

PSYCHOMETRIC PROPERTIES OF A SOCIAL SKILLS ASSESSMENT USING VIRTUAL  
ENVIRONMENT

by:

THIEN-AN LE

B.S. Wofford College, Spartanburg, 2012

B.A. Wofford College, Spartanburg, 2012

A thesis submitted in partial fulfillment of the requirements  
for the degree of Master of Science  
in the Department of Psychology  
in the College of Sciences  
at the University of Central Florida  
Orlando, Florida

Spring Term  
2016

© 2016 Thien-An Le

## **ABSTRACT**

**Background:** Social Anxiety Disorder (SAD) is a marked and persistent fear of social and/or performance situations in which embarrassment or scrutiny from others may occur. In children, this marked and persistent fear must be present in peer settings and is not exclusive to interactions with adults (American Psychiatric Association [APA], 2013). Behaviorally, children with SAD may avoid eye contact and exhibit other behavioral symptoms such as stooped shoulders, nail biting, trembling voice, avoidance of social and performance situations, muffled voice, longer speech latency, inappropriate tone or low voice volume, and lack of spontaneous speech (Beidel & Turner, 2007; Ollendick, Benoit, & Grills-Taquechel, 2014; Spence, Donovan, & Brechman-Toussaint, 1999). Currently, there are several methods of assessing symptom severity of SAD, such as structured and semi-structured interviews supplemented by self- and parent-report forms, as well as behavioral assessment of social skills, such as RPTs. However, RPTs inherently present with feasibility concerns as there are several obstacles for its implementation. Thus, the current study will examine the psychometric properties of a VE based social skills assessment as it compares to the traditional RPT. **Methods:** Participants were 46 children, ages 7 to 14, who underwent two assessment conditions: RPT and VE BAT. Participants were assessed prior to the assessment conditions using the ADIS-C/P and completed several self- and parent-report forms. Participants reported self-ratings of anxiety and acceptability, while blinded observers rated social skills and overall social anxiety. **Results:** A paired-samples t-test revealed (a) no significant difference in acceptability between the two tasks ( $t(36) = .209, p > .05$ ); (b) the VE BAT elicited somewhat less anxiety and somewhat more

skilled social behavior than a comparable and traditional RPT; (c) the VE BAT demonstrated moderate concurrent validity with the SPAI-C ( $r = .422, p = .004$ ); (d) behaviors were rated as consistent across assessment tasks for speech latency ( $r = .367, p = .016$ ), overall effectiveness ( $r = .541, p = .000$ ), overall social anxiety ( $r = .638, p = .000$ ), and SAM ratings ( $r = .730, p = .000$ ) and; (e) VE BAT was more feasible to implement than the RPT in terms of personnel time ( $t(45) = 12.87, p = .00, d = 2.69$ ) and costs ( $t(45) = 12.88, p = .00, d = 1.83$ ). **Conclusion:** The current study addresses many of the discussed limitations of conducting RPTs and, overall, supports the utilization of VE BATS as a viable alternative to behaviorally assessing social skills in children. Overall, the current study demonstrates acceptability, validity, and feasibility of implementing such a novel method, where a formal RPT is not possible. Further implications for the current study include that VEs have potential in the armamentarium for social skills training with children with SAD.

## TABLE OF CONTENTS

LIST OF TABLES .....	vii
INTRODUCTION .....	1
METHODS: .....	8
Study Design .....	8
Participants .....	8
Clinician Administered Assessment Measures .....	9
Self-Report and Parental Report Measures .....	9
Assessment Task 1: RPT (see Appendix A) .....	11
Assessment Task 2: Virtual Environment BAT (see Appendix B) .....	12
RESULTS .....	15
Acceptability of VE BAT .....	15
Observer Ratings of Social Skill .....	15
Relationship of Social Skill and Anxiety in the Two Behavioral Tasks .....	16
Observer and Self-Report Ratings of Anxiety .....	17
Discriminative Validity of the VE BAT .....	18
Concurrent Validity of the VE BAT .....	18
Cost Analysis .....	20
DISCUSSION .....	22

CONCLUSION.....	28
APPENDIX A: BEHAVIORAL ASSESSMENT TASK (BAT) SCRIPT .....	29
APPENDIX B: EXAMPLE OF VIRTUAL ENVIRONMENT BEHAVIORAL ASSESSMENT TASK (VE BAT) SCRIPT .....	32
APPENDIX C: OBSERVER RATING FORM: SOCIAL ANXIETY .....	35
APPENDIX D: OBSERVER RATING FORM: OVERALL SOCIAL EFFECTIVENESS.....	37
APPENDIX E: OBSERVER RATING FORM: CONVERSATIONAL BEHAVIORS .....	39
APPENDIX F: ASSESSMENT ACCEPTABILITY FORM .....	41
APPENDIX G: SELF-RATINGS OF ANXIETY: SELF-ASSESSMENT MANIKIN .....	43
APPENDIX H: ASSESSMENT FEASIBILITY FORM.....	45
APPENDIX I: IRB APPROVAL LETTER.....	48
REFERENCES .....	51

## **LIST OF TABLES**

Table 1: Correlation of Corresponding Observer and Self-Report Ratings During RPT and VE BAT .....	16
Table 2: RPT and VE BAT mean scores and standard deviations for social skills and anxiety ratings .....	17
Table 3 Correlation of VE BAT Observer Ratings and Self-Report Measures .....	19
Table 4: Correlation of RPT Observer Ratings and Self-Report Measures .....	20
Table 5: RPT and VE BAT mean personnel time and costs (per assessment).....	21

## **INTRODUCTION**

Social Anxiety Disorder (SAD) is a marked and persistent fear of social and/or performance situations in which embarrassment or scrutiny from others may occur. In children, this marked and persistent fear must be present in peer settings and is not exclusive to interactions with adults (American Psychiatric Association [APA], 2013). Children with SAD typically fear making mistakes or acting in a way that will lead to feelings of embarrassment or humiliation in front of peers. In young children, this anxiety response may manifest as crying, screaming, or clinging to familiar persons or objects. Behaviorally, children with SAD may avoid eye contact and exhibit other behavioral symptoms such as stooped shoulders, nail biting, trembling voice, avoidance of social and performance situations, muffled voice, longer speech latency, inappropriate tone or low voice volume, and lack of spontaneous speech (Beidel & Turner, 2007; Beidel, Turner, & Morris, 1999; Ollendick, Benoit, & Grills-Taquechel, 2014; Spence, Donovan, & Brechman-Toussaint, 1999). With a prevalence rate of 5% in youth (Beidel & Turner, 2007), SAD is the most common anxiety disorder (Kessler et al., 1994) and the third most common psychiatric disorder in the United States (Kashdan & Herbert, 2001). While the onset of SAD typically occurs during mid to late adolescence, it can be present in children as young as 8 years of age (Beidel et al., 1999).

Speaking, reading, writing, or eating in front of others, talking on the telephone, and engaging in basic social interactions are situations that commonly elicit distress in children with SAD. In addition to distress and anxiety, deficits in these skills can result in impaired social and academic functioning. Social impairments include fewer friendships, increased feelings of



loneliness, and limited social relationships (Beidel et al., 1999). Children with SAD are also at increased risk for school refusal, depression, social isolation, substance use disorders, and fear of failure and criticism (Beidel & Morris, 1995; Strauss & Last, 1993).

SAD, especially in children, rarely remits without intervention (Davidson, 1993). Effective treatment requires thorough and accurate assessment of the child's clinical status. In order to target specific goals for treatment, reliable and valid means of assessment are necessary. Currently, there are several methods for the assessment of SAD, such as structured and semi-structured interviews supplemented by self- and parent-report forms. Structured and semi-structured interviews provide probes related to the diagnosis, allowing consistency and standardization, while requiring clinician judgments (Beidel & Turner, 2007; Edelbrock & Costello, 1988; Rogers, 1995; Segal & Hersen, 2010). Self-reports, on the other hand, are particularly helpful in quantifying symptoms and may be essential for assessing aspects of social behavior that are observable (Dunning, Johnson, Ehrlinger, & Kruger, 2003). However, the validity of information obtained from self-reports is less reliable since individuals may be poor judges of their social behavior (Dunning et al., 2003).

As noted by Beidel and Turner (1998), children who deny anxiety or difficulty making friends may actually be unable to demonstrate their friendship-making skills in a behavioral test, again illustrating the limits of self-report. Thus, behavioral assessments can have important roles in identifying presenting difficulties (Beidel & Turner, 1998; Rappee & Sweeney, 2005), but also present several challenges.

Thus, Behavioral Assessment Tests (BATs) represent a useful strategy to directly observe social anxiety and competence, especially in children who might verbally deny having social difficulties (Beidel & Turner, 1998; Bellack et al., 2006). Behavioral observation strategies have been used for over forty years, been evaluated numerous times, and evidence several advantages (Bellack et al., 2006). This method of assessment allows increased flexibility and opportunity to assess social behaviors that are not readily available by other assessment strategies, such as nonverbal and paralinguistic behaviors (Bellack et al., 2006). Although not widely used in assessing social skills in children, studies with adults indicate that BATs provide information relevant to treatment planning, outcome evaluation (Rapee & Sweeney, 2005), and allow observation of actual social behaviors rather than relying on self-report (Beidel & Turner, 1998).

One type of BAT is a Role-Play Task (RPT). During an RPT, children are instructed to imagine a series of brief, social scenarios described by a clinician and then respond to scripted prompts by a same-age peer as if these situations were actually happening (e.g., giving and receiving a compliment, and receiving help; (Beidel et al., 1999; Ollendick, 1981). Although the role-play scenes are based on real-world scenarios, the peer prompts are pre-scripted and standardized. While prepared scripts provide standardization across assessments, the rigid structure poses a challenge since the scripts do not always follow the patient's initial response, therefore resulting in awkward verbal interactions.

To address this limitation of RPTs, Beidel and colleagues developed an unstructured peer interaction task designed to assess social skills (e.g., Mesa, Beidel, & Bunnell, 2014; Scharfstein & Beidel, 2014). The task presented the child or adolescent the opportunity to play the Wii™

with a peer. During the interactive task, participants were not provided any instructions other than to play the Wii™ and to have fun, therefore allowing naturalistic social behaviors. Researchers then independently rated for the presence of specific social behaviors and assessed each participant's social skills (Scharfstein & Beidel, 2014). The study demonstrated high internal validity as social behaviors during the tasks differentiated children with SAD from those with no disorder. The peer interaction task also addresses questions of external validity by simulating scenarios that better represent reality while also providing clinicians the opportunity to observe a patient's social skills in a less structured, non-scripted method. One limitation of these less scripted tasks, however, is that they may not target specific social skills that clinicians need to assess.

Both forms of analogue assessments also present challenges to transportability and dissemination. Conducting RPTs in community mental health facilities may not be feasible due to the lack of resources. RPTs often require two-way mirrors, observation rooms, or video recording devices, as well as personnel costs for the clinician and the same-age peer (Beidel & Turner, 2007). Challenges of implementing unstructured tasks also require the need for adequate resources, such as a trained peer, as well as equipment needed to simulate a task such as playing interactive games and personnel costs for the clinician's time to recruit the peer, which also exists for non-structured tasks.

An alternative to the problems of recruiting and training live peers is the use of virtual environments (VEs). VEs allow the presentation of situational cues not easily reproduced either through imagination or in real life (*in vivo*) (Wong-Sarver, Beidel, & Spitalnick, 2013). With

respect to the behavioral tasks discussed above, VEs could allow users to interact with computer-based avatars in various socially related scenarios replacing the need for actual peers (Parsons & Mitchell, 2002).

Like RPTs, VEs provide clinicians the opportunity to assess different types of social behaviors such as greetings, asking questions, receiving or asking for help, receiving or giving compliments, and assertive communication (Wong-Sarver et al., 2013), but allow clinicians to manipulate environments rather than relying on the child's imagination. Additionally, the variety of virtual scenarios is particularly beneficial in assessing children with SAD because they rarely present with fears specific to one setting. Overall, as an added benefit, clinicians can control the behavior of virtual audience members in a way that allow maximum control over elements of the scenario (Klinger et al., 2005; Klinger et al., 2003; Roy et al., 2003) that are individualized to each patient. One VE developed specifically for children with SAD, *Pegasys-VR*, includes avatars that differ in age, sex, race, and roles (i.e., teacher, principal, classmate, or bully) in various social settings (i.e., hallways and classrooms) (Wong-Sarver et al., 2013). The ability to use a range of characters as interpersonal partners could allow clinicians to comprehensively and carefully assess each patient's unique social skills deficits by observing engagement in multiple social situations with multiple characters.

Irrespective of assessment modality, acceptability and feasibility are challenges for conducting reliable and valid behavioral assessments. Because they are contrived, some patients may not fully engage in RPTs and only partially demonstrate natural social behavior. This may be an even greater concern when avatars are used as interpersonal partners. However, recent data

indicate that both socially anxious children and adolescents accept the use of avatars and VEs as elements of assessment and treatment (Parrish, Oxhandler, Duron, Swank, & Bordnick, 2015; Wong-Sarver et al., 2014) While these studies provide initial support for the acceptability of VEs, additional studies are required to specifically investigate their acceptability as an RPT.

Not only should assessment approaches demonstrate strong acceptability, but they should also be feasible to conduct. Feasibility concerns include economic and personnel costs. Costs associated with conducting RPTs include the personnel time to construct and standardize appropriate role-play scenarios, as well as costs associated with the time required to acquire, train, schedule and compensate peers. Costs associated with VEs as an assessment method include the actual cost of the equipment and software, training time required for clinician proficiency, and lost time for equipment malfunction. Wong-Sarver et al. (2014) indicated that 8 hours of face-to-face training and 4 hours of follow-up telephone consultation with the program designer was needed for the clinician to become proficient at using a VE program. Thus, it can be expected that similar, or less, training and consultation time will be required to use a similar VE program as a RPT.

Another important psychometric property is validity. RPTs retain an important role for assessing social skills in children with SAD (Beidel & Turner, 2007). However, some studies concluded that during analogue role-plays, children with SAD display less interpersonal skill than normal control children (Beidel et al., 1999; Rao et al., 2007; Spence et al., 1999). Of these RPTs, an interactive Wii™ task also demonstrated validity in differentiating children with SAD from children with no disorder (Mesa et al., 2014; Scharfstein & Beidel, 2014). While existing

studies support discriminative validity of BATs, there are currently no studies investigating similar capacity for VEs.

In summary, behavioral assessments represent an important and underused method of understanding social behavior in children. However, BATs require substantial resources which may not be feasible in many clinical settings. Contrastingly, VEs have the potential to become a cost-effective, feasible, and psychometrically sound method of assessing social skills, thus allowing for direct observation in a controlled environment and with reduced “personnel” costs. The current study will examine the psychometric properties of a VE based social skills assessment. Specifically, this study will assess: (a) acceptability; (b) validity and; (c) feasibility of a VE social skills assessment in comparison to an RPT.

This study had the following hypotheses:

- 1.) VE BATs would have adequate acceptability but acceptability scores will be less than scores of the RPT.
- 2.) According to observer ratings, social skills demonstrated during VE BATs will be rated similarly to social skills demonstrated during the RPT.
- 3.) The VE BAT will have moderate to high discriminative validity with measures unrelated to social anxiety and moderate to high concurrent validity with other measures of SAD.
- 4.) VE BATs would be less costly to conduct than RPTs.

## **METHODS:**

### *Study Design*

This study used a within-subjects design in which each participant underwent two assessment conditions: RPT and VE BAT. Based on a previous study with adults (Owens & Beidel, 2014), there was a significant task order effect when an in vivo task was followed by a virtual reality task. Specifically, ratings of presence, engagement, and Subjective Units of Distress (SUDs) were significantly lower when the virtual reality task followed the in vivo task. Participants felt that the task was much “easier” having been through the “real thing.” To avoid this order effect, the task order of the current study was not counterbalanced (Owens & Beidel, 2014).

### *Participants*

This study included 46 children, ages 7 to 14 years. Participants were recruited via flyer advertisements posted in the community (e.g., grocery stores, coffee shops) and as an ongoing larger family study funded by the Department of Defense. Free treatment was also offered to participants that met diagnostic criteria for SAD.

*Telephone Screen* A Telephone Interview Form (TIF) developed specifically for the project screened parents who called the clinic. Children who met criteria for suicidal ideation, Bipolar Disorder, Schizophrenia, Major Depressive Disorder, Autism Spectrum Disorder, or Attention Deficit Hyperactive Disorder were excluded from the study.

### *Clinician Administered Assessment Measures*

Following consent, parent and children were administered the *Anxiety Disorders Interview Schedule for Parents and Children* (ADIS-C/P, Silverman & Albano, 1996). The ADIS-C/P is a semi-structured interview that has good to excellent inter-rater reliability with kappa coefficient ranging from .67 to .86 (Brown, DiNardo, Lehman & Campbell, 2001). Diagnoses were established based on information from both parent and child. To ensure diagnostic reliability, 20% of the videotaped interviews were reviewed by a second blinded clinician to assess diagnostic agreement. The primary and secondary assessors were doctoral students in clinical psychology and the current study obtained high inter-rater reliability ( $r = 1.00$ ).

### *Self-Report and Parental Report Measures*

Children completed the *Social Phobia and Anxiety Inventory for Children* (SPAI-C, Beidel, Turner, & Morris, 1995). The SPAI-C is a 26-item self-report instrument designed to assess social anxiety in children (Beidel, Turner, & Morris, 1995). Items on the SPAI-C list numerous social situations (reading aloud, performing in a play, eating in the school cafeteria) and assesses physical and cognitive characteristics of social anxiety disorder as well as avoidance behaviors. Each of the 26 items is rated on a 3-point scale (*never or hardly ever, sometimes, most of the time or always*). The SPAI-C has high internal consistency ( $\alpha = .95$ ), high 2-week test-retest reliability ( $r = .86$ ), and good external and discriminative validity (Beidel, Turner, & Fink, 1996).



The *Child Behavior Checklist* (CBCL, Achenbach, 1991) provided both broad-band and narrow-band scores of a child's emotional and behavioral functioning. The broad-band scores (i.e., Internalizing, Externalizing, and Total Problems) provided information regarding more general areas of behavior problems. Additionally, the CBCL is composed of subscales, one of which is the Social Competence scale, which provides an assessment of the child's involvement in social activities. The Social Competence scale has high internal consistency ( $\alpha = .68$ ) and high 1-week test re-test reliability ( $r = .93$ ) (Achenbach & Rescorla, 2001).

The *Multidimensional Anxiety Scale for Children 2<sup>nd</sup> Edition* (MASC-2, March, 2013) is a 39-item assessment distributed across four major factors, three of which can be parsed into two subfactors, including (1) physical symptoms (tense/restless and somatic/autonomic), (2) social anxiety (humiliation/rejection and public performance fears), (3) harm avoidance (perfectionism and anxious coping), and (4) separation anxiety (March et al., 1997). The social anxiety subscale has a high internal consistency ( $\alpha = .88$ ) and high 1 to 4-week test re-test reliability ( $r = .90$ ) for the self-report version and has a high internal consistency ( $\alpha = .88$ ) and high 1 to 4-week test re-test reliability ( $r = .91$ ) for the parent-report version (March, 2013). The MASC-2 is typically suitable for children ages 8 to 19 years, but for the purposes of the current study it was also used for participants 7 years of age with assessor assistance as needed.

*Acceptability of the BAT* Children completed a newly constructed questionnaire using a format administered in a previous study (Wong-Sarver et al., 2013) (see Appendix F). The modified evaluation assessed the quality of the BAT based on participants' impressions. Children rated the overall quality of the BAT and indicated how realistic the scenarios felt on a

rating scale of *poor, fair, good, very good* or *excellent*. More specifically, questions indicated how engaged and comfortable children felt during the BAT. To assess participant acceptability, the questionnaire included items such as “how real did the scenario feel?” and “how comfortable was it for you to share information with the other characters?” Previously, Wong-Sarver et al. (2013) successfully used this questionnaire with children in the same age range.

*Self-Ratings of Anxiety.* Children rated their own level of anxiety using a pictorially adapted version of the Self-Assessment Manikin (SAM; Bradeley & Lang, 1994) (see Appendix G). The SAM consists of five pictures depicting various levels of distress, corresponding to a 5-point Likert scale ranging from 0 (*little or no anxiety*) to 4 (*extreme anxiety*). Participants were asked to use the SAM scale immediately following each scene to describe how anxious or nervous they felt during the interaction.

*Assessment Feasibility* (see Appendix H) A questionnaire was completed by the clinician logging total time required for assessment completion. Variables documented include: (a) time required recruiting peers; (b) time spent training peers; (c) amount of compensation for peer participation; (d) number of cancelled assessments due to peers failing to attend; (e) required training time for the clinician; (f) time required addressing technological issues and; (g) time required conducting each assessment.

#### *Assessment Task 1: RPT (see Appendix A)*

Each child participated in a structured role-play (Beidel, et al., 1999). The RPT consisted of five brief scenarios requiring interaction with a same-age peer including greetings, asking questions, giving compliments, receiving compliments, and assertiveness. Each participant was

instructed to imagine a specific scenario, described by the researcher, and to respond as if the situation were actually happening. The researcher described the situation and a same-age peer initiated the interaction by reciting a scripted prompt. The participant responded to the same-age peer. The same-age peer then provided a second scripted prompt, which the participant responded to accordingly. This procedure was repeated until all five role-play scenarios were presented to the participant.

When same-aged peers failed to arrive, “younger” looking undergraduate research assistants played the role of the confederate. However, undergraduate research assistants were only used when the participant was in the adolescent age range.

A practice scene was presented first to ensure that the participant understood the procedure. Peers were instructed to respond in a friendly but neutral manner and not to lead the conversation. If the child did not respond approximately ten seconds after the prompt, the peer provided the second prompt. Peers did not meet diagnostic criteria for a DSM-V disorder and were female. Social skills were assessed based on the responses and behaviors of the child during the conversation with the peers (see Blinded Observer Ratings).

#### *Assessment Task 2: Virtual Environment BAT (see Appendix B)*

The behavioral assessment described above was replicated using the *Pegasys-VR*. To assess social skills using VE, each child sat at a desktop computer where the VE BAT was administered. While the participant sat at the computer, the researcher initiated the VE BAT in a separate room, introducing the VE elementary school. During this assessment, the researcher used a “Wizard of Oz” interface to guide the interactions and the responses of the avatars. The

VE BAT consisted of five scenarios with topics identical to the RPT: greetings, asking questions, giving compliments, receiving compliments, and assertiveness. Each participant was instructed to imagine that they were a student at the elementary school and to respond to the avatar as if the scenario was actually happening. Once the researcher introduced the situation, the avatar spoke. To demonstrate the flexibility of VE, the avatar responded based on the participant's initial response, prompting the participant to respond again. This procedure was repeated until all five role-play scenarios were presented to the participant.

A practice scene was presented first to ensure that the participant understood the procedure. The researcher allowed approximately ten seconds for the participant to respond to the prompt before initiating conversation again. Social skills were assessed based on the responses and behaviors of the child during the conversation with the virtual avatars.

*Blinded Observer's Ratings:* Each RPT and VE BAT was video-taped and coded by trained observers. Each observer was trained by the first author. Two blinded observers were used to establish inter-rater reliability (i.e., Pearson's  $r$ , at least  $r = .80$  on all ratings) with the researcher. Observers were first trained on how to code each social behavior (see Appendix C) and social effectiveness (see Appendix D) using five videotaped social interactions. Feedback was provided to both observers, as necessary. Observers were given ten videotaped assessments to code. To establish inter-rater reliability between raters, a set of ten videotapes composed of sample behaviors similar to that in this study was coded.

Behaviors that observers rated include: (a) voice volume; (b) speech latency; (c) number of words spoken; (d) social effectiveness and; (e) overall social anxiety (see Appendix E).

Measures such as voice volume, speech latency, and words spoken were assessed objectively by measuring decibels, reaction time, and number of words spoken, respectively.

## **RESULTS**

### *Acceptability of VE BAT*

To determine the acceptability of the VE BAT, a paired-samples t-test indicated that there was no significant difference between acceptability ratings for the RPT and the VE BAT tasks ( $t(36) = .209, p > .05$ ). The mean rating for the VE BAT task was 2.81, whereas the mean rating for the RPT task was 2.84, indicating that both formats were rated as *good* to *very good* and were equally acceptable.

### *Observer Ratings of Social Skill*

Prior to analyses, all variables were examined for normal distribution. All variables were normally distributed with the exception of the average number of words spoken during assessment tasks.

The distribution for the average number of words spoken during the RPT and VE scores were examined for skewness and kurtosis values. Based on standardized values for skewness (2.43) and kurtosis (8.59), the number of words spoken during the RPT was positively skewed and peaked. One outlier was identified and removed using a threshold value of  $Z = 2.00$ . Also, skewness (1.49) and kurtosis (2.86) data for the number of words spoken during the VE task was positively skewed and peaked. Again, using a threshold value of  $Z = 2.00$ , two outliers were removed and the distribution was normalized. All other variables reflected a normal distribution and did not require removal of outliers.

*Relationship of Social Skill and Anxiety in the Two Behavioral Tasks*

To determine consistency across the two tasks, observer ratings of a) voice volume, b) speech latency, c) number of words spoken, d) social effectiveness, e) overall social anxiety and f) SAM ratings reported during the VE BAT were correlated with the corresponding social skill and social anxiety ratings observed during the RPT. Results revealed a significant positive correlation for the variables of speech latency ( $r = .367, p = .016$ ), overall effectiveness ( $r = .541, p = .000$ ), overall social anxiety ( $r = .638, p = .000$ ), and SAM ratings ( $r = .730, p = .000$ ). There was no significant correlation for voice volume or number of words spoken.

Table 1: Correlation of Corresponding Observer and Self-Report Ratings During RPT and VE BAT

Variable	Voice Volume	Speech Latency	Number of Words	Overall Effectiveness	Overall Social Anxiety	SAM Ratings
<i>r</i>	.037	.367*	-.016	.541**	.638**	.730**
<i>p</i>	.825	.016	.920	.000	.000	.000
<i>N</i>	39	43	43	44	44	45

\*  $p < .05$

\*\*  $p < .001$

A series of paired-samples t-tests was used to determine if observer ratings of a) voice volume, b) speech latency, c) number of words spoken, and d) effectiveness were significantly different across these two assessment formats (RPT vs. VE BAT). Means and standard deviations are depicted in Table 1. Results revealed a statistically significant difference for overall effectiveness  $t(43) = 2.70, p = .01, d = 0.41$ , where participants were rated as more effective when interacting with the virtual avatars than when interacting with a confederate.

Table 2: RPT and VE BAT mean scores and standard deviations for social skills and anxiety ratings

Variables	RPT		VE BAT		<i>P</i>
	Mean	SD	Mean	SD	
Voice Volume	3.87	.27	3.92	.39	.54
Speech Latency	1.76	.43	1.94	.67	.138
Number of Words	8.48	14.17	9.94	19.82	.70
Overall Effectiveness	1.85	.81	1.56	.58	.01**
Overall Social Anxiety	2.20	.74	1.68	.52	.00***
SAM Ratings	.81	.58	.85	.74	.66

\*  $p < .05$

\*\*  $p < .01$

\*\*\*  $p < .001$

There were no significant task differences for number of words spoken or speech latency.

#### *Observer and Self-Report Ratings of Anxiety*

A paired-samples t-test compared observer ratings of overall social anxiety during the RPT and VE BAT. Results revealed a statistically significant difference  $t(43) = -6.07, p = .00, d = .82$ , with participants rated as more anxious when interacting with a confederate than when interacting with the avatars. In contrast, a paired-samples t-test comparing children's self-reported anxiety ratings during the RPT and the VE BAT did not reveal a statistically significant difference between the two tasks.



### *Discriminative Validity of the VE BAT*

Pearson correlations revealed no significant correlations between SAM ratings during the VE BAT and either the MASC OCD subscale ( $r = .26, p > .05$ ) or the MASC Panic Disorder subscale ( $r = .29, p > .05$ ). These results indicate that there is no relationship among self-ratings of anxiety during the VE BAT social interaction task and self-reported symptoms of OCD and Panic Disorder.

### *Concurrent Validity of the VE BAT*

To determine the concurrent validity of the VE BAT as a tool to assess social skill and social anxiety, observer ratings of a) voice volume, b) speech latency, c) number of words spoken, d) social effectiveness, e) overall social anxiety and f) SAM ratings reported during the VE BAT were correlated with the SPAI-C, the MASC SAD subscale, and the Social Competency subscale of the CBCL.

There was a significant positive correlation between SAM VE BAT ratings and the SPAI-C,  $r = .422, p = .004$ , indicating that children who reported higher anxiety on the SPAI-C also reported experiencing more anxiety during the VE BAT.

There were no statistically significant correlations between any observer rating of social skills or social anxiety during the VE BAT and the SPAI-C, MASC SAD subscale, or the Social Competency subscale of the CBCL.

Table 3 Correlation of VE BAT Observer Ratings and Self-Report Measures

Variables		SPAI-C	MASC SAD	CBCL- Social Competence
Voice Volume	<i>r</i>	-.076	-.017	-.247
	<i>p</i>	.627	.913	.106
	<i>N</i>	43	44	44
Speech Latency	<i>r</i>	-.187	-.115	.005
	<i>p</i>	.230	.456	.975
	<i>N</i>	43	44	44
Number of Words	<i>r</i>	-.011	-.047	-.057
	<i>p</i>	.946	.763	.717
	<i>N</i>	42	43	43
Effectiveness	<i>r</i>	.141	.153	.099
	<i>p</i>	.368	.322	.523
	<i>N</i>	43	44	44
Overall Social Anxiety	<i>r</i>	.073	.130	-.081
	<i>p</i>	.641	.399	.601
	<i>N</i>	43	44	44
SAM VE BAT	<i>r</i>	.422	.279	.085
	<i>p</i>	.004**	.064	.577
	<i>N</i>	44	45	45

\*  $p < .05$   
\*\*  $p < .005$

Subsequently, a Pearson correlation was conducted to examine the relationship among observer ratings of a) voice volume, b) speech latency, c) number of words spoken, d) social effectiveness, e) overall social anxiety and f) SAM ratings and the SPAI-C, the MASC SAD subscale, and the Social Competency subscale of the CBCL.

There was a significant positive correlation between SAM RPT ratings and the SPAI-C,  $r = .315$ ,  $p = .037$ , indicating that children who reported higher anxiety on the SPAI-C also reported experiencing more anxiety during the RPT. There were no statistically significant

correlations between any observer ratings of social skills or social anxiety during the RPT and the SPAI-C, MASC SAD subscale, or the Social Competency subscale of the CBCL.

Table 4: Correlation of RPT Observer Ratings and Self-Report Measures

Variables		SPAI-C	MASC SAD	CBCL- Social Competence
Voice Volume	<i>r</i>	-.061	-.011	-.130
	<i>p</i>	.718	.505	.431
	<i>N</i>	38	39	39
Speech Latency	<i>r</i>	-.205	-.122	.029
	<i>p</i>	.194	.436	.856
	<i>N</i>	42	43	43
Number of Words	<i>r</i>	.115	.158	.071
	<i>p</i>	.464	.306	.648
	<i>N</i>	43	44	44
Effectiveness	<i>r</i>	.005	.087	-.002
	<i>p</i>	.976	.576	.988
	<i>N</i>	43	44	44
Overall Social Anxiety	<i>r</i>	.037	.033	-.091
	<i>p</i>	.816	.831	.557
	<i>N</i>	43	44	44
SAM RPT	<i>r</i>	.315*	.203	-.062
	<i>p</i>	.037	.181	.688
	<i>N</i>	44	45	45

\*  $p < .05$

### Cost Analysis

A paired-samples t-test compared personnel time required to conduct the RPT and VE BAT assessment tasks. Means and standard deviations are depicted in Table 3 and results revealed a statistically significant difference between the time necessary to prepare for and conduct the RPT than the VE BAT  $t(45) = 12.87, p = .00, d = 2.69$ .

To compare personnel costs, a paired-samples t-test was conducted. Clinician cost in the community was calculated at \$150.00 per hour. There was a statistically significant difference

between RPT and VE BAT personnel cost  $t(45) = 12.88, p = .00, d = 1.83$ . These results suggest that RPT assessment tasks cost significantly more than VE BAT assessments in terms of billable clinician time.

In addition, confederates required for the RPT assessment task was compensated for their time with a \$10.00 gift card. Subsequently, the total cost for confederates equated to \$460.00. Contrastingly, the VE program cost and all necessary equipment were quoted as \$1,295.00, which includes: software for PC or Apple laptop, iPhone 4S, PC laptop with Wi-Fi and necessary specifications, manuals and other associated accessories. However, no confederates or additional compensation was required when conducting the VE BAT. Overall, when accounting for personnel cost and program costs, each assessment costs \$66.82 per RPT assessment, compared to \$42.92 per VE BAT assessment.

Table 5: RPT and VE BAT mean personnel time and costs (per assessment)

	RPT		VE-BAT		<i>p</i>
	Mean	SD	Mean	SD	
Personnel Time ( <i>in minutes</i> )	18.94	7.34	4.77	1.19	.00***
Billable Personnel Cost ( <i>in dollars</i> )	56.82	22.03	14.77	1.96	.00***
Other Associated Costs	10.00	0.00	28.15	0.00	--
Overall Costs ( <i>Personnel cost plus associated costs</i> )	66.82	0.00	42.92	0.00	--

\*  $p < .05$

\*\*  $p < .01$

\*\*\*  $p < .001$

## **DISCUSSION**

This study sought to examine psychometric properties of a VE based social skills assessment in terms of (a) acceptability, (b) validity, and (c) feasibility in comparison to an RPT. The results indicated that VE BAT was as equally accepted by participants as the RPT, elicited behaviors that were consistent with behaviors elicited during an RPT, demonstrated moderate concurrent validity with the SPAI-C, but not correlated with unrelated measures of SAD, elicited somewhat less anxiety and somewhat more skill than a comparable RPT, and was more feasible to implement than the RPT in terms of program and personnel costs.

The first aim of the study determined whether VE BATs would have adequate acceptability by child participants. Consistent with other investigations (Parrish et al., 2015; Wong-Sarver et al., 2014), the results support the feasibility and utility of VE BATs for the assessment of social skills. The average acceptability rating for both VE and RPT tasks fell just below the cut-off for *very good*, suggesting that the children found both formats to be acceptable. Specifically, with high acceptability, clinicians and researchers can have confidence that children responded to the prompts in the same fashion and with the same amount of effort across both conditions. Of course, the exact relationship of VE to behavior in naturalistic settings remains to be established, but these results support use of VE BAT to feasibly assess social skills without the challenges of more traditional RPTs.

The second aim of the study compared children's social performance across these two formats and assessed both social skill and social anxiety. When correlating corresponding social skills and social anxiety ratings between both conditions, results revealed significant positive

correlations for speech latency, effectiveness, overall social anxiety, and SAM ratings. These results indicate behavioral consistency across the two tasks. Whatever their level of social skill or social anxiety, they demonstrated behavioral consistency across the formats.

However, when comparing level of performance across the two tasks, there was a significant task difference in observer ratings of overall social anxiety. Although anxiety was rated as minimal to mild in both conditions, children were rated as more anxious when interacting with a live confederate than with avatars. The overall low ratings of anxiety in both tasks is likely due to the fact that only 6% of this unselected sample ( $n = 3$ ) met diagnostic criteria for social anxiety disorder. Similarly, there were significant differences between overall effectiveness on the VE BAT and the RPT, with better performance during the VE BAT. Thus, although children's behaviors were consistent across conditions, they exhibited significantly more anxiety and significantly less effectiveness when interacting with a live peer, when compared to an avatar.

These results are consistent with other recent investigations comparing virtual reality (VR) to live audiences in the assessment of social anxiety disorder. According to Owens and Beidel (2015), differences in social performance between these two conditions (VR versus a live audience) condition may be accounted for by the difference in the presented stimulus. Although participants demonstrated some physiological arousal when giving a speech to a virtual audience, the distress elicited was significantly less than an actual live audience. Overall, participants judged the VR to be less immersive than real life confederates, and indicated that they were less anxious when speaking to the avatars because "the avatars were not real and could not judge my

behavior”. Given that the core of social anxiety disorder is a fear that others will evaluate one’s behavior in a negative light, the VR cannot sufficiently elicit this fear when compared to a live audience (Owens & Beidel, 2015). The same contingencies were no doubt at play in this investigation, and combined with the observer ratings of more anxiety during the RPT, these results demonstrate that although children respond very similarly across conditions, the RPT task remains a somewhat more demanding assessment of social performance, even for this unselected sample. Whether these results remain consistent for treatment seeking samples, remains open for further investigation.

The third aim of the study was to examine the validity of the VE as an assessment of social anxiety and social skill. Consistent with our hypotheses, there was no correlation between self-report of social anxiety during the VE and self-report measures of obsessive-compulsive disorder and panic disorder, suggesting some degree of discriminative validity.

With respect to concurrent validity, there was a significant positive correlation between self-reported anxiety on the VE task and scores on the SPAI-C as there was between the RPT and the SPAI-C. Children who reported higher anxiety on the SPAI-C also reported more anxiety during these conditions. However, there were no statistically significant correlations between observer ratings of social skills or self-report ratings of anxiety during the VE BAT or RPT and (a) self-report measures of social anxiety or (b) parental measures of social competence. These data stand in contrast to previous investigations (Beidel, Turner, Hamlin, & Morris., 2001). One explanation for the different outcome of this investigation is the different sample characteristics. This sample reflected a general population sample where only a few children had significant

social fears and met criteria for social anxiety disorder, whereas previous samples have used much larger samples of children who met diagnostic criteria of social anxiety disorder. Thus, while representing a general child population, the small number of children with social anxiety disorder may have restricted the range of the assessed behaviors, thus leading to the lack of statistically significant correlations for these variables. Future studies that included a significant number of children with SAD might provide further validation of this assessment strategy.

The fourth aim compared the costs of the two assessments. The results indicated that VEs require significantly less personnel time and therefore, significantly less personnel costs, when compared to RPT. When clinician cost in the community was calculated at \$150.00 per hour, VE BAT costs significantly less money to conduct than an RPT. When all costs are calculated, a clinician could conduct four VE BATs in the same amount of time required to conduct one RPT. This difference supports the conclusion that conducting VE BATs are far more efficient than RPTs and thus, consistent with Wong-Sarver et al. (2014), provides the clinician with a non-intrusive cost-efficient method by which to directly observe a child's social skills and social anxiety.

Thus, the current study demonstrates acceptability and feasibility of the VE BAT as an alternative strategy for the assessment of social skills. Yet, despite its feasibility, validity and cost-effectiveness, further studies are required to assess its validity. Performance on the VE BAT was significantly better than the RPT on measures of social anxiety and social skills, suggesting that the VE provided a less rigorous assessment. Thus, although it may not provide the most optimal assessment of social performance, VE BATs could be a plausible alternative for



assessing social skills where it is too burdensome or challenging for clinicians to find peer-confederates or when community mental facilities are not equipped for RPTs.

Virtual reality and virtual environments are playing an increasing role in the treatment of anxiety and trauma-related disorders and the VE described in the study has had preliminary success in the treatment of childhood social anxiety disorder (Wong-Sarver et al., 2013). Extant studies that do not include social skills training (Compton et al., 2010; Ginsburg et al., 2011; Hudson et al., 2015; RUPP, 2002) consistently report attenuated treatment outcomes for children with SAD using standard pharmacological or CBT (exposure, cognitive restructuring); there is a need to consider adding additional interventions such as social skills training in order to achieve optimal outcome (Hudson et al., 2015). Indeed, the data indicate that interventions that include social skills training, such as Cognitive-Behavioral Group Therapy for Adolescents (CBGT-A) and Social Effectiveness Therapy for Children (SET-C) provide efficacious treatment outcomes for children with SAD and further emphasizes the need to incorporate social skills training within the treatment program. In addition to its assessment utility, the results of the current study provide some support for use of VE/VR as a tool for social skills training, as multiple behavioral rehearsal opportunities can be provided to solidify newly learned skills in a safe, but variable environment. Particularly when skills training is done on an individual basis (rather than in a group setting) finding other individuals who can practice in order to promote skill generalization can be a challenge. Thus, VE may serve as an intermediate step from practice with a clinician in the office to practice (without the clinician) in the community. Consistent with the belief of Parsons and Mitchell (2002), however, it is important to note that the use of VEs does not aim to

minimize social interactions, which would be counter-productive for the generalization of such skills.

This study was not without its limitations. One limitation was the inability to consistently recruit same-age peers as the confederate for RPTs. The difficulty in recruiting same-age peers illustrates the overall difficulty community mental health facilities would face in attempting to conduct RPTs. For this study, we used “younger” looking undergraduate research assistants to play the role of confederate when the participant was in the adolescent age range. Although this limitation poses as a potential confound to the current study, a paired samples t-test revealed no significant differences in observer ratings of social skills when the confederate was a same-age peer or an undergraduate research assistant.

Additionally, the sample used in the current study was not drawn from a treatment seeking population, but rather represented an unselected sample of children who chose to participate in this research study. It may be expected to see even greater significant differences in observer ratings of social skills, as well as self-reported ratings of anxiety in a clinical population between the RPT and VE BAT. This could occur because people with SAD fear negative evaluation by others and thus, the presence of a real life confederate (rather than an avatar) would tap directly into this core fear, thus exacerbating any anxiety experienced throughout the interaction tasks (Owens & Beidel, 2015). Additionally, it is possible that a replication of this study using a treatment seeking sample may yield significant correlations between social skills and social anxiety and self- and parent-report measures of similar behaviors.

## **CONCLUSION**

Individuals with SAD may be poor judges of their social behavior (Dunning et al., 2003), or in the case of children, may deny social anxiety or difficulty making friends even when their social deficits are apparent to others (Beidel & Turner, 1998). Thus, not only do BATs have important roles in identifying social skill deficits, they can provide information relevant to treatment planning and outcome evaluation (Beidel & Turner, 1998; Rapee & Sweeney, 2005). However, RPTs, a type of BAT, present several challenges in their implementation. Thus, it is important to examine the psychometric properties, feasibility, and cost effectiveness of alternative strategies such as a VE.

In summary, the current study is the first to examine the acceptability, validity, and feasibility of a virtual environment to directly observe and assess social skills. Overall, these findings indicate that VE BATs are not an exact analogue to RPTs, but may be an acceptable alternative to traditional RPTs in behaviorally assessing social skills among children in cases where conducting a true RPT is not feasible. Long term effects of utilizing VE BATs have the potential to reduce personnel costs involved with conducting these assessments, thus allowing greater dissemination and in an increasing range of clinical settings. Future studies using clinical populations and incorporating test-retest reliability procedures would be valuable to further demonstrate the utility of such technology in designing optimal assessment and treatment strategies for this chronic and disabling condition.

**APPENDIX A: BEHAVIORAL ASSESSMENT TASK (BAT) SCRIPT**

## **APPENDIX A: Behavioral Assessment Task (BAT) Experimenter Script**

We are going to do role-plays today, and after we are done, I am going to ask you to look at this sheet. We call this little guy “SAM” and what I want you to do is to point to the picture of SAM that best describes how you felt when you were doing the role-plays. So if you felt very nervous, you would point to the picture of the very nervous SAM (point to #5), and if you did not feel nervous at all, you would point to the not very nervous SAM picture (point to #1). So, SAM #1 is like eating an ice cream cone where you are not nervous at all and SAM #5 is like being chased by a bear where you are really, really nervous.

“Today we are going to do some little skits, called role-plays. I am going to describe situations and (actor’s name) is going to say some things that someone your age may or may not say to you in real life. What I want you to do is respond just how you would in real life, and if you wouldn’t say anything in real life, that’s OK too. We are going to do a practice scene first and if you have any questions, you can ask me at that time.”

### Practice Scene:

Imagine that you are at the movies and you are buying some popcorn. You pay the cashier and receive your popcorn. There is a boy/girl standing behind you and he/she says:

Actor: How’s the popcorn?

Actor: I would really like to have some, can I have a taste?

### Scene 1:

You are riding your bike in front of your house. A boy/girl is standing next to his/her bike and it looks like he/she had a crash and is looking down at a flat tire. You approach him/her. He/she looks at you, and with a sad voice, he/she says:

Actor: How am I going to get this darn bike home?

Actor: I guess I ought to call my dad.

### Scene 2:

In gym class, you are learning how to play basketball and how to shoot free throws. You are having trouble making some shots from the free throw line. Another boy/girl who is a good basketball player says:

Actor: Would you like for me to help you with your free throws?

Actor: Well, it was hard for me to learn at first. Would you like for me to give you some pointers?

Scene 3:

A boy/girl who sits next to you in math class is having some trouble with his/her math test. He/she's been working hard to get his/her grade up. The class gets back the most recent test with grades on them. He/she gets a big smile on her/her face and says:

Actor: I finally got an A!

Actor: I've been studying so hard.

Scene 4:

You've been working hard to memorize a poem to recite in English Literature class. You finish reciting the poem in front of the class and return to your seat. The boy/girl sitting next to you says:

Actor: You did a great job.

Actor: You remembered every word and you looked so calm and cool.

Scene 5:

You are reading a comic book during recess. Pretty soon another kid takes your comic and says.

Actor: I'm going to read it myself.

Actor: Go find another one.

**COMPLETE the practice scene and Scenes 1-5**

After completing Scenes 1-5, pick up the SAM and ask the participant to point to the picture that best describes how he/she felt during the role-play.

**APPENDIX B: EXAMPLE OF VIRTUAL ENVIRONMENT BEHAVIORAL  
ASSESSMENT TASK (VE BAT) SCRIPT**

## **APPENDIX B: Example of Virtual Environment Assessment Behavioral Assessment Task**

### **(VE BAT) Script**

We are going to do skits today, and after we are done, I am going to ask you to look at this sheet. We call this little guy “SAM” and what I want you to do is to point to the picture of SAM that best describes how you felt when you were doing the role-plays. So if you felt very nervous, you would point to the picture of the very nervous SAM (point to #5), and if you did not feel nervous at all, you would point to the not very nervous SAM picture (point to #1). So, SAM #1 is like eating an ice cream cone where you are not nervous at all and SAM #5 is like being chased by a bear where you are really, really nervous

“Today we are going to do some little skits with the computer. I am going to describe situations and the characters on the computer will say some things that someone your age may or may not say to you in real life. What I want you to do is respond just how you would in real life, and if you wouldn’t say anything in real life, that’s OK too. We are going to do a practice scene first and if you have any questions, you can ask me at that time.”

#### Practice Scene:

Imagine that you are at school and walking through the hall. You see a girl that you have class with. She stops you and says:

Avatar (Cool girl): What’s up?

Avatar (Cool girl): Want to go to the gym?

#### Scene 1:

Imagine that it is the first day of school and you are in a class with other kids you have never met before. You take a seat in your first class of the day and the girl sitting next to you says:

Avatar (Smart girl): Oh, Hi Hi Hi

Avatar (Smart girl): Hey my friend

#### Scene 2:

You get to school and take your seat in your Language Arts class. When you sit down, the girl next to you turns and says:

Avatar (Smart girl): Did you read anything good lately?



Avatar (Smart girl): When are we gonna study together?

Scene 3:

A girl who sits next to you in math class is having some trouble with her math test. she's been working hard to get her grade up. The class gets back the most recent test with grades on them. she gets a big smile on his/her face and says:

Avatar (Smart Girl): Oh good, I tried really hard

Avatar (Smart Girl): This is one of my favorite classes

Scene 4:

You've been working hard to memorize a poem to recite in English Literature class. You finish reciting the poem in front of the class and return to your seat. The girl sitting next to you says:

Avatar (Smart Girl): Oh my gosh, you rocked that assignment!

Avatar (Smart Girl): uh huh, everyone was clapping for you!

Scene 5:

You have been working really hard on a class project for the past week. The day before the assignment is due, another kid sees you in the hallway and says:

Avatar (Cool girl): I had practice last night and then went to bed. Let me copy your homework.

Avatar (Cool girl): You are so smart-- can I copy your math homework?

**COMPLETE the practice scene and Scenes 1-5**

After completing Scenes 1-5, pick up the SAM and ask the participant to point to the picture that best describes how he/she felt during the role-play.

**APPENDIX C: OBSERVER RATING FORM: SOCIAL ANXIETY**

## APPENDIX C: Observer Rating Form: Social Anxiety

**Severely anxious:** Uncomfortable, gross motor signs of anxiety exhibited consistently (hand wringing, or turning, leg shaking, fidgety). Also could be manifested as extreme inhibition (“frozen with fear”).

**Moderately anxious:** Clear signs of discomfort, awkward, some gross motor movements as above, but less extreme and/or less consistent than above.

**Mildly anxious:** Occasional signs of anxiety, which consist primarily of facial apprehension (furrowed brow, eyes wide open), or awkward body movement (slight hand wringing, awkward seating position).

**Not at all anxious:** No overt signs of anxiety, smiles at conversational partner, appears interested and/or enjoys the interaction.

Rate each scene separately

Scene 1	4	3	2	1
Scene 2	4	3	2	1
Scene 3	4	3	2	1
Scene 4	4	3	2	1
Scene 5	4	3	2	1

Score (Average of all scenes) \_\_\_\_\_

**APPENDIX D: OBSERVER RATING FORM: OVERALL SOCIAL  
EFFECTIVENESS**

## APPENDIX D: Observer Rating Form: Overall Social Effectiveness

**Not effective at all:** Looks awkward, no response or one word response, does not ask questions, mumbling, barely audible speech.

**Minimally effective:** Clearly awkward, answers questions but mainly gives two or three word responses, and no further participation in conversation.

**Moderately effective:** Only mild awkwardness, able to respond to questions fully, some degree of fluidity, and moderate effort to keep conversation going, voice volume moderate.

**Effective:** No awkwardness, carries part of the conversation, may self-disclose, appears to enjoy to the interaction, voice strong and clear.

Rate each scene separately

Scene 1	4	3	2	1
Scene 2	4	3	2	1
Scene 3	4	3	2	1
Scene 4	4	3	2	1
Scene 5	4	3	2	1

Score (Average of all scenes) \_\_\_\_\_

**APPENDIX E: OBSERVER RATING FORM: CONVERSATIONAL  
BEHAVIORS**

**APPENDIX E: Observer Rating Form: Conversational Behaviors**

**Voice Volume**

	<u>SCENE 1</u>	<u>SCENE 2</u>	<u>SCENE 3</u>	<u>SCENE 4</u>	<u>SCENE 5</u>
	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

- 1 = Inappropriate voice volume; too loud or inaudible
- 2 = Voice volume somewhat too loud or barely inaudible
- 3 = Slightly too loud or moderately inaudible
- 4 = Appropriate volume

**Average of All Scenes** \_\_\_\_\_

**Latency to First Utterance:** Record the number of seconds between when the child actor finishes each line and when the target child begins to speak (.1-10 seconds).

	<u>SCENE 1</u>	<u>SCENE 2</u>	<u>SCENE 3</u>	<u>SCENE 4</u>	<u>SCENE 5</u>
Response time from Line 1:	_____	_____	_____	_____	_____
Response time from Line 2:	_____	_____	_____	_____	_____

**Number of Words Spoken**

	<u>SCENE 1</u>	<u>SCENE 2</u>	<u>SCENE 3</u>	<u>SCENE 4</u>	<u>SCENE 5</u>
Not include utterances (e.g., eh, uh, um, like)	_____	_____	_____	_____	_____

**Average of All Scenes** \_\_\_\_\_

**Appropriateness of Response**

	<u>SCENE 1</u>	<u>SCENE 2</u>	<u>SCENE 3</u>	<u>SCENE 4</u>	<u>SCENE 5</u>
	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4	1 2 3 4

(Degree to which the emotion displayed is appropriate to the social scenario; facial expressions; overt behaviors)

- 1 = No response to either prompt; response is not at all appropriate
- 2 = Minimally appropriate response
- 3 = Moderately appropriate

**Average of All Scenes** \_\_\_\_\_

4 = Appropriate response; both responses are appropriate  
(e.g., says “thank you” when complimented; asserts oneself with a bully)

**APPENDIX F: ASSESSMENT ACCEPTABILITY FORM**



## APPENDIX F: Assessment Acceptability Form

### Assessment Acceptability Form

<b>HOW WOULD YOU RATE:</b>	<b>Excellent</b>	<b>Very Good</b>	<b>Good</b>	<b>Fair</b>	<b>Poor</b>
1. Quality of the conversation	4	3	2	1	0
2. Quality of what you saw	4	3	2	1	0
3. How real did it feel when you were talking to the other characters	4	3	2	1	0
4. How real did the scenario feel	4	3	2	1	0
5. How comfortable was it for you to share information with the other character	4	3	2	1	0
6. How comfortable were you when you were talking to the other character	4	3	2	1	0
7. How natural did your interactions feel	4	3	2	1	0
8. How likely do you think these scenarios would occur naturally	4	3	2	1	0
9. How involved were you during the experience	4	3	2	1	0
10. How engaged were you during the experience	4	3	2	1	0

Thinking about today's activities and experience, please provide honest opinions for each item.

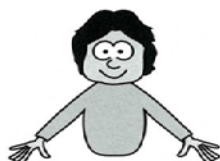
**APPENDIX G: SELF-RATINGS OF ANXIETY: SELF-ASSESSMENT**  
**MANIKIN**

**APPENDIX G: Self-Ratings of Anxiety: Self-Assessment Manikin**

# SAM

---

Point at the picture that shows how scared you were when this happened



---

**0**

**1**

**2**

**3**

**4**

---

Copyright © 2003 Multi-Health Systems Inc. All rights reserved.

**APPENDIX H: ASSESSMENT FEASIBILITY FORM**

**APPENDIX H: Assessment Feasibility Form**

**Assessment Feasibility Form  
Virtual Environment Log**

Total VE Clinician Training Time (minutes): \_\_\_\_\_

Date	Total Assessment Time (minutes)	Time Spent Addressing Technological Issues (minutes)	Time Spent in Contact with IT or VBI (minutes)	Problem	Solution	Notes

**APPENDIX H: Assessment Feasibility Form**

**Assessment Feasibility Form  
Role-Play Task Log**

<b>Date</b>	<b>Total Assessment Time (minutes)</b>	<b>Total Time Recruiting Peers (minutes)</b>	<b>Time Spent Training Peers (minutes)</b>	<b>Amount of Compensation for Peer Participation</b>	<b>Number of Cancelled Assessments Due to Peers Failing to Attend</b>	<b>Notes</b>

**APPENDIX I: IRB APPROVAL LETTER**

## APPENDIX I: IRB Approval Letter



University of Central Florida Institutional Review Board  
Office of Research & Commercialization  
12201 Research Parkway, Suite 501  
Orlando, Florida 32826-3246  
Telephone: 407-823-2901 or 407-882-2276  
[www.research.ucf.edu/compliance/irb.html](http://www.research.ucf.edu/compliance/irb.html)

### Approval of Human Research

**From:** UCF Institutional Review Board #1  
FWA00000351, IRB00001138

**To:** Thien-An Le

**Date:** November 04, 2014

Dear Researcher:

On 11/4/2014, the IRB approved the following human participant research until 11/03/2015 inclusive:

Type of Review: UCF Initial Review Submission Form  
Project Title: Psychometric Properties of a Social Skills Assessment Using a Virtual Environment  
Investigator: Thien-An Le  
IRB Number: SBE-14-10634  
Funding Agency:  
Grant Title:  
Research ID: N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30 days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at <https://iris.research.ucf.edu>.

If continuing review approval is not granted before the expiration date of 11/03/2015, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the [Investigator Manual](#).

On behalf of Sophia Dziegielewska, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:



*Joanne Muratori*

Signature applied by Joanne Muratori on 11/04/2014 09:15:31 AM EST

IRB Coordinator

## REFERENCES

- Achenbach, T. (1991). Manual for the Child Behavior Checklist/4 - 18 and 1991 Profile. Burlington, VT: University of Vermont Department of Psychiatry.
- Achenbach, T. M., & Rescorla, L. (2001). ASEBA school-age forms & profiles. Burlington: Aseba.
- Albano, A. M., Marten, P. A., & Holt, C.S. (1991). *Cognitive-Behavioral Group Treatment of Adolescent Social Phobia: Therapist's Manual*. Unpublished manuscript, State University of New York at Albany.
- Alfano, C. A., Beidel, D. C., & Turner, S. M. (2006). Cognitive correlates of social phobia among children and adolescents. *Journal of Abnormal Child Psychology*, 34(2), 182-194.
- American Psychiatric Association (APA) (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edition, Arlington, VA, APA.
- Beidel, D.C., Alfano, C.A., Kofler, M.J., Rao, P.A., Scharfstein, L., & Wong-Sarver, N. (2014). The impact of social skills training for social anxiety disorder: A randomized controlled trial. *Journal of Anxiety Disorders*, 28, 908-918.
- Beidel, D. C., Fink, C. M., & Turner, S. M. (1996). Stability of anxious symptomatology in children. *Journal of Abnormal Child Psychology*, 24(3), 257-269.
- Beidel, D. C., Rao, P. A., Scharfstein, L., Wong, N., & Alfano, C. A. (2010). Social skills and social phobia: an investigation of DSM-IV subtypes. *Behaviour research and therapy*, 48(10), 992-1001.

- Beidel, D.C., & Turner, S.M. (2007). *Shy Children, Phobic Adults: The Nature and Treatment of Social Anxiety Disorder*. Washington, DC: American Psychological Association.
- Beidel, D. C., Turner, S. M., Hamlin, K., & Morris, T. L. (2001). The Social Phobia and Anxiety Inventory for Children (SPAI-C): external and discriminative validity. *Behavior Therapy*, 31(1), 75-87.
- Beidel, D. C., Turner, S. M., Jacob, R. G., & Cooley, M. R. (1989). Assessment of social phobia: Reliability of an impromptu speech task. *Journal of Anxiety Disorders*, 3(3), 149-158.
- Beidel, D.C., Turner, S.M., & Morris, T.L. (1995). A new instrument to assess childhood social anxiety and phobia: The Social Phobia and Anxiety Inventory for Children. *Psychological Assessment*, 1, 73-79.
- Beidel, D.C., Turner, S.M., & Morris, T.L. (1999). Psychopathology of childhood social phobia. *Journal American Academy Child Adolescence Psychiatry*, 38, 643-650.
- Bellack, A. S. (1979). A critical appraisal of strategies for assessing social skill. *Behavioral Assessment*, 1, 157-176.
- Bellack, A. S., Brown, C. H., & Thomas-Lohrman, S. (2006). Psychometric characteristics of role-play assessments of social skill in schizophrenia. *Behavior Therapy*, 37, 339-352.
- Bellack, A. S., Hersen, M., & Turner, S.M. (1978). Role-play tests for assessing social skills: Are they valid? *Behavior Therapy*, 9, 448-461.
- Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry*, 52, 976 - 986.

- Bouton, M. E., & King, D. A. (1986). Effect of context on performance to conditioned stimuli with mixed histories of reinforcement and nonreinforcement. *Journal of Experimental Psychology. Animal Behavior Processes*, 12, 4 – 15.
- Brown, T. A., DiNardo, P., Barlow, D. H., & DiNardo, P. A. (1994). Anxiety Disorders Interview Schedule for DSM-IV.: Client interview schedule. Oxford University Press.
- Brown, T.A., DiNardo, P.A., Lehman, C.L., & Campbell, L.A., (2001). Reliability of *DSM-IV* anxiety and mood disorders: Implications for the classification of emotional disorders. *Journal of Abnormal Psychology*, 110, 85–599.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. Social phobia: Diagnosis, assessment, and treatment, 41(68), 00022-3.
- Compton, S. N., Walkup, J. T., Albano, A. M., Piacentini, J. C., Birmaher, B., Sherrill, J. T., ... & March, J. S. (2010). Child/adolescent anxiety multimodal study (CAMS): Rationale, design, and methods. *Child and adolescent psychiatry and mental health*, 4(1), 1.
- Davidson, J. R. T. (1993, March) Childhood histories of adult social phobics. Paper presented at the Anxiety Disorders Association of America Annual Convention, Charleston, SC.
- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J.(2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12, 83-87.
- Edelbrock, C., & Costello, A. J. (1988). Structured psychiatric interviews for children. In M. Rutter, A. H. Tuma & I. S. Lann (Eds.), *Assessment and diagnosis in child Psychopathology* (pp. 87–112). *New York: Guilford*.
- Eisler, R. M., Hersen, M., Miller, P. M., & Blanchard, E. G. (1975) Situational determinants of assertive behaviors. *Journal of Consulting and Clinical Psychology*, 43, 330-340.

- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.
- Ginsburg, G. S., Kendall, P. C., Sakolsky, D., Compton, S. N., Piacentini, J., Albano, A. M., ... & March, J. (2011). Remission after acute treatment in children and adolescents with anxiety disorders: findings from the CAMS. *Journal of consulting and clinical psychology*, *79*(6), 806.
- Hayward, C., Varady, S., Albano, A. M., Thienemann, M., Henderson, L., & Schatzberg, A. F. (2000). Cognitive-behavioral group therapy for social phobia in female adolescents: Results of a pilot study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(6), 721-726.
- Herbert, J. D., Gaudiano, B. A., Rheingold, A. A., Myers, V. H., Dalrymple, K., & Nolan, E. M. (2005). Social skills training augments the effectiveness of cognitive behavioral group therapy for social anxiety disorder. *Behavior Therapy*, *36*(2), 125-138.
- Hersen, M. (Ed.). (2011). *Clinician's handbook of child behavioral assessment*. Academic Press.
- Hersen, M., Bellack, A. S., & Turner, S. M. (1978). Assessment of assertiveness in female psychiatric patients: motor and autonomic measures. *Journal of Behavior Therapy and Experimental Psychiatry*, *9*, 11-16.
- Hudson, J. L., Keers, R., Roberts, S., Coleman, J. R., Breen, G., Arendt, K., ... & Eley, T. C. (2015). Clinical Predictors of Response to Cognitive-Behavioral Therapy in Pediatric Anxiety Disorders: The Genes for Treatment (GxT) Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*: 454–63.

- Kashdan, T. B., & Herbert, J. D. (2001). Social anxiety disorder in childhood and adolescence: Current status and future directions. *Clinical Child and Family Psychology Review*, 4(1), 37-61.
- Kazdin, A. E. (1981). Behavioral Observation. In M. H. Hersen & A.S. Bellack (Ed.), *Behavioral Assessment: a practical handbook* (pp. 110-111). Elmsford, NY: Pergamon Press Inc.
- Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H. U., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States: Results from the National Comorbidity Survey. *Archives of General Psychiatry*, 51, 8–19.
- Klinger, E., Bouchard, S. S., Le´geron, P. P., Roy, S. S., Lauer, F. F., Chemin, I. I., & Nugues, P.P. (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *CyberPsychology & Behavior*, 8, 76–88.
- Klinger, E., Chemin, I., Le´geron, P., Roy, S., Lauer, F., & Nugues, P. (2003). Designing virtual worlds to treat social phobia. In B. Wiederhold, G. Riva & M. D. Wiederhold (Eds.), *Cybertherapy* (pp. 113–121). San Diego, CA: Interactive Media Institute.
- March, J.S. (2013). The Multidimensional Anxiety Scale for Children Second Edition (MASC 2): Technical Manual. NY: Multi-Health Systems Inc.
- March, J. S. (2013). *Multidimensional Anxiety Scale for Children* (2nd ed.). Toronto, Ontario, Canada: Multi-Health Systems.
- Mesa, F., Beidel, D. C., & Bunnell, B. E. (2014). An Examination of Psychopathology and Daily Impairment in Adolescents with Social Anxiety Disorder. *PloS one*, 9(4), e93668.

- Muris, P., Meesters, C., & Fijen, P. (2003). The self-perception profile for children: Further evidence for its factor structure, reliability, and validity. *Personality and Individual Differences, 35*(8), 1791-1802.
- Ollendick, T. H. (1981). Assessment of social interaction skills in school children. *Behavioral Counseling Quarterly, 1*, 227-243.
- Ollendick, T. H., Benoit, K. E., & Grills-Tauechel, A. E. (2014). Social Anxiety Disorder in Children and Adolescents. *The Wiley Blackwell Handbook of Social Anxiety Disorder*, 179-200.
- Orvaschel, H. (2006). Structured and semistructured interviews. In M. Hersen (Ed.), *Clinician's handbook of child behavioral assessment* (pp. 159–179). *San Diego, CA: Elsevier Academic Press*.
- Owens, M. E., & Beidel, D. C. (2015). Can Virtual Reality Effectively Elicit Distress Associated with Social Anxiety Disorder?. *Journal of Psychopathology and Behavioral Assessment, 37*(2), 296-305.
- Parsons, S. & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research, 46*, 430-443.
- Rapee, R. M., & Sweeney, L. (2005). Social phobia in children and adolescents: Nature and assessment. *The essential handbook of social anxiety for clinicians*, 123-131.
- Rao, P., Beidel, D., Turner, S., Ammerman, R., Crosby, L., & Sallee, F. (2007). Social anxiety disorder in childhood and adolescence: Descriptive psychopathology. *Behaviour Research and Therapy, 45*, 1181–1191.

- Rodebaugh, T. L., Holaway, R. M., & Heimberg, R. G. (2004). The treatment of social anxiety disorder. *Clinical Psychology Review, 24*(7), 883-908.
- Rogers, R. (1995). Diagnostic and structured interviewing: A handbook for psychologists. *Odessa, FL: Psychological Assessment Resources.*
- Roy, S. S., Klinger, E. E., Le'geron, P. P., Lauer, F. F., Chemin, I. I., & Nugues, P. P. (2003). Definition of a VR-based protocol to treat social phobia. *CyberPsychology & Behavior, 6*, 411–420.
- RUPP Anxiety Study Group. (2002) Treatment of pediatric anxiety disorders: An open-label extension of the Research Unites on Pediatric Psychopharmacology Anxiety Study. *Journal of Child and Adolescent Psychopharmacology, 12*, 175-188.
- Scharfstein, L. A., & Beidel, D. C. (2014). Social skills and social acceptance in children with anxiety disorders. *Journal of Clinical Child & Adolescent Psychology*, (ahead-of-print), 1-13.
- Scharfstein, L. A., Beidel, D. C., Sims, V. K., & Finnell, L. R. (2011). Social skills deficits and vocal characteristics of children with social phobia or Asperger's disorder: a comparative study. *Journal of abnormal child psychology, 39*(6), 865-875.
- Segal, D. L., & Hersen, M. (2010). Diagnostic Interviewing (pp. 31–48). *New York, NY: Springer Publishing Co.*
- Shaffer, D., Fisher, P. W., & Lucas, C. P. (1999). Respondent-based interviews. In D. Shaffer, C.P. Lucas & J. E. Richters (Eds.), *Diagnostic assessment in child and adolescent Psychopathology* (pp. 3–33). *New York: Guilford.*



- Silverman, W.K., & Albano, A.M. (1996) Anxiety Disorders Interview Schedule for DMS-IV: Child Version. New York: Oxford University Press, Inc.
- Spence, S., Donovan, C., & Brechman-Toussaint, M. (1999). *Social skills, social outcomes, and cognitive features of childhood social phobia. Journal of Abnormal Psychology, 108*, 211–221.
- Spence, S.H., Donovan, C., & Brechman-Toussaint, M. (2000). The treatment of childhood social phobia: The effectiveness of a social skills training-based, cognitive-behavioral intervention, with and without parental involvement. *Journal of Child Psychology & Psychiatry, 41*, 713-726.
- Strauss, C. C., & Last, C. G. (1993). Social and simple phobias in children. *Journal of Anxiety Disorders, 7*(2), 141-152.
- Wong-Sarver, N., Beidel, D.C., & Spitalnik, J. (2013). The Feasibility and Acceptability of Virtual Environments In the Treatment of Childhood Social Anxiety Disorder. *Journal of Clinical Child & Adolescent Psychology, 43*(1), 63–73, 2014.
- Zaider, T. I., Heimberg, R. G., Fresco, D. M., Schneier, F. R., & Liebowitz, M. R. (2003). Evaluation of the clinical global impression scale among individuals with social anxiety disorder. *Psychological medicine, 33*(04), 611-622.