Participants were all juniors and seniors, with the exception of one sophomore. There were 25 females and 2 males. With the exception of one returning student in her thirties, all participants were traditional college-aged students between the ages of 19-21. We randomly assigned two intact classrooms to one of two groups: A) VRP ($N=14$): practiced skills with virtual role-play, or B) RP (role-play) ($N=13$): practiced skills through reflecting on the case study. Both classes were taught by the same instructor, who was also one of the researchers for the study.

Data Collection

The corpus of data included two videos (pre- and post-) from each participant, and one reflection paper from each participant in the VRP. Before receiving instructional content related to communication or bullying, participants in both courses read a case about Eli, a 4th grader who struggled in school socially and academically and was mostly a loner. Participants learn from the case, that although some students got along with Eli when they were with him one-on-one, Eli tended to be the target of teasing when he was around students in groups. We intentionally wanted to incorporate relational bullying since this poses more challenges for teachers (Voulgaridou & Kokkinos, 2015). In the case, the teacher, Mrs. Fischer, walks across the playground during recess and hears Eli, who appears to be throwing a tantrum. He is near three other students who appear to be laughing at him.

After reading the case, participants in both groups thought about how they would respond. In groups of three, participants engaged in traditional role-play in which they simulated how they would talk to a bully-victim in their class. One person played Eli, one played Mrs. Fischer, and one used a smartphone to video record the role-play. Roles were rotated until each participant had an opportunity to serve in each role. We labeled these the “pre-instruction” videos. The purpose of the pre-instruction videos was to establish a baseline for each group and to determine if there were differences between the VRP group and the RP only group.

Following instruction on bullying and communication (described below) and additional practice or reflection, each participant videotaped a second role-play performance based on the same case study as the first role-play. We labeled these the “post-instruction” videos. All videos were transcribed. All participants wrote a final reflection: VRP compared their experience using the virtual role-play with the in-class role-play, and the RP group compared their first and second in-class role-play performances. The final dataset analyzed included 12 videos for VRP and 11 videos for the RP group. Two videos from each group were excluded from the analysis due to limitations in video data quality. See Figure 1.

	Before Instruction	Instruction	Practice	After Instruction
RP	Read Eli case <i>Pre- Instruction videos (n=11)</i>	Bullying, NVC	Reflection on case	<i>Post- Instruction videos (n=11)</i> Reflection comparing RP videos
VRP	Read Eli case <i>Pre- Instruction videos (n=12)</i>	Bullying, NVC	VRP	<i>Post- Instruction videos (n=12)</i> Reflection comparing VRP, RP (n=13)

Figure 1. Study Timeline & Data. Data that were analyzed for this study are italicized.

Bullying instruction. Both groups received the same instruction in how to use the principles of Non-Violent Communication (NVC) (Rosenberg, 2003) and bully prevention (Olweus, 1993; Rigby & Griffiths, 2010) to mediate student conflict around bullying. Instruction was delivered by the same study investigator on both occasions. NVC is designed to prevent and manage potentially volatile situations by helping individuals make observations rather than judgments, pinpoint unfulfilled needs, and develop empathy (Rosenberg, 2003). Instruction on bullying included information about the types of bullying that occur—verbal, physical, social exclusion—negative effects of bullying for everyone involved, and, most importantly, a communication protocol for addressing a child suspected of being bullied. One

of the study investigators developed the protocol by adapting the National Institute of Child Health and Human Development's protocol for the investigative interviewing of children (see Lamb, La Rooy, Malloy, & Katz, 2011) and incorporating principles from the evidence-based "Steps to Respect" and "Method of Shared Concern" anti-bullying programs (Committee for Children, 2001; Rigby & Griffiths, 2010) and the [stopbullying.gov](http://www.stopbullying.gov) website (<http://www.stopbullying.gov/respond/index.html>). The bullying communication protocol includes four main domains: develop trust with the child, teach description, discuss the problematic event, assure child by discussing next steps. The protocol is simultaneously prescriptive and flexible. This is to provide guidance on what to say and when to say it while also allowing the teacher to take into account the context of the specific situation. Both classes had equal exposure to instructional material.

Virtual role-play intervention. We designed a VRP pedagogical agent, virtual "Eli," that could simulate a student who had been bullied. By leveraging a self-paced approach and opportunities for sustained practice afforded by virtual role-play, we intended to provide teacher candidates with targeted practice to increase their fluency when intervening in bullying situations. The "brain" of Eli was constructed to respond with anger or to withdraw when teacher candidates interacted with Eli in ways that were problematic. It was also constructed to respond positively, by being willing to communicate, expressing relief, or thanking the teacher when teacher candidates interacted according to principles of effective communication they were taught¹.

Teacher candidates in VRP communicated with virtual Eli for approximately 1+ hours after receiving classroom instruction on bullying and communication (see Appendix B for sample transcripts). They were to engage in at least three conversations with virtual Eli, from an opening greeting to ending the conversation. Interactions with VRP were text-based, with the users typing in their reply as if they were Mrs. Fischer, the teacher, and the chatbot's responses appearing onscreen. The chatbot responds to each sentence/question as if they were independent of each other, which sometimes can result in a less cohesive dialogue. Teacher candidates were to review their simulation transcripts. Instead of using VRP, candidates in the RP did a more extensive case study analysis as homework, which took approximately one hour to complete. Teacher candidates in both groups wrote a reflective essay. The VRP candidates compared doing in-person role-play with virtual role-play. The RP participants compared their second in-person role-play performance with their first role-play performance.

1 For more information on the development of virtual Eli, see Frank, Lee, Schussler, Wright, Yu (in progress).

Data Analysis

One PI (also the course instructor) and one research assistant, who was blind to condition, coded the quality of teacher candidate communication during live role-play activities. The coding protocol for the video analysis consisted primarily of the bullying communication protocol described above as our purpose was to build fluency in these communication skills. The protocol includes four domains (i.e., Trust Building, Teach Description, Discussing the Problematic Event, Assuring the Child) each of which contains about 3 indicators. For example, the indicators in the domain “Assuring the Child” include: 1) provide empathy and find solution with child’s involvement, 2) inform of future discussions with the other children involved, 3) thank the child for talking. For a complete list of all indicators in each domain, see Tables 1-5. In addition to the four domains specific to the bullying protocol, there were two additional domains: 1) General Principles, which were criteria for communicating about bullying that could have occurred anytime in the role-play, 2) Nonverbal and Other Communication Issues. Both of these domains included indicators of favorable criteria we wanted to see as well as indicators of unfavorable criteria we did not want to see. For example, in the “General Principles” there were 5 favorable indicators including: 1) convey respect for the student’s confidentiality, and 2) get the story from those involved. There were 6 unfavorable indicators including: 1) talking to the children who were involved all together, 2) trying to immediately sort out the facts. The “Nonverbal and Other Communication Issues” domain included 2 favorable indicators and 4 unfavorable indicators.

Because this is an exploratory study investigating whether use of VRP improves communication fluency, we chose to conduct a content analysis measuring the communication indicators described above. Indicators were coded in the following manner:

Favorable=favorable indicator present

Mix=both favorable and unfavorable evidence of indicator present

Unfavorable=unfavorable indicator present OR opposite of favorable indicator present

N/A=there was not evidence of the indicator present

It should be noted that this coding system meant the unfavorable indicators could only be coded as a “Unfavorable,” “Mix,” or “N/A” (see Tables 1-4). In addition, the coders read and coded de-identified written transcripts first, before watching videos. This was to minimize potential bias on the part of the instructor who was aware of the groups assigned to the students.

Table 1
Virtual Role-Play Pre-Video Coding

Domains	Indicators	Favorable ¹	Mix ²	Unfavorable ³	N/A ⁴
General Principles	+Separate kids involved	1	0	0	11
	+Ensure everyone's safety (meet any medical/ mental health needs)	0	0	0	12
	+Stay calm and respectful	12	0	0	0
	+Convey respect for confidentiality	1	0	4	7
	+Get the story from those involved	4	0	0	8
	-Talking to kids involved together	0	1	3	8
	-Try to immediately sort out facts	0	0	5	7
	-Attempt conflict resolution; get kids to apologize	0	0	1	11
	-Force kids to speak publicly	0	0	1	11
	-Blaming; judging; mention bullying	0	0	5	7
-Tell to ignore the bullying	0	0	0	12	
Non-verbals & Other Communication Issues	+Makes eye contact	12	0	0	0
	+Positive touching or posture	11	0	0	1
	-Giggling	0	0	3	9
	-Distracted by others; loud noises around	0	0	0	12
	-Reading off paper; robotic, scripted tone	0	0	0	12
	-"Eli" leading conversation	0	0	0	12
Interview Protocol					
Trust Building	+Explain role of teacher	1	0	0	11
	+Show support	5	0	0	7
	+Explain how to tell truth [kg-gr.3]	0	0	0	12
Teach Description	+Explain opinion vs. fact	0	0	0	12
	+Can Mrs. Fischer tell me about event?	0	0	0	12
Discussing Problematic Event	+Gather information:[Where; When; Who was involved/ What was specifically said, done]	11	0	0	1
	+Determine patterns to bullying	2	0	0	10
	+Use target's own words	9	0	0	3
Assuring Child	+Empathy, find solution with child	6	1	0	5
	+Inform discussion future discussion	7	0	0	5
	+Thank child	1	0	0	11

¹Favorable indicator present; ²Both favorable and unfavorable evidence of indicator present; ³Unfavorable indicator present OR opposite of favorable indicator present; ⁴There was not evidence of the indicator present

Table 2
Role-Play Pre-Video Coding

Domains	Indicators	Favorable ¹	Mix ²	Unfavorable ³	N/A ⁴
General Principles	+Separate kids involved	2	1	1	7
	+Ensure everyone's safety (meet any medical/ mental health needs)	2	0	0	9
	+Stay calm and respectful	8	1	0	2
	+Convey respect for confidentiality	0	0	0	11
	+Get the story from those involved	1	0	0	10
	-Talking to kids involved together	0	1	2	8
	-Try to immediately sort out facts	0	0	2	9
	-Attempt conflict resolution; get kids to apologize	0	0	1	10
	-Force kids to speak publicly	0	0	0	11
	-Blaming; judging; mention bullying		0	4	7
	-Tell to ignore the bullying	0	0	1	10
Non-verbals & Other Communication Issues	+Makes eye contact	11	0	0	0
	+Positive touching or posture	8	3	0	0
	-Giggling	0	3	2	6
	-Distracted by others; loud noises around	0	0	1	10
	-Reading off paper; robotic, scripted tone	0	0	0	11
	-“El!” leading conversation	0	0	0	11
Interview Protocol					
Trust Building	+Explain role of teacher	1	0	0	10
	+Show support	5	0	0	6
	+Explain how to tell truth [kg-gr.3]	0	0	0	11
Teach Description	+Explain opinion vs. fact	0	0	0	11
	+Can Mrs. Fischer tell me about event?	0	0	0	11
Discussing Problematic Event	+Gather information:[Where; When; Who was involved/ What was specifically said, done]	5	1	0	5
	+Determine patterns to bullying	0	0	0	11
	+Use target's own words	4	0	0	7
Assuring Child	+Empathy, find solution with child	0	0	1	10
	+Inform discussion future discussion	2	0	0	9
	+Thank child	0	0	0	11

¹Favorable indicator present; ²Both favorable and unfavorable evidence of indicator present; ³Unfavorable indicator present OR opposite of favorable indicator present; ⁴There was not evidence of the indicator present

Table 3
Virtual Role-Play Post-Video Coding

Domains	Indicators	Favorable ¹	Mix ²	Unfavorable ³	N/A ⁴
General Principles	+Separate kids involved	1	0	0	11
	+Ensure everyone's safety (meet any medical/ mental health needs)	4	0	0	8
	+Stay calm and respectful	11	1	0	0
	+Convey respect for confidentiality	12	0	0	0
	+Get the story from those involved	10	1	0	1
	-Talking to kids involved together	0	0	0	12
	-Try to immediately sort out facts	0	0	3	9
	-Attempt conflict resolution; get kids to apologize	0	0	0	12
	-Force kids to speak publicly	0	0	0	12
	-Blaming; judging; mention bullying	0	0	5	7
-Tell to ignore the bullying	0	0	1	11	
Non-verbals & Other Communication Issues	+Makes eye contact	9	2	0	1
	+Positive touching or posture	5	2	0	5
	-Giggling	0	0	4	8
	-Distracted by others; loud noises around	0	0	1	11
	-Reading off paper; robotic, scripted tone	0	0	2	10
	-"Eli" leading conversation	0	0	0	12
Interview Protocol					
Trust Building	+Explain role of teacher	2	0	0	10
	+Show support	10	0	0	2
	+Explain how to tell truth [kg-gr.3]	0	0	0	12
Teach Description	+Explain opinion vs. fact	1	0	0	11
	+Can Mrs. Fischer tell me about event?	0	0	0	12
Discussing Problematic Event	+Gather information:[Where; When; Who was involved/ What was specifically said, done]	11	0	0	1
	+Determine patterns to bullying	5	0	0	7
	+Use target's own words	8	0	0	4
Assuring Child	+Empathy, find solution with child	7	1	0	4
	+Inform discussion future discussion	11	0	0	1
	+Thank child	1	0	0	11

¹Favorable indicator present; ²Both favorable and unfavorable evidence of indicator present; ³Unfavorable indicator present OR opposite of favorable indicator present; ⁴There was not evidence of the indicator present

Table 4
Role-Play Post-Video Coding

Domains	Indicators	Favorable ¹	Mix ²	Unfavorable ³	N/A ⁴
General Principles	+Separate kids involved	0	0	0	11
	+Ensure everyone's safety (meet any medical/ mental health needs)	3	0	0	8
	+Stay calm and respectful	11	0	0	0
	+Convey respect for confidentiality	2	0	0	9
	+Get the story from those involved	9	0	0	2
	-Talking to kids involved together	0	0	1	10
	-Try to immediately sort out facts	0	0	1	10
	-Attempt conflict resolution; get kids to apologize	0	0	0	11
	-Force kids to speak publicly	0	0	0	11
	-Blaming; judging; mention bullying	0	1	1	9
-Tell to ignore the bullying	0	0	0	11	
Non-verbals & Other Communication Issues	+Makes eye contact	1	9	1	0
	+Positive touching or posture	3	0	0	8
	-Giggling	0	0	1	10
	-Distracted by others; loud noises around	0	0	0	11
	-Reading off paper; robotic, scripted tone	0	0	9	2
	-"Eli" leading conversation	0	0	0	0
Interview Protocol					
Trust Building	+Explain role of teacher	8	2	0	1
	+Show support	10	0	0	1
	+Explain how to tell truth [kg-gr.3]	0	0	0	11
Teach Description	+Explain opinion vs. fact	3	0	0	8
	+Can Mrs. Fischer tell me about event?	0	0	0	11
Discussing Problematic Event	+Gather information:[Where; When; Who was involved/ What was specifically said, done]	8	0	0	3
	+Determine patterns to bullying	8	0	0	3
	+Use target's own words	6	0	0	5
Assuring Child	+Empathy, find solution with child	5	1	0	5
	+Inform discussion future discussion	9	1	0	1
	+Thank child	3	0	0	8

¹Favorable indicator present; ²Both favorable and unfavorable evidence of indicator present; ³Unfavorable indicator present OR opposite of favorable indicator present; ⁴There was not evidence of the indicator present

There were 46 usable videos, 23 participants with both pre- and post-videos. The two coders followed an accepted procedure to establish reliability (see Scott, 1955). First, they coded three videos separately, compared the reliability of codes and conferred about any codes that were not in agreement. They added these decisions to a codebook. Then coders completed two additional cycles of coding, each time refining the codebook to reflect nuances of distinction. For example, the indicator “inform of future discussions with the other individuals involved and future discussion with all children” included two things: 1) letting the child know Mrs. Fischer would talk to the other children, 2) letting the child know Mrs. Fischer would be following up. It was decided that if the participant did only one, we would code it as the favorable indicator was present. After three rounds of coding and comparing, for a total of seven videos, the coders coded 19 videos separately and then compared codes. Using a consensus estimate for establishing reliability was appropriate given the data are “nominal in nature and different levels of the rating scale represent qualitatively different ideas” (Stemler, 2004). Coders were in agreement 92.1% of the time, which is well within the range of acceptability (Neuendorf, 2002). Given the reliability of coding, the remaining 20 videos were divided equally between the two coders.

The aggregated results of the coding are presented in Tables 1-4 which show the VRP Pre-Instruction, RP Pre-Instruction, VRP Post-Instruction, and RP Post-Instruction data. Indicators preceded with a “+” are favorable indicators that participants should be doing, while those preceded with a “-” are unfavorable indicators that we do not want to see. To make sense of these data, three different comparisons were conducted. First, to establish differences at baseline, we compared differences between the VRP Pre- and the RP Pre- videos. Second, to examine pre- and post- differences between the two groups, we compared VRP Pre- videos to VRP Post- as well as RP Pre- to RP Post-. Third, we wanted to see if there were substantive differences between the VRP Post- and RP Post- as this would help us better understand the value-added of VRP in addition to instruction in bullying and communication. We looked for differences of more than 2 in each cell focusing especially on the “Favorable” and the “Unfavorable” columns, as these columns showed the presence of either favorable or unfavorable indicators.

Reflections where VRP participants compared the use of VRP with traditional role-play (RP) were open-coded by two members of the research team. To develop the codes, the researchers focused on teacher candidates' perceptions of the benefits and drawbacks of VRP and RP and which aspects of either pedagogy helped them feel more confident about dealing

with bullying. The reflections of the RP group were not coded since they had experience with RP only. Twenty-three categories in an initial list were collapsed into 14 categories and included: realistic, unrealistic, no time to think, think time, felt anxious, and others. To capture whether comments referred to benefits or drawbacks of VRP, RP or both, categories were developed to capture this, including: VRP Benefits, VRP Drawbacks, VRP Neutral Observation, RP Benefits, RP Drawbacks, RP Neutral Observation, Neutral Observation both, Direct Comparison. Once the categories were established one researcher coded all reflections in NVivo 10 to capture both the content of what was said and the pedagogy to which it referred (i.e., RP or VRP). See Table 5.

Table 5
Written Reflections Content Analysis

	VRP Benefits	VRP Drawbacks	RP Benefits	RP Drawbacks	Neutral Observation of both
Realistic	6	0	3	0	1
Not realistic	0	0	0	1	0
Technology problems	0	0	1	0	0
Conversations too similar	0	1	0	0	0
Limited responses	0	1	0	0	0
Wording question well	0	2	1	0	0
No time to think; fast-paced	0	0	1	1	0
Think time	1	2	0	0	0
Felt Anxious	0	3	0	1	0
Practice	3	0	0	1	3
Improved over time	1	0	0	0	0
Better questions	3	0	0	0	1
Focused on right things	5	0	0	1	2
Other	2	1	0	3	6

RESULTS

Videos

Analysis 1: Baseline data. At baseline, there were a few differences between the VRP and RP pre-instruction videos. In the “General Principles” domain the VRP participants showed more negative indicators for “Convey respect for confidentiality” (VRP 4, RP 0) and “Try to sort out facts immediately” (VRP 5, RP 2). However, VRP participants scored better in “stay calm and respectful” (VRP 12, RP 8) though substantively this difference was fairly minor. None of the participants were openly disrespectful toward Eli though some appeared to be self-conscious during the role-play. This was also evident in the analysis of the “Non-verbal” domain. Three RP participant videos were coded in the “Mix” column, which indicates favorable and unfavorable evidence of an indicator, while all but one of the VRP videos were coded as favorable. In terms of the domains from the interview protocol, the participants were fairly close across the indicators with the exception of three. The VRP group at baseline did a better job of attempting to “Gather information” with 11 favorable ratings compared to 5 for the RP participants. They also were more likely to “Use the child’s own words.” (VRP 9, RP 4). VRP participants also were more likely to “show empathy and find solution with child” (VRP 6, RP 0) “inform child of future discussion” (VRP 7, RP 2). This means that prior to instruction regarding the protocol, the VRP group was already engaging in some effective aspects of communication when trying to handle an incident of potential bullying. It also indicates that the RP participants had more room for improvement in these areas than the VRP participants.

Analysis 2: Changes from pre- to post-. We were especially interested in the changes from pre- to post- for both groups. These changes are summarized in Table 6 where “Fav” is a change in the “Favorable” column and “Unfav” is a change in the “Unfavorable” column. Numbers with a “+” indicate they increased from pre- to post- while numbers with a “-” indicate a decrease. A “0” indicates no change. Ideally, favorable indicators should increase and unfavorable “ should decrease. For the VRP there were substantial increases in some of the indicators in the “General Principles” domain including “Convey respect for confidentiality” (+11), “Get the story from those involved” (+6), and “Ensure everyone’s safety” (+4). All domains from the interview protocol saw at least one indicator with positive change, including “Show support” (+5), “Determine patterns to bullying” (+3) and “Inform of future discussions” (+4). The VRP

participants also showed decreases in some negative indicators. For example, some were less likely to “talk to kids involved together” (-3) and try to “immediately sort out facts” (-2) before ensuring the safety of the child.

Table 6
Change in Pre- to Post-Videos

Domains	Indicators	VRP Change		RP Change	
		Fav	Unfav	Fav	Unfav
General Principles	+Separate kids involved	0	0	-2	-1
	+Ensure everyone’s safety (meet any medical/ mental health needs)	+4	0	+1	0
	+Stay calm and respectful	-1	0	+3	0
	+Convey respect for confidentiality	+11	-4	+2	0
	+Get the story from those involved	+6	1	+8	1
	-Talking to kids involved together	0	-3	0	-1
	-Try to immediately sort out facts	0	-2	0	-1
	-Attempt conflict resolution; get kids to apologize	0	-1	0	-1
	-Force kids to speak publicly	0	-1	0	0
	-Blaming; judging; mention bullying	0	0	0	-3
	-Tell to ignore the bullying	0	-1	0	-1
Non-verbals & Other Communication Issues	+Makes eye contact	-3	0	-10	+1
	+Positive touching or posture	-6	0	-5	0
	-Giggling	0	+1	0	-1
	-Distracted by others; loud noises around	0	0	0	+1
	-Reading off paper; robotic, scripted tone	0	+2	0	+9
	-“Eli” leading conversation	0	0	0	0
Interview Protocol					
Trust Building	+Explain role of teacher	+1	0	+7	0
	+Show support	+5	0	+5	0
	+Explain how to tell truth [kg-gr.3]	0	0	0	0
Teach Description	+Explain opinion vs. fact	+1	0	+3	0
	+Can Mrs. Fischer tell me about event?	0	0	0	0
Discussing Problematic Event	+Gather information:[Where; When; Who was involved/ What was specifically said, done]	0	0	+3	0
	+Determine patterns to bullying	+3	0	+8	0
	+Use target’s own words	-1	0	+2	0
Assuring Child	+Empathy, find solution with child	+1	0	+5	-1
	+Inform discussion future discussion	+4	0	+7	0
	+Thank child	0	0	+3	0

There were also some changes in the wrong direction for the VRP participants, most of these which fell in the “Nonverbals and other Communication issues” domain. There was a decrease in the number of participants who showed “positive touching or posture” (-6) and “made eye contact” (-3) as well as an increase in the number of students who “read off their paper” (+2) and were “giggling” (+1). Some indicators that showed no change or very little change were due to a ceiling effect, like “gather information” (0), “use target’s own words” (-1), and “show empathy and find solution with child” (+1).

The RP only participants experienced some positive changes from pre- to post- especially in some domains of the interview protocol. In the “trust building” domain they showed improvement in “explaining the role of the teacher” (+7) and “show support” (+5). In “Discussing the Problematic Event” they showed improvements in all areas, especially, “determine whether there are patterns to bullying” (+8). They also experienced increases in all indicators of the “Assuring the Child” domain with “inform of future discussions” having the largest improvement (+7). These were consistent with some of the changes in the “General Principles” domain where there were increases in positive indicators like participants’ explicit efforts to “get the story from those involved” (+8) and a decrease in negative indicators like “blaming; judging; or explicitly mentioning bullying” (-3). The changes in the wrong direction occurred predominantly in the “Nonverbals and other Communication issues” domain where ten participants experienced a decrease in their “making good eye contact” and five fewer used “positive touching or posture”. In addition, nine more participants at post were “reading off their paper.”

Analysis 3: Comparison of VRP to RP. Since there were some baseline differences, including ceiling effects for the VRP group, it was important both to consider how the VRP post- (Table 3) compared to the RP post- (Table 4) and to examine the comparison of VRP change vs. RP change (Table 6). We assumed that the favorable indicators would increase more and the unfavorable criteria would decrease more for the VRP while the RP would experience less change. Since the VRP and RP groups did not have exactly the same number of participants, we describe the cells of Tables 3 and 4 in percentages for the most precise comparisons.

On a number of indicators in the post- role-plays, participants in both VRP and RP performed about equally. These included favorable criteria such as “staying calm and respectful,” “getting the story from those involved,” and “building trust and showing support.” The two groups also performed about equally in avoiding unfavorable criteria, such as “talking

to kids together,” “attempting to get kids to apologize,” and “telling kids to ignore the bullying.”

There were some important differences in a number of indicators in the post-instruction videos. Teacher candidates who practiced with VRP did a moderately better job in the “Assuring the Child” domain. Specifically, they were more likely to “show empathy,” “find solution with child” (58% VRP vs. 45% RP) and “inform of discussion with other students and a future discussion with all involved” (92% virtual role-play vs. 82% non-virtual role-play). They also performed better in the domain “Discussing the Problematic Event” domain. Ninety-two percent of teacher candidates using VRP asked appropriate questions to determine what had happened to the child whereas only 73% of the other teacher candidates did so. However, it should be noted that VRP participants were equally as high in this indicator on their pre-instruction videos. As mentioned previously, there were similar ceiling effects for a number of indicators (e.g., “inform of discussion with other students and a future discussion with all involved” (7), “use target’s own words” (9), “positive touching or posture” (11), “makes good eye contact” (12) “stay calm and respectful” (12)) whereby there was not much room for improvement.

The teacher candidates who just did RP scored better on some indicators than the candidates who practiced with VRP. For example, far more participants in RP only (73%) explained their role than participants who practiced with VRP (17%). They were also more likely to “determine patterns to bullying” (73% RP vs. 42% VRP), and less likely to “blame or judge the potential victim or mention ‘bullying’” (9% RP vs. 42% VRP).

One of the most striking differences involved eye contact (or lack thereof) and the use of notes during the post- role-play. Only 9% of teacher candidates who had RP only received a favorable rating for “makes eye contact” while 75% of participants who practiced with VRP received a favorable rating on this indicator. This finding was highly correlated with 82% of RP participants reading off their notes whereas only 17% of VRP participants did the same. Despite less reliance on their notes, we also noted that candidates who received VRP practice demonstrated more refined word choice that more closely followed the communication protocol taught in class. This finding suggests that VRP participants more readily internalized the communication techniques as they did not require the continued aid of the written protocol at post.

Written Reflections

Content analysis of written reflections suggests that the VRP practice did help teacher candidates very consciously choose their words, though some also cited this as a drawback of the activity as sometimes the virtual child did not respond in what seemed to be an appropriate manner. As can be seen in Appendix B, one of Virtual Eli's default responses when it was confused about the input was, "I have no answer for that." The top benefits of VRP candidates cited were that it was realistic (n=6), it helped them to focus on the right things (n=5), they asked better questions (n=3), and they appreciated being able to practice multiple times (n=3). Ingrid² describes a few of these in her response:

While Virtual Role Play did give me some frustrations in the beginning, trying different methods of talking to Virtual Eli helped me to understand better how it works. Being able to talk to Virtual Eli multiple times was a lot of help for me.... In my beginning sessions, I asked Virtual Eli simple questions of what happened and why and what we can do to fix the problem. By my final sessions, I was asking more questions on how he feels and how I can help him feel safe, which are things we talked about in class. Having the opportunity to talk to Virtual Eli as many times as I wanted was most interesting to me because I found myself reacting differently emotionally every time the conversation began steering in a different direction.

Some said interacting with Virtual Eli was more realistic than having their friend imitate a middle school student. Specifically, the responses let them know when they said something problematic, while their traditional role-play partner typically did not provide this kind of immediate feedback.

In terms of VRP drawbacks, some said it made them nervous (n=3), it was more difficult to word their questions well (n=2), or that they felt more pressure to not "say the wrong thing" because they had more time (n=2). In addition one person said there was a lack of depth of responses from Virtual Eli, and one said the conversations were too similar and not helpful.

Teacher candidates thought a benefit of regular role-play was that it was in real time, though one teacher candidate said they felt awkward and self-conscious and one said there was not time to think about what she was going to say. Lexi said, "While being filmed, I had to react to what Eli said right away, and I felt like I said certain things that I did not think were helpful to him. However, during the virtual dialogue, I had much more time to

2 All names are pseudonyms.

consider what Eli said and I could carefully decide how I was going to answer him.” When asked explicitly if practicing with VRP multiple times precipitated a change in the quality of their interactions, 9 said yes, 2 said no, 2 neglected to respond to this question in their reflection. Teacher candidates also noted that having a protocol for what to say and how to say it was helpful in knowing how to communicate effectively with a child and gather appropriate information about what happened.

DISCUSSION

The purpose of this study was to examine how practice with VRP affected teacher candidates’ communication in a situation involving relational bullying. Specifically, we wanted to know whether candidates who engaged with traditional RP and additional practice in VRP demonstrated more fluency compared to candidates who engaged with traditional RP but whose additional practice consisted of reflection. We also wanted to know how teacher candidates compared their experience using VRP with their experience engaging in traditional RP.

After coding teacher candidates’ pre- and post-videos using a bullying communication protocol, our results appear to be somewhat mixed. Teacher candidates in traditional RP experienced greater gains on some indicators like “determining patterns to bullying” and “explaining the role of the teacher.” However, the ceiling effects in some of the indicators at baseline made it more difficult for the candidates who practiced with VRP to show the same kinds of gains. Despite some of the gains not being as great, there were more candidates in the VRP group who followed some aspects of the protocol than the candidates in the RP only group. Most notably, these were “gathering information about the bullying,” “using target’s own words,” “inform of discussion with other students and a future discussion with all involved.” There were also many more VRP candidates who “conveyed respect for confidentiality,” “made good eye contact,” and did not engage in the negative indicator of “reading off their paper.” This final indicator is significant. The students who used the communication protocol or their notes during their post- role-plays, essentially had the information they were to follow right in front of them. It would be like having the answers to a test while taking the test. Those who did not use notes relied on their knowledge and comfort level with the protocol. In addition, VRP participants indicated in their reflections that they considered their word choice more carefully.

That students who practiced with VRP had a more refined word choice, maintained better eye contact, and did not rely on their notes during regular

role-play suggests that their comfort level in having a conversation with a child who likely has been bullied is probably greater than those whose practice consisted of reflection. One possible explanation is that repeated practice afforded by VRP builds candidates' fluency in a way that traditional role-play can not, especially given the time and logistical constraints for conducting regular role-play. Teachers need to know not only what kinds of questions to ask, but how to phrase questions and probe for information in a manner that is objective and supportive (Rigby & Griffiths, 2010). VRP may help teacher candidates develop these skills both by allowing for repetition of the same case-based dialogue and by compelling the users to refine their word choice through chatbot responses. Although candidates who practiced with VRP became frustrated when the chatbot defaulted to a particular response (aka, "I have no answer for that"), it did force them to re-think their word choice and sometimes the content of what they said. The chatbot also responded in ways that sometimes surprised the candidates, also forcing them to re-think how they responded.

It is also possible that the VRP group's comfort level with the communication protocol, as demonstrated in their overall lack of reliance on their notes and refined word choice, lead to a greater sense of both efficacy and empathy. Although we did not measure empathy or efficacy directly for this study, there is some evidence in the VRP participant's written reflections that VRP influenced both. As stated above, Ingrid noted that she found herself "reacting differently emotionally every time the conversation began steering in a different direction" and by her final sessions she was "asking more questions on how [Eli] feels and how I can help him feel safe." Relational bullying, the kind Eli experienced in the case study, is the most difficult to identify and to intervene effectively (Voulgaridou & Kokkinos, 2015). However, empathy is a significant predictor of teacher candidates labeling witnessed social exclusion as bullying and labeling and intervening in verbal aggression (Craig, Henderson, et al., 2000), both of which occurred in the Eli case study. That some teacher candidates articulated feeling more empathetic as a result of interacting with VRP is a finding that warrants further investigation. Efficacy is also an important predictor determining whether teachers respond to bullying (Yoon, 2004). VRP may be helpful in both building empathy so teachers know when to identify students' interactions as bullying and also fostering communication fluency so they feel confident and have the skills to respond effectively.

Given some of the findings of this exploratory study, we suggest a few directions for future research and refinement of the VRP technology. First, given the important relationship between empathy, efficacy and interven-

tion, additional research should investigate how practice with VRP influences teachers' empathy, efficacy, and fluency. Second, we speculate that the VRP would have added pedagogical benefit if candidates reflected on the content and nuance of their VRP dialogues and how they could improve. Although we asked candidates to review their transcripts, we did not require any formal reflection whereby they had to identify places their communication skills were strong or problematic. Adding more targeted reflection coupled with instructor feedback could make the intervention even more robust. Finally, finding ways to create a virtual "Eli" with enhanced qualities that mimic an actual child, either through voice or animation, may help minimize the absence of any nonverbal cues and increase the authenticity.

There were some limitations to this study. As already stated, there were some baseline differences between the VRP and RP groups that created some ceiling effects for some of the indicators for the VRP teacher candidates, making the comparisons between them more challenging. Relatedly, the students were not randomly assigned to practice with VRP or not, but rather were part of intact classes. Class differences could have affected the outcomes on the role-play videos. In addition, the size of each class was relatively small, and each class lost some video footage, either due to student absence or video that couldn't be coded due to poor quality. Therefore, conclusions about the effectiveness of VRP should be taken with some caution.

CONCLUSION

Teacher candidates entering the field can never be completely prepared for every situation they will encounter. The more practice they receive, especially for situations that are critical to the safety of the students in their care, such as bullying, the more likely they will be to act effectively and efficaciously. Presenting students with not only information about bullying, but also a protocol for responding to bullying, was helpful in communicating effectively. In addition, practice time that gave them opportunities to think was also helpful.

Overall, supplementing or replacing traditional role-play with virtual role-play, may help overcome the limitations of traditional role-play like lack of time and relying on the quality of teacher candidates' acting skills. It also affords more students the ability to engage in such practice and to capture the simulated conversation for review and feedback later. Virtual role-play can also create a more developmental approach to building teacher candidates' communication skills as candidates can engage in the same scenario

multiple times in an attempt to improve their previous performance. Virtual role-play shows promise as an innovative technology for teaching targeted communication skills.

Acknowledgement

The authors would like to thank the TLT Studio for development and support of the first iteration of the Eli chatbot, as well as Brian Shook (code), Serena Epstein (design), and Emily Chukusky (content). The authors also would like to thank the reviewers for their helpful feedback on this manuscript.

References

- American Educational Research Association. (2013). *Prevention of bullying in schools, colleges, and universities: Research report and recommendations*. Retrieved from Washington, DC:
- Committee for Children. (2001). *Steps to Respect Program Guide*. Seattle, WA: Second Step.
- Cornell, D., Gregory, A., Huang, F., & Fan, X. (2013). Perceived prevalence of bullying and teasing predicts high school dropout rates. *Journal of Educational Psychology, 105*, 138-149.
- Craig, W. M., Henderson, K., & Murphy, J. G. (2000). Prospective teachers' attitudes toward bullying and victimization. *School Psychology International, 21*(1), 5-21.
- Craig, W. M., & Pepler, D. J. (1997). Observations of bullying and victimization in the school yard. *Canadian Journal of School Psychology, 13*(2), 41-59.
- Craig, W. M., Pepler, D. J., & Atlas, R. S. (2000). Observations of bullying in the playground and in the classroom. *School Psychology International, 21*(1), 22-38.
- Espelage, D. L., Low, S., & De La Rue, L. (2012). Relations between peer victimization subtypes, family violence, and psychological outcomes during early adolescence. *Psychology of Violence, 2*(4), 313-324.
- Espelage, D. L., & Swearer, S. M. (2003). Research on school bullying and victimization: What have we learned and where do we go from here? *School Psychology Review, 32*(3), 365-384.
- Girod, M., & Girod, G. R. (2008). Simulation and the need for practice in teacher preparation. *Journal of Technology and Teacher Education, 16*(3), 307-337.
- Holt, M. K., & Keys, M. A. (2004). Teachers' attitudes toward bullying. In D. Espelage & S. M. Swearer (Eds.), *Bullying in American schools: A social-*

- ecological perspective on prevention and intervention* (pp. 121-140). Mahwah, NJ: Erlbaum.
- Hoque, M., Courgeon, M., Martin, J.-C., Mutlu, B., & Picard, R. W. (2013). *MACH: My Automated Conversation coach*. Paper presented at the UbiComp'13, Zurich, Switzerland.
- Janson, G. R., Carney, J. V., Hazler, R. J., & Oh, I. (2009). Bystanders' reactions to witnessing repetitive abuse experiences. *Journal of Counseling and Development, 87*(3), 319-326.
- Jeffries, P. (2005). A FRAMEWORK for designing, implementing, and evaluating simulations used as teaching strategies in nursing. *Nursing Education Perspectives, 26*(2), 95-103.
- Jia, J., & Ruan, M. (2008). Use chatbot CSIEC to facilitate the individual learning in English instruction: A case study. In B. Woolf et al. (Ed.), *Proceedings of the International Conference on Intelligent Tutoring Systems, 2008* (pp. 706-708). New York, NY: Springer.
- Kinnebrew, J., & Biswas, G. (2011). Modeling and measuring self-regulated learning in teachable agent environments. *Journal of e-Learning and Knowledge Society, 7*(2), 19-35.
- Kinnebrew, J., Biswas, G., Sulcer, B., & Taylor, R. (2012). Investigating self-regulated learning in teachable agent environments. In R. A. V. Alevén (Ed.), *International Handbook of Metacognition and Learning Technologies* (Vol. 99). New York, NY: Springer.
- Lamb, M. E., La Rooy, D. J., Malloy, L. C., & Katz, C. (Eds.). (2011). *Children's testimony: A handbook of psychological research and forensic practice* (2nd ed.): John Wiley & Sons.
- Mehta, S., Cornell, D., Fan, X., & Gregory, A. (2013). Bullying climate and school engagement in ninth grade students. *Journal of School Health, 83*, 45-52.
- National Council for Accreditation of Teacher Education. (2010). *Transforming teacher education through clinical practice: A national strategy to prepare effective teachers*. Retrieved from Washington, DC:
- Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, California: Sage Publications.
- Newton, S. (2010). *Preservice performance assessment and teacher early career effectiveness: Preliminary findings on the Performance Assessment for California Teachers*. Retrieved from Stanford, CA:
- Olweus, D. (1993). *Bullying at school: What we know and what we can do*. Oxford, UK: Blackwell.
- Pepler, D., Jiang, D., Craig, W., & Connolly, J. (2008). Development trajectories of bullying and associated factors. *Child Development, 79*(2), 325-338.
- Rigby, K., & Griffiths, C. (2010). *Applying the method of shared concern in Australian schools: An evaluative study*. Retrieved from <http://www.deewr.gov.au/Schooling/NationalSafeSchools/Documents/covertBullyReports/MethodOFSharedConcern.pdf>:

- Rose, C. A., & Espelage, D. L. (2012). Risk and protective factors associated with the bullying involvement of students with emotional and behavioral disorders. *Behavioral Disorders, 37*, 133-148.
- Rose, C. A., Monda-Amaya, L. E., & Espelage, D. L. (2011). Bullying perpetration and victimization in special education: A review of the literature. *Remedial and Special Education, 32*, 114-130.
- Rosenberg, M. B. (2003). *Nonviolent communication: A language of life*. Encinitas, CA: Puddle Dancer Press.
- Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *The Public Opinion Quarterly, 19*(3), 321-325.
- Segedy, J. R., Kinnebrew, J. S., & Biswas, G. (2013). The effect of contextualized conversational feedback in a complex open-ended learning environment. *Educational Technology Research and Development, 61*(1), 71-89.
- Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation, 9*(4).
- Stevenson, K., & Sander, P. (2002). Medical students are from Mars – business and psychology students are from Venus – University teachers are from Pluto? . *Medical Teacher, 24*(1), 27-31.
- U.S. Department of Education. (2013). *Indicators of School Crime and Safety: 2012*. National Center for Education Statistics.
- Voulgaridou, I., & Kokkinos, C. M. (2015). Relational aggression in adolescents: A review of theoretical and empirical research. *Aggression and Violent Behavior, 23*, 87-97. doi:10.1016/j.avb.2015.05.006
- Wang, J., Iannotti, R. J., & Nansel, T. R. (2009). School bullying among adolescents in the United States: Physical, verbal, relational, and cyber. *Journal of Adolescent Health, 45*(4), 368-375.
- Yoon, J. S. (2004). Predicting teacher interventions in bullying situations. *Education & Treatment of Children, 27*(1), 37-45.

APPENDIX B

Transcript from Shaun's Session 1 as Mrs. Fischer

Mrs. Fischer: Eli what is wrong?

Eli: They were being so mean to me!

Mrs. Fischer: Well what were they doing?

Eli: They called me a loser and said I was dumb!

Mrs. Fischer: Why would they do that?

Eli: I don't know. I was just playing by myself. I want them to leave me alone!!

Mrs. Fischer: Did you try asking them nicely to leave you alone?

Eli: I tried but they just made fun of me even more. I feel like I can't do anything about it.

Mrs. Fischer: They continued to "make fun" of you even after you asked them to stop?

Eli: I have no answer for that.

Mrs. Fischer: Did you say please when you asked them to stop?

Eli: I didn't do anything wrong! I was just sitting there.

Mrs. Fischer: Eli I am not saying you did anything wrong I am just trying to find out what happened

Eli: I have no answer for that.

Mrs. Fischer: Do you want to try talking to them again?

Eli: No. I just freeze up when they're around me.

Mrs. Fischer: Do you want me to come with you and talk to them?

Eli: I don't think that will work. They don't like me at all.

Mrs. Fischer: What makes you think that?

Eli: I have no answer for that.

Mrs. Fischer: Why makes you think they dont like you?

Eli: I have no answer for that.

Mrs. Fischer: Why wont it work?

Eli: I have no answer for that.

Transcript from Tanya's Session 3 as Mrs. Fischer

Mrs. Fischer: Hello, Eli. I hear you may have had some trouble on the playground today and maybe even before that. One of my most important jobs is to make sure our school is a safe place where everybody can learn and be happy. If that's not the case for everybody, I want to know how I can help make it better.

Eli: Hey Ms. Fischer. I don't know. They just don't seem to like me at all. It would help if they stopped making fun of me. I have no answer for that.

Mrs. Fischer: I want to hear more about what happened to you, so it's really important that you tell the truth. You should only tell me things that really happened to you. If I ask you a question and you don't know the answer just say "I don't know." Okay?

Eli: When they came up to me, Brandon called me "loser-face". Julio and DeShaun told Brandon he was funny and pointed at me while they laughed. Ok, I promise to tell the truth. Ok. Ok.

Mrs. Fischer: Okay, thank you. Now, do you know what an opinion is?

Eli: I have no answer for that. No, not really.

Mrs. Fischer: Okay. Stating an opinion is different from describing what happened. An opinion is what you think about what happened. If I asked you what I was like and you said, "You're cool," that would be your opinion because somebody else might say, "He's not cool." If you said, "You sat in your chair and asked questions," that would be telling me about what happened. First I would like you to try to tell me about things that happened, instead of your opinion, ok? Can you tell me everything that happened when you walked in my office?

Eli: Okay. Okay. They wouldn't stop laughing at me! I know. I really want them to stop. What should I do? They were calling me names like crybaby. I have no answer for that. Brandon was calling me names like "loser face". The other kids stood around and pointed at me while they laughed. I got really angry and yelled at them to go away but they didn't. All three of them started calling me a cry baby. I hate them!

Mrs. Fischer: Can you tell me exactly what happened that makes you say, "I hate them?"

Eli: They told me I look stupid and have no friends!

Mrs. Fischer: What else did they do?

Eli: They told me I would never fit in. They told me the cool kids all think I'm a loser!

Mrs. Fischer: Where did that happen?

Eli: It happens on the playground almost every day.

Mrs. Fischer: How long has this been going on for now?

Eli: I have no answer for that.

Mrs. Fischer: Where at recess does this happen?

Eli: It happened where I was playing by the swings. It usually happens there.

Mrs. Fischer: Okay. What happened right before Brandon called you "loser face?"

Eli: Okay. They were being so mean to me!

Mrs. Fischer: Can you tell me exactly what they were doing that makes you say that?

Eli: They were calling me names like crybaby.

Mrs. Fischer: How did you respond to that?

Eli: I have no answer for that.

Mrs. Fischer: Now I want to ask you a question that does ask for Mrs. your opinion. Do you have any idea why Brandon is behaving this way?

Eli: I have no answer for that. I don't know why!! I wasn't doing anything to them!

Mrs. Fischer: I understand. So you mentioned that Julio, DeShaun, and Brandon were calling you names like crybaby and loser face. Is there anything else that would help me understand?

Eli: Thanks for understanding. I have no answer for that. When I started crying they started laughing even harder!

Mrs. Fischer: I am really sorry that you are going through this. It is not your fault. Because it is so important to me that everybody feels safe and happy at our school, I will do everything I can to prevent this from happening again. Let's think about what you could do if you are with Brandon, Julio, and DeShaun at recess again.

Eli: I just want them to go away and leave me alone! How do I get them to stop? I just want to fight back. I could hit them. Well then what should I do about it?

Mrs. Fischer: You could try to use your words, instead of fighting, to explain how you feel.

Eli: Too much recursion in AIML

Mrs. Fischer: You could try to use your words to explain how you feel.

Eli: Too much recursion in AIML

Mrs. Fischer: I will I will be talking to Brandon, Julio, and DeShaun and some other people at the school individually so that I know everybody's description of the situation and so that I can help make things better. I will not tell them what we talked about today, but I will ask them some similar questions that I asked you. Is there anything else I should know before I speak with them?

Eli: Well then what should I do about it? Okay, thanks Ms. Fischer! Well then what should I do about it?

Mrs. Fischer: You could try explaining to them how you feel when they call you "loser face" or "crybaby"

Eli: Well then what should I do about it?

Mrs. Fischer: You could try to stay calm and explain how it makes you feel when they call you certain names

Eli: I could try that.

Mrs. Fischer: I understand yo I will always be at recess and I will make sure to keep an eye out for you.

Eli: Thanks for understanding.

Mrs. Fischer: I really appreciate you talking to me today. Let's talk again soon.

Eli: I have no answer for that. Okay, we can do that.