Published in final edited form as:

Health Educ Res. 2005 February ; 20(1): 61–70. doi:10.1093/her/cyg103.

Psychometric properties of virtual reality vignette performance measures: a novel approach for assessing adolescents' social competency skills

Mallie J. Paschall[1,4], **Diana H. Fishbein**[2], **Robert C. Hubal**[3], **Diana Eldreth**[2]

[1]Prevention Research Center, Pacific Institute for Research and Evaluation, Berkeley, CA 94704

[2]Transdisciplinary Behavioral Science Program, RTI International, Baltimore, MD 21224

[3]Technology Assisted Learning Division, RTI International, Research Triangle Park, NC 27709, USA

Abstract

This study examined the psychometric properties of performance measures for three novel, interactive virtual reality vignette exercises developed to assess social competency skills of at-risk adolescents. Performance data were collected from 117 African-American male 15–17 year olds. Data for 18 performance measures were obtained, based on adolescents' interaction with a provocative virtual teenage character. Twelve of the 18 performance measures loaded on two factors corresponding to emotional control and interpersonal communication skills, providing support for their factorial validity. The internal reliability coefficients for the two multi-item measures were 0.88 and 0.91, respectively. Additional analyses with established measures of three psychosocial factors (beliefs supporting aggression, aggressive conflict-resolution style and hostility) and behavioral criteria (e.g., self-reported behavioral misconduct and drug use) provided limited support for the construct and criterion-related validity of the performance measures. Study findings suggest that the virtual reality vignette exercises may represent a promising approach for assessing adolescents' social competency skills.

Introduction

Interventions designed to improve adolescents' social competency skills have proliferated in recent decades to prevent or reduce behaviors such as interpersonal violence and substance abuse (Botvin, 2000; Gottfredson and Wilson, 2003; Winters *et al.*, 2003). However, assessment of these skills has largely been limited to self-ratings or external ratings by teachers and parents and archival records, which are limited in a number of ways. Self-ratings may be inflated or subject to recall and social desirability biases, while external ratings by teachers or parents may be based on limited observation of adolescents' behavior and/or stereotyping bias (Frankfort-Nachmias and Nachmias, 2000). Archival records of disciplinary actions may fail to capture behaviors that reflect poor social competency skills. Although role-play exercises may be used to gauge adolescents' social competency skills,

[4]Correspondence to: M. J. Paschall; paschall@pire.org.

they may be difficult to implement with large numbers of adolescents. These limitations point to the need for novel measures of social competency skills that can be administered relatively easily, simulate real interactions with other people and be used to evaluate prevention programs.

Responsive virtual human (RVH) technology is a rapidly advancing method for assessing technical and social competency skills (Kizakevich *et al.*, 1998; Olsen, 2001; Hubal *et al.*, 2003a,b). RVH technology allows individuals to engage in seemingly real verbal discourse with virtual characters (i.e. computerized human-like figures). Thus, virtual reality (VR) simulation affords a more realistic social encounter than less interactive paper- or computer-based assessments and circumvents ethical or legal issues surrounding actual provocative role-play encounters.

The present study is one of the first to utilize RVH technology to develop a social competency skills assessment tool for at-risk adolescents. The software developed for this study is based on the types of social competency skills that have been associated with violent behavior and drug use, and that are targeted by many adolescent violence and drug abuse prevention programs. These skills include emotional control, information seeking, expressing one's own preferences, negotiation and being non-provocative.

Previous research has demonstrated that questionnaire vignettes representing hypothetical social situations can discriminate between adolescents at high and low risk for violent behavior [e.g. (Slaby and Guerra, 1988; Paschall and Flewelling, 1997; Paschall and Hubbard, 1998)]. Vignettes used in these studies asked adolescents what they would do in potentially provocative social situations (e.g. 'What would you do if somebody pushes you out of line at the water fountain?'), with possible response options corresponding to social competencies such as seeking information and stating one's preferences. Those studies found that violent adolescents were more likely to choose verbally and/or physically aggressive responses than non-violent adolescents, and thus provide empirical support for the potential of VR simulations as a means of assessing adolescents' social competency skills.

The present study extends prior research by using RVH technology and hypothetical social situations to evaluate adolescents' social competency skills. Due to the relatively high cost and complexity of developing a totally immersive VR system, a non-immersive or 'desktop' form of VR was developed for this study (Robertson *et al.*, 1993). A virtual teenage character of the same gender, ethnicity and age of study participants was created to simulate provocative interpersonal social situations in a school setting. Because this innovative technology is untested, this study examines adolescents' level of engagement and subjective impressions of the VR vignette exercises, and psychometric properties of VR vignette performance measures.

Psychometric properties of VR vignette performance measures examined in this study include factorial, construct and criterion-related validity, and internal reliability. Factorial validity is assessed through factor analysis to determine whether measures that appear to be conceptually similar or different actually load on the same or different factors

(DeVellis, 1991). The internal reliability of multi-item factors derived from the VR vignette performance measures indicates the extent to which multiple items loading on a single factor represent measures of a single underlying construct (Nunally, 1978; DeVellis, 1991). In order for the VR vignette exercises to be a useful measurement tool, the internal reliability of multi-item measures of social competency skills derived through factor analysis must be adequate, using a Cronbach's α coefficient of 0.80 or higher as the criterion (DeVellis, 1991).

The construct validity of a novel measure is supported if the measure is associated with established measures of theoretically related constructs measured at the same time (Nunally, 1978; De-Vellis, 1991). This study examines the associations between multi-item VR vignette performance measures derived through exploratory factor analysis and multi-item measures of beliefs supporting aggression, aggressive conflict-resolution style and hostility. We hypothesized that the multi-item VR vignette performance measures would be inversely associated with the three established psychosocial measures.

The criterion-related validity of a novel measure is supported if the measure is associated with other objective (criterion) measures of the same or similar constructs (Nunally, 1978; DeVellis, 1991). We hypothesized that the multi-item VR vignette performance measures would be inversely associated with an overall rating of adolescents' behavioral misconduct and their self-reported drug use.

Methods

Subjects

Subjects were 117 African-American male 15–17 year olds (mean = 16.07, SD = 0.25) in Grade 10 who were participating in a longitudinal epidemiological prevention study. A purposive sample of adolescents was selected for this study based on self-report and parent ratings of behavioral misconduct from 1999 to 2003. Approximately half of the subjects met diagnostic criteria (DSM-IIIR) for conduct disorder during that period and half did not have symptoms of conduct disorder or any other reported behavioral problem. This sampling strategy was used to ensure adequate variation in behavioral problems and related psychosocial factors, and for assessment of psychometric properties of the VR vignette performance measures. The number of subjects enrolled into the study was based on results of statistical power analyses indicating the minimum sample size needed to detect subgroup differences in levels of psychosocial factors and related social competency skills.

VR vignettes

The VR vignettes were developed with three-dimensional graphical software, and speech recognition and language processing technology to simulate real interpersonal verbal interactions with appropriate body language (Hubal *et al.*, 2003b). Based in part on the content of role-plays in the Positive Adolescent Choices Training violence prevention program developed for urban African-American youth (Yung and Hammond, 1998), a script was developed for the virtual teenage character (shown in Fig. 1) to entice the study participant to engage in several types of risky behavior. The script included provocative

introductory statements and multiple response options for different types of verbal feedback from a study participant, including questions intended to elicit additional information from the character and statements expressing the participant's position or preference. An algorithm was developed by which the virtual teenager would initially entice the risky behavior, but gradually back off if the participant demonstrated appropriate avoidance and/or de-escalation behavior (i.e. information seeking, negotiation). The language the virtual teenager used was modeled on that used by youth in the study sample, although for the present study we used computer-generated speech. The virtual teenager remained in one position because the entire interaction was based primarily on conversation, so the gestures were not complex, relying mostly on beat gestures and idle motions (Cassell *et al.*, 2000). However, when the virtual character became agitated or aggressive, the gestures became more representational (e.g. pointing, placing hands on hips, backing away), depending on the content of the conversation.

Three virtual confrontational situations were developed for this study. In one situation, the virtual teenage character asked the study participant to keep a gym bag with something inside in his school locker, but did not provide any information about what was in the gym bag. Participants who asked about the contents within the bag were told by the virtual character not to worry about what's inside the bag and to keep the bag. However, if the participant inquired further, he discovered that a pair of sneakers was in the gym bag. Although never explicitly stated by the virtual teenager, the sneakers within the gym bag were stolen and participants who agreed to keep the bag were taking a risk by keeping them. Participants who declined to keep the bag were enticed with a $50 bribe from the virtual character. In a second situation, the virtual teenager invited the participant to come to a drinking party with girls at his home. A request for further information or a rejection of the offer elicited persuasive, high-pressure responses from the virtual character such as: 'These girls are going to be hot', 'We can't do this any other time' and 'You don't have a girlfriend'. In a third situation, the virtual teenager accused the participant of bumping into him in the hallway and tried to provoke a fight. If the participant backed off, rejecting offers to fight and defusing the situation, the virtual character attempted to exacerbate the situation by making the following remarks: 'I'll kick your butt', 'Don't cry' or 'You must be afraid of your own shadow.' Participants responding to the provocation in an inflammatory manner were aggravated further by the virtual character. These situations represented actual experiences of urban minority adolescents that are used for role-plays in prevention programs such as Positive Adolescent Choices Training (Yung and Hammond, 1998) and Life Skills Training (Botvin *et al.*, 1995; Botvin, 2000).

Before the VR vignette exercises began, study participants were told that they would be interacting with a teenage character on the computer and that they should behave just as they would in a real-life situation in school or on the street. Participants were led to believe that the character could hear them and would respond directly to their utterances; they were shown the speakers on the computer and told to speak clearly into them so the character could understand them. They were also told that there were no right or wrong ways to respond to the teenage character, only those in which they were inclined. At the beginning of the first exercise, the virtual teenage character introduced himself as 'Greg', and asked the study participant his name and the time, providing participants with some practice and

familiarization prior to the actual exercises. Participants were also encouraged to ask the facilitator what Greg said if for some reason they were unable to understand him. The VR vignette exercises were administered in a random order to minimize any influence that ordering of exercises might have on subjects' performance.

Measures

Measures of engagement and social competency skills were based on ratings of adolescents' verbal and non-verbal interaction with the virtual teenage character along seven dimensions: general engagement, verbalizations, emotional control, information seeking, expressing own preferences, compromise/negotiation and being non-provocative. A five-point scale for each dimension ranged from 'very low' to 'moderate' to 'very high.' Because of substantive, technical and logistical issues that have delayed the development of an automated rating system, two trained researchers observed each adolescent's performance on the three VR vignette exercises and recorded their ratings on separate forms, yielding a total of 18 performance measures (participants' level of engagement was considered separately from the other six observational measures). The observers sat behind the adolescents so they could not be seen during the exercises. The correspondence between the two raters' scoring for the six social competency dimensions was good (Cohen's $\kappa = 0.92–0.97$). More detail about level of engagement and the six VR vignette performance ratings is provided in the Appendix.

A semi-formal debriefing was subsequently conducted with study participants to assess their subjective impressions of the VR vignette exercises. Participants were asked the following questions: How did you like the computer exercises? Were they fun? Were they boring? Were they interesting? Were they realistic? Did they keep your attention? Were they silly/stupid? Were they helpful? In addition, the facilitator questioned participants to determine whether they felt their responses to the virtual character reflected behaviors that would accurately characterize them in similar real-life situations. Participants' responses were recorded on a response form.

Study participants then completed a written questionnaire that included the multi-item psychosocial measures described below.

A 10-item scale for beliefs supporting aggression was adapted from a study by Slaby and Guerra (Slaby and Guerra, 1988). Participants were asked how strongly they agreed or disagreed with such statements as: 'It makes you feel big and tough when you push someone around' and 'A guy who doesn't fight back when other kids push him around will lose respect'. A mean response score was computed for each participant with a higher score indicating more beliefs supporting aggressive behavior (mean = 2.21, SD = 0.43; Cronbach's $\alpha = 0.72$).

Four vignettes originally developed by Slaby and Wilson-Brewer (Slaby and Wilson-Brewer, 1992) were used to measure participants' conflict-resolution style. Youth were presented with a hypothetical confrontational situation and then asked what they would probably do in that situation. For each vignette, participants were presented with verbally aggressive,

information seeking, passive, verbally assertive and physically aggressive response options. A four-point response scale was created as follows: 1 = neither a verbally nor a physically aggressive response was chosen; 2 = a verbally aggressive response was chosen, but not a physically aggressive response; 3 = a physically aggressive response was chosen, but not a verbally aggressive response; and 4 = both verbally and physically aggressive responses were chosen. A mean response score was computed for each participant with a higher score indicating a more aggressive conflict-resolution style (mean = 2.07, SD = 0.88; Cronbach's $\alpha = 0.76$).

A five-item hostility scale was based on the Symptom Checklist-90 (Derogatis *et al.*, 1976). Participants were asked such questions as, 'How often do you have temper outbursts that are hard to control?' with four possible response options ranging from 'never' to 'most of the time'. An average score was computed for each participant with a higher score indicating more hostility (mean = 1.92, SD = 0.57; Cronbach's $\alpha = 0.76$).

Behavioral criterion measures were based on data collected for the ongoing epidemiological prevention study. As noted above, adolescents were classified as meeting DSM-IIIR criteria for conduct disorder based on self-reports and parent ratings of behavioral misconduct (e.g. physical aggression) from 1999 to 2003. Dichotomous measures of lifetime alcohol, cigarette and marijuana use were based on adolescent self-reports in a written questionnaire administered in Grade 9.

Data analysis

Descriptive statistics were used to examine subjects' level of engagement in the VR vignette exercises, and qualitative data were reviewed to determine whether subjects thought the VR vignette exercises reflected real-life situations, whether they were taking the exercises seriously, and if they would have behaved differently in real-life situations.

Principal components analysis was conducted to assess the nature and number of latent constructs underlying the VR vignette performance measures, and the extent to which measures with similar face validity would load on the same factor. Orthogonal axis rotation was invoked to assess the independence of measures. A minimum factor loading of approximately 0.70 was used as the criterion for retaining items. The internal reliability of multi-item VR vignette performance measures was assessed by computing a Cronbach's α coefficient for each measure.

The construct and criterion-related validity of VR vignette performance measures was assessed by examining their associations with measures of psychosocial constructs and behavioral problems, respectively. Pearson product-moment correlation coefficients and *t*-tests were used for these analyses. Because of the directional hypotheses, one-tailed significance tests were used.

Results

Level of engagement and qualitative feedback

Descriptive statistics for each level of engagement measure are provided in Table I. Approximately half of the subjects were rated as being highly or very highly engaged in all three VR vignette exercises, while at least 80% were at least moderately engaged in the exercises. Mean values suggest that subjects were most engaged in the exercise in which the virtual teenage character attempted to provoke a fight.

Many participants stated that the gym bag vignette was most likely to happen to them in school. One participant mentioned that he had gotten in trouble for keeping a bag in his locker for a friend because the bag contained illicit drugs. Participants who refused to keep the bag often stated that they were afraid that drugs were in the bag. A frequent remark from participants regarding the drinking party vignette was that they would go to the party and not drink. To get out of going to the party without offending the virtual teenager, some participants would say they were busy and could not go. Of the participants who were provoked by the hallway confrontation vignette, many said that they do not tolerate anyone pushing them or talking about their mother in a negative way. Participants who tried to defuse the situation would often tell the virtual character they were sorry and it would never happen again. Overall, many participants were enthralled by the fact that the virtual character was able to respond to them, which may explain why engagement levels were so high. Most participants stated that they would have behaved the same in similar real-life situations as they did with the virtual character.

Exploratory factor analysis

Principal components analysis results are provided in Table II. Five factors with eigenvalues greater than 1.0 were initially observed, but there was a substantial drop in the percent of total variance explained after the second factor, as indicated by the eigenvalues and by an elbow in the scree plot. Thus, a two-factor solution was obtained with 12 of the 18 VR vignette performance measures. Measures of emotional control and being non-provocative from all three VR vignette exercises loaded strongly on one factor, which appeared to represent adolescents' emotional control or composure in a confrontational situation. Measures of expressing own preferences and verbalizations from all three exercises loaded on a second factor, which appeared to represent adolescents' interpersonal communication skills in a confrontational situation. The information-seeking and negotiation measures did not load strongly on any factor, and were dropped from analyses.

The internal reliability of the multi-item emotional control and interpersonal communication skills factor measures was very good (Cronbach's αs = 0.88 and 0.91, respectively). Factor scores for the two multi-item measures (mean = 0, SD = 1) were used in subsequent analyses to accommodate differential loading of items on each factor.

Construct and criterion-related validity

Results of bivariate analyses conducted to assess the construct and criterion-related validity of the emotional control and interpersonal communication skills factor measures are

provided in Table III. The emotional control factor was inversely associated with all three psychosocial variables (beliefs supporting aggression, aggressive conflict-resolution style and hostility), and the interpersonal communication skills factor was inversely associated with aggressive conflict-resolution style and hostility. The emotional control factor was not associated with any of the behavioral criterion measures at the 0.05 level, but was inversely related to lifetime alcohol use at the 0.10 level. The interpersonal communication skills factor was inversely associated with conduct disorder, cigarette and marijuana use.

Discussion

Novel approaches are needed for assessment of adolescents' social competency skills due to the limitations of self-ratings, ratings by parents and teachers, and archival records. Interactive RVH technology represents a potentially promising approach for assessing adolescents' social competency skills as it can simulate real-life social situations with verbal interaction and appropriate body language. A necessary step in the development of this innovative technology is pilot testing with samples of at-risk adolescents and evaluation of the psychometric properties of social competency skills measures derived from subjects' performance.

This study examined the extent to which a sample of at-risk African-American male adolescents would be engaged by a novel set of interactive and confrontational VR vignette exercises that featured a virtual teenage character in a school setting. This study also examined the validity and reliability of 18 social competency skills performance measures related to the exercises. Our findings suggest that the majority of study participants were at least moderately engaged by the VR vignette exercises, while half were highly engaged. Qualitative feedback from study participants also indicated that the vignettes were representative of real-life events and that the virtual character was real enough for most participants to engage in each vignette as they would in real life.

Results of exploratory factor analysis indicated that most of the performance measures (12) have adequate factorial validity. Six measures of adolescents' emotional control loaded on one factor, while six measures of interpersonal communication loaded on a second factor. Contrary to expectations, measures of information seeking and negotiation did not load strongly on any factor, raising questions about the validity of those measures. Additional descriptive analyses revealed that the majority of adolescents did not seek information from or attempt to bargain with the virtual teenage character during any of the three exercises, which may help to explain the lack of correspondence between these measures and other measures of interpersonal communication. One possible explanation for these unexpected findings is that few of the adolescents had received appropriate training in these key dimensions of social competency skills. Further qualitative analysis is needed, however, to better understand whether and how utterances classified as information-seeking or negotiation may have differed from other types of verbal interaction.

The Cronbach's α coefficients obtained for the multi-item measures of emotional control (0.88) and interpersonal communication skills (0.91) provided support for the internal

reliability of each measure, indicating that a substantial proportion of the total variance of each multi-item measure was attributable to a common underlying construct.

Correlational analyses with the established measures of beliefs supporting aggression, aggressive conflict-resolution style and hostility provided support for the construct validity of both the emotional control and interpersonal communication skills measures, as they were inversely associated with nearly all three of the psychosocial variables. Insofar as the three established psychosocial measures reflect adolescents' propensity for aggressive behavior (Paschall and Flewelling, 1997; Paschall and Hubbard, 1998), it is not surprising that the factor measure for emotional control was more strongly associated with two of the three psychosocial variables than the interpersonal communication skills measure.

A different pattern of results was observed for the behavioral criterion measures, as the interpersonal communication skills measures was inversely associated with three of the behavioral criteria (conduct disorder, cigarette and marijuana use), while the emotional control measure was not significantly associated with any of the behavioral criteria at the 0.05 level. However, the emotional control measure was inversely associated with lifetime alcohol use at the 0.10 level, and trends in the expected directions also were observed for conduct disorder and marijuana use.

Findings of this study should be considered in light of several potential limitations. Although our sample size met minimum requirements for multivariate analyses, we would be more confident about factor analysis results with a larger sample. Ideally, the sample size should be large enough to divide in half to determine whether factor analysis results are reliable (DeVellis, 1991). Another possible study limitation is our reliance on behavioral criterion measures that were based on data collected from study participants in Grade 9 and from their parents in previous years. Using concurrent (Grade 10) behavioral measures might have yielded greater support for the criterion-related validity of the emotional control and interpersonal communication skills measures.

Although findings of this study provide preliminary evidence that the novel VR vignette exercises can be used to assess at-risk adolescents' social competency skills, additional development and research is needed. More qualitative work is needed to ensure that the exercises are representative of risky social situations experienced by urban minority youth. Further development of an automated performance rating system is necessary so that external observers are not needed to gauge adolescents' performance on the exercises. Additional research with longitudinal data is also needed to further evaluate the psychometric properties of the VR vignette performance measures. Such research will help to determine whether the vignette performance measures have adequate test–re-test reliability and predictive validity with respect to substance abuse and violence-related behaviors that are targeted by prevention programs.

Acknowledgements

This study was supported by a grant from the National Institute on Drug Abuse (R01 DA14813). The authors would like to thank Dr Nicholas Ialongo, Director of the Johns Hopkins University Preventive Intervention Research Center, for his assistance in providing access to study participants and data already collected from study participants for other prevention research projects.

Appendix

Table AI.

VR vignette performance rating criteria

Performance measure	Scoring
Engagement	*Very low* = Participant does not seem to enjoy the vignettes and does not take the virtual character seriously. Seems to only be responding to the virtual character because he has to. May tell the experimenter, 'I can't talk to a computer'. *Low* = Participant is doing the vignettes, but not really getting into them. *Moderate* = Participant does not seem to like or dislike doing the vignettes. *High* = Participant seems to like the vignettes and is engaged. May tell the experimenter he likes the vignettes. *Very high* = Participant is extremely engaged in the vignette exercises. May tell the experimenter the vignettes were awesome and that he wishes they were longer.
Verbalizations	*Very low* = Two responses—participant responds by only saying yes/no. *Low* = Three responses—participant says only a little more than yes/no. *Moderate* = Four short responses—participant does not go into too much detail when responding to virtual character. *High* = Five responses—participant may have said yes or no, but also provided virtual character with detailed information (i.e. questions, preferences, excuses). *Very high* = Six responses—participant holds a conversation with virtual character, providing him with information about preferences, lots of negotiation, perhaps even provocative statements. This participant never limits his responses to a simple yes/no and stands out from other kids because of the amount he says.
Emotional control	*Very low* = Not calm at all, very emotional. *Low* = Not too calm but was able to control emotions to a small degree. *Moderate* = Somewhat calm/emotional state went both ways. *High* = Calm/not too emotional. *Very high* = Very calm/not emotional at all.
Information seeking	*Very low* = Participant did not ask virtual character any questions. *Low* = One question asked. *Moderate* = Two questions asked. *High* = Three questions asked. *Very high* = Four or more questions asked.
Providing information about own preferences	*Very low* = Study participant says only 'yes' or 'no' to the virtual character. *Low* = Participant provides one reason why he does not wish to go along with virtual character. *Moderate* = Participant provides two reasons why he does not wish to go along with virtual character. *High* = Participant provides three reasons why he does not wish to go along with virtual character. *Very high* = Participant provides four or more reasons why he does not wish to go along with virtual character.
Negotiation	*Very low* = Study participants does not use any negotiation or bargaining statements (e.g. asking virtual character what can be done to avoid a physical fight). *Low* = One negotiation or bargaining statement. *Moderate* = Two negotiation or bargaining statements. *High* = Three negotiation or bargaining statements. *Very high* = Four or more negotiation or bargaining statements.
Being non-provocative	*Very low* = Four or more provocative or oppositional statements by study participant and/or study participant curses at the virtual character multiple times. *Low* = Three provocative/oppositional statements and/or study participant may curse at virtual character one time. *Moderate* = One or two provocative/oppositional statements and study participant does not curse at virtual character. *High* = No provocative/oppositional statements, but study participant's tone could be more calm. *Very high* = No provocative/oppositional statements and study participant is not provocative in any way.

References

Botvin GJ (2000) Preventing drug abuse in schools: social and competence enhancement approaches targeting individual-level etiologic factors. Addictive Behaviors, 25, 887–897. [PubMed: 11125777]

Botvin GJ, Schinke SP, Epstein JA, Diaz T and Botvin EM (1995) Effectiveness of culturally focused and generic skills training approaches to alcohol and drug abuse prevention among minority adolescents: two-year follow-up results. Psychology of Addictive Behaviors, 9, 183–194.

Cassell J, Bickmore T, Vilhjálmsson, H. and Yan, H. (2000) More than just a pretty face: affordances of embodiment. In Proceedings of 2000 International Conference on Intelligent User Interfaces, New Orleans, LA, pp. 52–59.

Derogatis LR, Rickels K and Rock AF (1976) SCL-90 and the MMPI: a step in the validation of a new scale. British Journal of Psychiatry, 128, 280–289.

DeVellis RF (1991) Scale Development: Theory and Applications Sage, Newbury Park, CA.

Frankfort-Nachmias C and Nachmias D (2000) Research Methods in the Social Sciences, 6th edn. Worth, New York.

Gottfredson DC and Wilson DB (2003) Characteristics of effective school-based substance abuse prevention. Prevention Science, 4, 27–38. [PubMed: 12611417]

Guerra NG and Slaby RG (1990) Cognitive mediators of aggression in adolescent offenders: 2. Intervention. Developmental Psychology, 26, 269–277.

Hammond WR and Yung B (1991) Preventing violence in at-risk African-American youth. Journal of Health Care for the Poor and Underserved, 2, 359–373. [PubMed: 1772948]

Hubal RC, Frank GA and Guinn CI (2003a) Lessons learned in modeling schizophrenic and depressed responsive virtual humans for training. In Proceedings of the 2003 International Conference on Intelligent User Interfaces, ACM Press, New York, pp. 89–92.

Hubal RC, Deterding RR, Frank GA, Schwetzke HF and Kizakevich PN (2003b) Lessons learned in modeling pediatric patients. In Westwood JD, Hoffman HM, Mogel GT, Phillips R, Robb RA and Stredney D (eds), NextMed: Health Horizon IOS Press, Amsterdam, pp. 127–130.

Kizakevich PN, McCartney ML, Nissman DB, Starko K and Smith NT (1998) Virtual medical trainer: patient assessment and trauma care simulator. In Westwood JD, Hoffman HM, Stredney D and Weghorst SJ (eds), Medicine Meets Virtual Reality—Art, Science and Technology: Healthcare (R)evolution IOS Press, Amsterdam, pp. 309–315.

Nunally JC (1978) Psychometric Theory McGraw-Hill, New York.

Olsen DE (2001) The simulation of a human for interpersonal skill training. In Proceedings of the 2001 Office of National Drug Control Policy (ONDCP) International Technology Symposium, San Diego, CA.

Paschall MJ and Flewelling RL (1997) Measuring intermediate outcomes of violence prevention programs targeting African-American male adolescents: an exploratory assessment of the psychometric properties of six psychosocial measures. Health Education Research, 12, 117–128. [PubMed: 10166899]

Paschall MJ and Hubbard ML (1998) Effects of neighborhood and family stressors on African-American male adolescents' self-worth and propensity for violent behavior. Journal of Consulting and Clinical Psychology, 66, 825–831. [PubMed: 9803701]

Robertson GG, Card SK and Mackinlay JD (1993) Three views of virtual reality: nonimmersive virtual reality. IEEE Computer, 26, 81–83.

Slaby RG and Guerra NG (1988) Cognitive mediators of aggression in adolescent offenders: 1. assessment. Developmental Psychology, 24, 580–588.

Slaby RG and Wilson-Brewer R (1992) Vignettes Developed for the Evaluation of Violence Prevention Curricula Targeting Middle School Students Education Development Center, Newton, MA.

Winters KC, August G and Leitten W (2003) Preventive interventions for externalizing disorders in adolescents. In Romer D. (ed.), Reducing Adolescent Risk: Toward an Integrated Approach Sage, Thousand Oaks, CA, pp. 139–146.

Yung B and Hammond WR (1998) Breaking the cycle: a culturally sensitive prevention program for African-American children and adolescents. In Lutzker J (ed.), Handbook of Child Abuse Research and Treatment New York, Plenum, pp. 319–340.

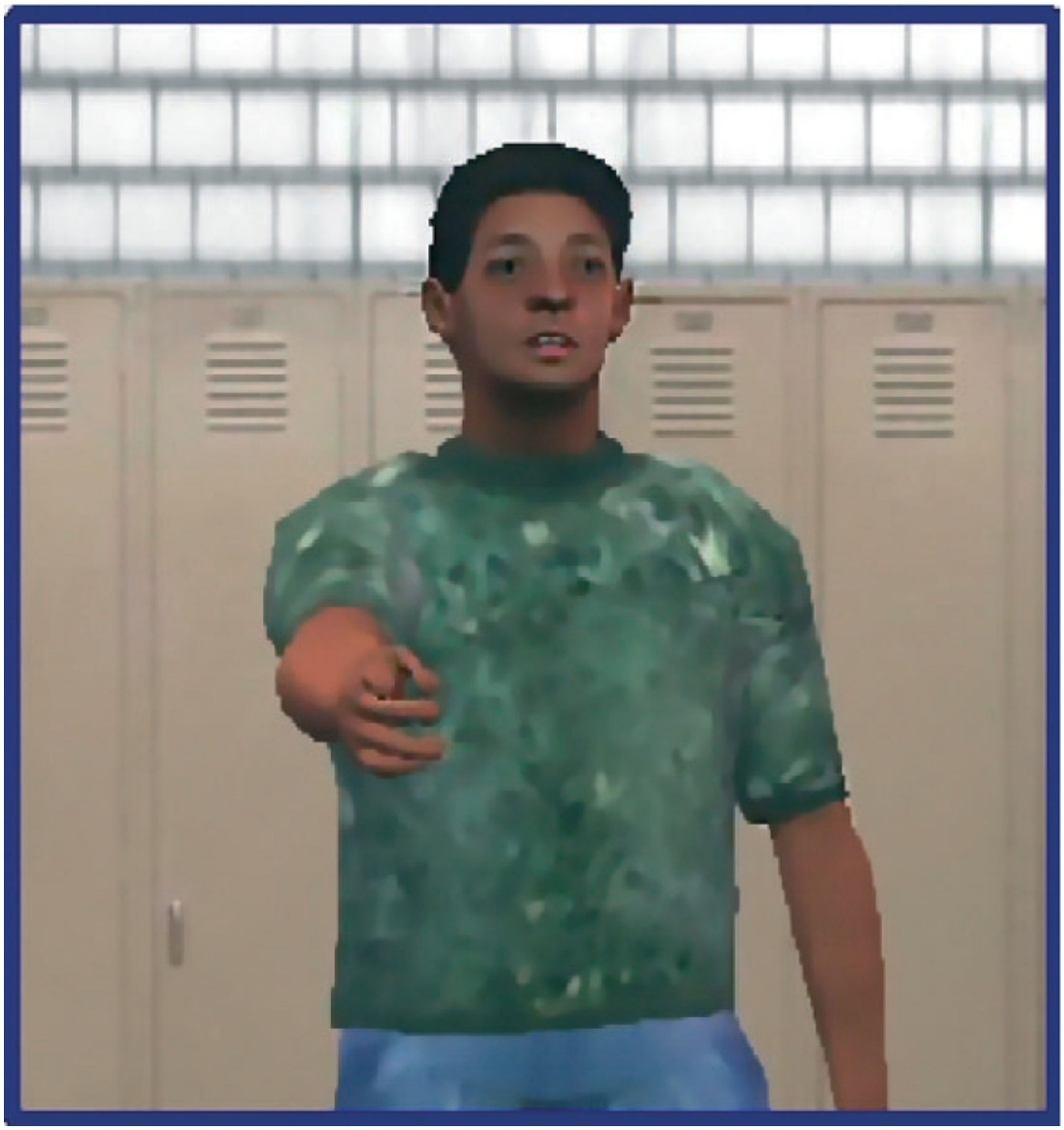

Figure 1.
Virtual teenage character.

Table I.

Frequency (%) distribution and mean (SD) level of engagement for each virtual reality vignette exercise (N = 117)

Engagement level	Exercise 1[a]	Exercise 2[a]	Exercise 3[a]
Very low	5 (4.3)	5 (4.3)	3 (2.6)
Low	17 (14.5)	13 (11.1)	11 (9.4)
Moderate	36 (30.8)	41 (35.0)	38 (32.5)
High	34 (29.1)	33 (28.2)	37 (31.6)
Very high	25 (21.4)	25 (21.4)	28 (23.9)
Mean	3.49 (1.1)	3.51 (1.1)	3.64 (1.0)

[a]The number beside each performance measure denotes the corresponding vignette exercise (1 = gym bag with stolen sneakers, 2 = invitation to drinking party, 3 = being provoked to fight).

Table II.

Results of exploratory factor analysis with virtual reality vignette performance measures to assess their factorial validity (N = 117)

Vignette performance measure[a]	Factor loading	
	1	2
Emotional control 1	0.840	−0.046
Emotional control 2	0.829	−0.021
Emotional control 3	0.805	0.045
Being non-provocative 1	0.842	0.051
Being non-provocative 2	0.839	0.059
Being non-provocative 3	0.691	0.145
Expressing own preferences 1	−0.053	0.900
Expressing own preferences 2	0.162	0.824
Expressing own preferences 3	0.030	0.828
Verbalizations 1	−0.084	0.843
Verbalizations 2	0.080	0.722
Verbalizations 3	0.049	0.703
Information seeking 1	0.058	0.540
Information seeking 2	0.011	0.132
Information seeking 3	−0.078	0.062
Negotiation 1	−0.068	0.597
Negotiation 2	0.202	0.475
Negotiation 3	0.541	0.329

[a]The number beside each performance measure denotes the corresponding vignette exercise (1 = gym bag with stolen sneakers, 2 = invitation to drinking party, 3 = being provoked to fight).

Table III.

Associations between multi-item factor scores, psychosocial and behavioral variables: assessment of construct and criterion-related validity[a]

Psychosocial and behavioral variables	Factors	
	Emotional control	Interpersonal communication skills
Beliefs supporting aggression	−0.29[d]	−0.14[b]
Aggressive conflict-resolution style	−0.32[d]	−0.17[c]
Hostility	−0.20[c]	−0.25[d]
Conduct disorder		
yes	−0.06	−0.17[c]
no	0.06	0.15
Alcohol use		
yes	−0.13[b]	−0.05
no	0.17	0.10
Cigarette use		
yes	−0.03	−0.25[c]
no	−0.03	0.16
Marijuana use		
yes	−0.12	−0.37[d]
no	0.02	0.21

[a]Sample sizes for each set of analyses were 117 (psychosocial and conduct disorder variables) and 104 (drug use variables). Correlation coefficients are reported for psychosocial variables and factor score means are reported for drug use variables.

[b]$P < 0.10$,

[c]$P < 0.05$,

[d]$P < 0.01$.