British Journal of Educational Technology Vol 49 No 4 2018 742–76
doi:10.1111/bjet.12646

Autistic youth in 3D game-based collaborative virtual learning: Associating avatar interaction patterns with embodied social presence

Xianhui Wang, Wanli Xing and James M. Laffey

Xianhui Wang is an assistant professor at School of Journalism and Communication, Central China Normal University. She earned her Ph. D. from University of Missouri-Columbia, USA. Her research interests include 3D collaborative learning, human–computer interaction, learning analytics and technology use in the young generations. Wanli Xing is an assistant Professor in Instructional Technology at Texas Tech University, USA with background in learning sciences, statistics and computer science. His research interests are learning analytics, and CSCL. James Laffey is the Professor Emeritus in the School of Information Science and Learning Technologies at the University of Missouri. He earned his Ph.D. from the University of Chicago and prior to MU, Laffey worked for Apple Computer, Inc. conducting research on learning and support systems and developing award winning interactive learning systems. Since coming to MU, Laffey has been the PI for grants of nearly $9.5 million and a collaborator on grants of an additional $4 million. Address for correspondence: Xianhui Wang, School of Journalism and Communication, Central China Normal University, Wuhan, Hubei 430079, China. Email: xw7t4@mail.missouri.edu

Abstract

This study examines interaction patterns in a series of game activities for learning social skills by autistic youth in a 3D game-based collaborative virtual learning environment (CVLE). Researchers studying collaborative learning have indicated the importance of social interactions and social presence. However, few studies have examined the relationship of avatar social interactions with embodied social presence in a 3D game-based CVLE. Specifically, we examined avatar-mediated verbal and nonverbal interactions by autistic youth in iSocial, a game-based 3D CVLE for developing social competencies. How are avatar-mediated verbal and nonverbal social interactions related to the extent of embodied social presence? Building on prior studies on embodied social presence (ESP) theory and the verbal and nonverbal social interaction framework, this paper aims to explore the link between the combination of verbal (appropriate and inappropriate) and nonverbal (avatar proximity, orientation, joint attention and gesture) interaction patterns and experienced ESP level. We report on the results of 3 cohorts of a total of 11 youth aged from 11 to 14 who were diagnosed with Autism Spectrum Disorder learning in 13 game activities in iSocial. The video data of the participants' screen recordings were analyzed and coded based on the ESP theory and verbal and nonverbal social interaction framework. Through cluster analysis, the results identify distinct patterns of verbal and nonverbal interaction that are associated with different levels of embodied social presence. The findings of this study (1) shed light on the link between social interactions and embodied social presence and (2) provide a deeper understanding of how the unique spatial and visual characteristics of 3D CVLE and the design of game activities in 3D CVLE may transform collaborative learning, especially for autistic youth.

Introduction

3D game-based collaborative virtual learning environments (CVLEs) are used to motivate and enrich learning across multiple disciplines (Barab, Thomas, Dodge, Carteaux, & Tuzun, 2005;

Practitioner Notes

What is already known about this topic

- Social interaction plays an important role in 3D game-based collaborative virtual learning.
- Sense of social presence in the avatar-mediated, game-based collaborative learning experiences is essential for learning gains.
- Utilizing 3D game-based collaborative learning virtual environment to teach young people with Autism Spectrum Disorder has promising preliminary results.

What this paper adds

- This study characterized autistic youth's verbal and nonverbal social interaction patterns in the game-based learning activities in 3D virtual environment.
- The author explored the link of the distinct verbal and nonverbal social interaction patterns with embodied social presence in 3D game-based collaborative learning virtual environment.
- This article discussed how the game-design principles, spatial and visual technology features of 3D virtual environment that could support autistic youth's collaborative learning.

Implications for practice and/or policy

- We need to re-conceptualize ideas around collaborative learning in 3D game-based collaborative learning virtual environment, in particular away from more traditional approaches and towards a notion of learning as more centered upon experience and exploration in order to support autistic youth's collaborative learning in the avatar embodied, game-embedded 3D virtual environment.
- The lessons from this study for practitioners and educators is the importance of considering the unique characteristics of social interactions in the 3D game-based collaborative learning virtual environment, and designing learning experiences that encourage the desired verbal and nonverbal interaction pattern.

Burgess, Slate, Rojas-LeBouef, & LaPrairie, 2010). The potential benefits of using a 3D game-based CVLE to teach social competence for autistic youth include engaging distant learners in the game-based role-playing learning activities and fostering social interactions among avatar-mediated learners in a simulated, curriculum-enriched environment without suffering real-world consequences and in some formats without trying the patience of their peers or teachers (Schmidt, Laffey, Schmidt, Wang, & Stichter, 2012; Standen & Brown, 2006; Wang, Laffey, Xing, Ma, & Stichter, 2016). Students with high functioning autism, are known to "have a desire to be social but do not yet have the knowledge or skills to successfully perform interactions in a complex and social environment" (DuCharme & Gullotta, 2003). In comparison to typically developing youth, autistic youth make and receive fewer social initiations, engage in shorter episodes of social interaction, have lower quality interactions, spend less time in proximity to peers, and spend more time in nonsocial play (McGee, Almeida, Sulzer-Azaroff, & Feldman, 1992; Sigman & Ruskin, 1999).

There is considerable evidence that students with autistic spectrum condition can learn targeted social skills in 3D virtual learning environment. Research supports that these individuals can learn and generalize social appropriateness in social settings (Parsons & Mitchell 2002; Parsons *et al.*, 2005; Rutten *et al.*, 2003), other targeted social skills such as emotion recognition through avatar representations (Fabri & Moore, 2005), and behaviors that facilitate social acceptance such as eye contact and attending (Cheng & Ye, 2010).

The context of this study—iSocial, a 3D game-based CVLE, is a translation of a face-to-face, clinic-based curriculum-social competence intervention-adolescents (SCI-A), for delivery over the Internet. SCI-A (Stichter *et al.*, 2010) targets impairments of youth with High Functioning Autism/Asperger Syndrome (HFA/AS) in three social cognition processes: theory of mind, emotion recognition and executive functioning. The learning outcome results of preliminary studies have shown promise for autistic youth to learn social competence in iSocial (Laffey, Stichter, & Galyen, 2014; Stichter, Laffey, Galyen, & Herzog, 2014).

However, there is much to learn about how autistic students interact with peers in a 3D game-based CVLE, and what are the mediating factors/constructs that affect learning outcomes. One mediating construct that was recently discovered to be associated with effective learning in 3D worlds is embodied social presence (ESP), proposed by Mennecke, Triplett, Hassall, Conde, and Heer (2011).

> ESP is premised on the notion that certain communication acts and interactions take place in the context of embodied states that create a sense of presence that is derivative of human cognitions associated with physical, real world body-to-body interactions. To achieve this sense of ESP, one must first achieve and perceive sufficient levels of embodied presence and copresence in the 3D CVLE.

Compared to embodied presence and copresence, ESP represents a higher level of perceptual engagement that learners experience as they engage in activity-based social interaction in the virtual environment. ESP focuses not only on an individual's development of cognitive engagement but also on how social interaction influences the development of cognitive engagement and one's perception of embodiment.

The researchers have confirmed that avatar-mediated social interactions can aid autistic youth in developing embodied social presence (Aymerich-Franch, 2010). Wang, *et al.* (2016) described autistic youth's embodied presence, copresence and social presence in iSocial game-based learning activities. The findings of this study also show that autistic youth experienced embodied social presence only in some of the activities in iSocial. Based on that study, we think it is important to know what are the social interaction characteristics that are associated with experienced embodied social presence, and what are the social interaction characteristics that are associated with only the lower levels of ESP: the embodied presence and copresence?

Learners' avatar-mediated interactions in online 3D immersive virtual worlds have been reported to resemble real-life interactions yet having their own unique characteristics (Allmendinger, 2010; Bailenson *et al.*, 2008; Yee, Bailenson, Urbanek, Chang, & Merget, 2007). This seems also to be the case with autistic students learning in the 3D CVLE. Parsons, Mitchell, and Leonard (2005) found that teenagers with autistic spectrum condition demonstrate a sufficient level of imagination to impute human behaviors to virtual people. For example, they typically said that two virtual people standing in proximity and oriented toward each other in a characteristic way, were "having a conversation." Schmidt *et al.* (2012) categorized autistic youth's interactions in the 3D CVLE into verbal interaction and nonverbal interactions such as movement, gesture, action, and text. Wang, Laffey, Xing, Galyen, and Stichter (2017) further built on this social interaction framework and identified nonverbal features (gesture, proximity, orientation and joint attention movement) that supplement the verbal communications. Through the close examinations of these discrete behaviors, researchers found that verbal communication and nonverbal signals via the avatars sometimes lead to mutual engagement of the task, but at other times cause distraction and diminished sense of self, others and the learning activity (Allmendinger, 2010; Bailenson *et al.*, 2008; Schmidt *et al.*, 2012). The focus of this paper is on whether certain patterns of verbal and nonverbal behaviors associate with the development of embodied social presence and thus help to better structure social interaction for autistic youth in game-based learning situations in 3D CVLE. The main research question of this study is:

What are the social interaction patterns that associate with better experiences of embodied social presence?

Theoretical background

3D game-based collaborative learning

Studies have been carried out to measure the outcome of 3D game-based collaborative learning (Barab *et al.*, 2005; Berns, Gonzalez-Pardo, & Camacho, 2013; Hamalainen, 2008). Merchant, Goetz, Cifuentes, Keeney-Kennicutt, and Davis (2014) did a meta-analysis of the effectiveness of virtual reality-based instruction. Results suggest games (FEM = 0.77; REM = 0.51), simulations (FEM = 0.38; REM = 0.41), and virtual worlds (FEM = 0.36; REM = 0.41) were effective in improving learning outcome gains. Games show higher learning gains than simulations and virtual worlds. Researchers also found that autistic young people enjoy playing games in the 3D virtual learning environment. Leonard *et al.* gave the youth tasks to find a place to sit in a virtual café and a virtual bus. The researchers reported on whether the autistic youth enjoy using the VE and to provide rich examples of participant/facilitator interaction in the VE. Participants reported that they felt the VE had helped them and that they thought it was useful.

They appeared to be motivated by the VE and stated that they enjoyed using it. When the participants made mistakes, they would laugh and joke about the mistakes, indicating that the social blunders they made may not have been interpreted as being as serious as in real life, nor as anxiety inducing.

With regard to the promising potential of 3D game-based collaborative learning for engaging autistic youth, they are known to have perceptual, cognitive and sensory tendencies that may cast some doubt on the potential effectiveness (Wallace *et al.*, 2010). Specifically, eye movement studies have shown that autistic youth tend to focus on (different) details of visual displays compared to typically developing youth (Riby & Hancock, 2009). Also, some autistic young people are known to experience sensory difficulties or overload in response to particular stimuli (Rogers, Hepburn, & Wehner, 2003). In addition, autistic young people have little aptitude for pretense, which might constrain opportunities for role-play (Lewis, Boucher, Lupton, & Watson, 2000). These factors could undermine the value of 3D game-based collaborative learning for autistic young people relative to their typical peers, or, at the very least, may hinder their chances of perceiving sufficient levels of presence in the 3D CVLE to benefit from the learning opportunities (Wallace *et al.*, 2010). For example, in iSocial, if a student does not have a sense of being there in the virtual world and participating with others in the activities, then they will not practice turn taking or sharing ideas in ways that elicit the cognitive and behavioral expectations of the curriculum (Stichter *et al.*, 2014).

Dalgarno and Lee (2010) conceptualized a theoretical model that shows learning benefits in a 3D CVLE being affected by learners' sense of embodiment, presence and co-presence in a 3D CVLE, which in turn are affected by technology representations of the environment and learner's social interaction within the environment. However, Dalgarno and Lee's theoretical model on 3D learning affordances was built on a review of previous literature, without empirical evidence. Extended from Dalgarno and Lee's work, Mennecke *et al.* (2011) examined sense of embodiment, the social context, presence and copresence, hence, developed embodied social presence theory.

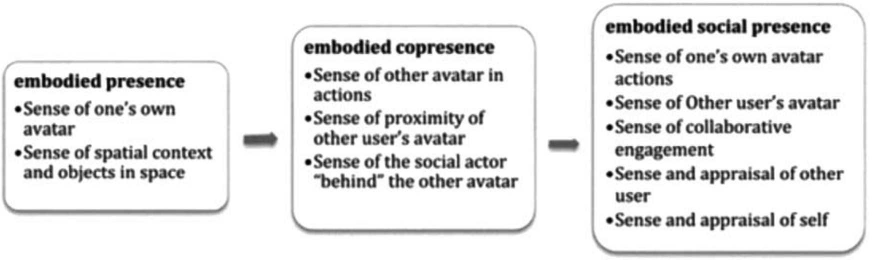

Figure 1: Embodied social presence (*Wang* et al., *2016*).

Embodied social presence theory

Mennecke *et al.* (2011) did a qualitative analysis of factors associated with ESP theory to identify the process(es) by which the experiences of ESP are derived and examined the results of this phenomenon on social engagement, collaboration, and interactions. Fifty-seven students' reflections in learning activities in the Second Life were collected and analyzed. Content analysis was used to analyze the text, and has been used to discover the psychological, attitudinal, and behavioral states of individuals and groups. The researchers analyzed the collected data by conducting focus analysis, followed by theme development and then operationalized ESP development. Three categories (ESP achieved, neutral and not achieved) were identified based on the distinction between the narrative mode of first-person, and third-person. ESP was achieved by 68% of the students. In addition, a step process model, focused on the conditions needed to achieve ESP, was developed.

Mennecke and colleagues found that in 3D collaborative environment, the learner is presented with stimuli representing the virtual learning environment, the objects in that environment, and their own avatar's representation. If the learner engages with these stimuli he or she will experience, to one degree or another, embodied presence. When leaners share a virtual space they have the opportunity to experience embodied copresence. For embodied social presence to occur, learners in 3D virtual environment must participate in goal-directed, shared activity mediated through embodied representations in a context (Allmendinger, 2010; Mennecke *et al.*, 2011). See Figure 1 for a schematic representation of the embodied social presence development framework.

Promoting autistic youth's embodied social presence in the gameplay learning activities in 3D CVLE is extremely important given that they need support and social interactions as part of the learning process. In addition, ESP is associated with high levels of cognitive engagement, a focus on shared activities and spaces, on the actions exhibited by the virtual and real bodies of self and others, and on perceptions and interpretations of intent and content from verbal and nonverbal communication.

Verbal and nonverbal social interactions

One consistent area of need for autistic students is in the area of social competence, particularly in regards to understanding and using basic rules of social engagement (Stichter *et al.*, 2014). One of the goals of face-to-face social intervention for youth or children with autistic spectrum condition is facilitating social interactions. A number of studies have examined whether the participants show an increase in the frequency and quality of social initiations and responses after or during a social intervention implementation (Freedman & Silverman, 2008; Odom & Strain, 1986).

Schroeder, Heldal, and Tromp (2006) pointed out, in 3D collaborative virtual learning, nonverbal discrete behaviors should be interpreted in the context of the social interactions and activity. The combination of meaning for the verbal communication and the non-verbal communication can give an enriched understanding of reciprocal social interaction in a 3D CVLE (Schmidt *et al.*, 2012; Yee *et al.*, 2007)

Researchers have developed methods to understand social interactions of autistic youth learning in game-based activities in iSocial (Schmidt *et al.*, 2012; Wang, *et al.*, 2017). It is common to observe and identify discrete behaviors first and then analyze social interaction behaviors with frequency and sometimes sequence (Ang, Bobrowicz, Siriaraya, Trickey, & Winspear, 2013; Tromp, Steed, & Wilson, 2003). Wang *et al.* (2017) examined verbal interaction and four nonverbal features: avatar orientation, avatar proximity, avatar gesture and avatar movement to indicate joint attention. This two-dimensional method that looks into the co-occurrence of verbal and nonverbal interactions, takes into consideration that interactions in 3D environments resemble face-to-face interactions with avatars as a substitute for the human body (Bauminger, 2002; Cobb *et al.*, 2002; Cole & Griffiths, 2007).

Social interactions in 3D collaborative game-based learning are essential for the development of sense of embodied social presence (Allmendinger, 2010; Aymerich-Franch, 2010; Mennecke *et al.*, 2011). However, avatar-mediated interaction behaviors in a 3D game-based CVLE often have been seen to be easily distracted from the collaborative tasks required for embodied social presence to develop (Micaela Esteves, Fonseca, Morgado, & Martins, 2011; Montoya, Massey, & Lockwood, 2011). We need to identify social interaction behavior patterns that are associated with higher quality of embodied social presence.

Linking interaction behavior patterns to experienced embodied social presence

Previous studies that examined social presence in immersive or video gaming virtual environment have discussed the importance of social interactions, but rarely identified patterns of social interactions that benefit the experience of social presence (Biocca, Harms, & Burgoon, 2003).

Social interactions in 3D game-based learning environments can be as complicated as face-to-face interactions, as or sometimes more complicated. For example, in the castle quest game activity in iSocial, student Jesse said “how are you doing?” in the meantime he was waving his avatar hand and walking his avatar body toward student Mike’s avatar. Mike was not responding because he was busy changing his avatar’s hair color. After a while, Mike responded “fine, you?” with his avatar not facing and was far away from Jesse’s avatar. As we can see from the example, interaction behaviors have both verbal and nonverbal elements, and should be interpreted within the social context (Bailenson *et al.*, 2008). However, prior studies and literature do not provide much guidance for the question of what type of activity-based, body-centered interactions among autistic youth in the 3D game-based CVLE will positively influence embodied social presence.

In this study we focus on investigating patterns of social interaction behaviors that facilitate the development of sense of embodied social presence in a 3D game-based CVLE.

Method

Participants

Participants are 11 youth aged from 11 to 14 who were diagnosed with

Asperger’s syndrome by Autism Diagnostic Interview Revised (ADI-R) (Le Couteur, Rutter, & Lord, 2003) and/or the Autism Diagnostic Observation Schedule (ADOS) (Lord, Rutter, DiLavore, & Risi, 2002). Additionally, an IQ of 75 or above and capable of speech are required for participation. These 11 participants are all male. Participants were from three different junior high or middle schools in the mid-west area of the United States. The iSocial study received IRB approval and student participation was approved by district administrators, principals, teachers, parents and youth. Selection for participation in the study was based on meeting diagnostic criteria for Asperger’s syndrome, determination of need for services by administrators and parents and assent

Figure 2: Screenshots of students at work in NP game-based learning activities in iSocial.
[Colour figure can be viewed at wileyonlinelibrary.com]

by students. All students who were invited to participate did so. Participants in the same school district were formed into a cohort and participated in the iSocial curriculum together. Cohort 1 has three students and Cohorts 2 and 3 have four students. These three cohorts were taught by the same teacher, an online guide (OG) in the 3D CVLE, who is highly trained in teaching social competencies for youth with autistic spectrum condition. The three cohorts went through the same iSocial curriculum.

iSocial curriculum and game-based naturalistic practice activities

As mentioned previously, iSocial is a 3D game-based intervention for social and behavioral outcomes for autistic adolescents. iSocial seeks to translate and implement the SCI-A curriculum (Stichter *et al.*, 2010) in the game-based virtual environment. The 10-week curriculum, SCI-A is based on a framework of Cognitive Behavioral Intervention (SCI-CBI). Initial results from face-to-face implementations of SCI-CBI indicate promising trends for growth (across pre- and post-intervention assessments) among autistic adolescents (Stichter, Randolph, Gage, & Schmidt, 2007). In iSocial lessons, each user is represented by one's own avatar. Autistic students can interact with peers and the OG using both verbal and nonverbal communications via avatars in the game-play learning activities. Game-based learning activities in iSocial provide opportunities for students to manipulate virtual objects and select options in the 3D space. Goal-oriented game tasks stimulate discussions and negotiations among students. For example, in the Unit 3 lesson 6, students were building the buffet menu in a virtual restaurant. Student took turns to manipulate the options tool in order to show the virtual dishes on the buffet table (see figure 2). To complete the task, they need to set all the menu items together for the restaurant.

The context for this case study was a field test of the iSocial curriculum in three school districts. The iSocial curriculum lasted for 4 months. Each cohort had approximately 2 lessons per week, with 45 minutes per lesson. Students were first trained on how to navigate avatars, interact with objects, use curricular tools, and follow behavior rules in the 3D game-based CVLE during two orientation sessions. Then, they went through 5 units of curriculum: facial expression, sharing ideas, turn taking in conversation, emotions and feelings, and problem solving, with each unit building upon the previous learnt skills. Each unit contained 5–7 lessons. In each unit, the lesson plan follows a consistent structure of learning and rehearsing skills culminating in a Naturalistic Practice activity where the students put their new competencies into practice in a challenging activity meant to engage them with their peers in a gameplay task, See the example screenshots in Figure 2. The 13 Naturalistic Practice (NP) game-based activities are selected as the data source for this study. The NP activities were primarily student-led, goal-oriented and game-based activities that took place in 3D virtual scenarios that were embedded with narratives, role-

Table 1: Description of NP activities in iSocial

#NP	Goal of the game	Game description
1	Facial expression scenarios	Role-playing the given scenarios, discuss facial expressions in group and take a selfie demonstrating the discussed facial expression
2	Share Out	Sharing with group on facial expressions and reach consensus
3	Lost at sea: take items	In a virtual ship, each player has to make decisions together on what items to take in order to survive a sinking ship
4	Lost at sea: go to island	On the deck of a sinking ship, players has to choose an island as a group to "survive"
5	Sell it!	Selling the idea like a real business salesman. Taking turns to do the pitch
6	Restaurant buffet I—selecting main dishes, beverages	As hypothetical restaurant owners, students are making decisions on main dishes and beverages for the buffet
7	Restaurant buffet II—selecting desserts, side dishes	As hypothetical restaurant owners, students are making decisions on desserts and side dishes for the buffet
8	Identify emotion status activity	Sharing emotion on a collaborative 3D sticky note
9	Emotional role-play	Discussing role plays for different emotions
10	Role-play planning and taping	Act out the play in front of a "real" camera in a virtual studio
11	Watch and rate role plays	Watch videos in the virtual world and rate the videos
12	Plan Quest activity	In a kings' castle, planning the quest on an interactive map as a group
13	King's Quest	Finding the king's missing items in an 18th-century castle as a group, and get knighthood in the game

playing and interactions with the virtual objects. In these activities, students were supported by the OG, to collaboratively complete certain learning tasks that require using the social skills learned in the prior units and lessons. See a description of the game-based NP activities in Table 1.

Data collection

The screens of every student and the OG were recorded for the entire set of lessons using ScreenFlow – a video recording software. For this study, the recordings of each student in every NP game-based learning activities were collected. There should have been 143 observations of video data (11×13). However, if two or more students were absent or having technical difficulties that prevented them from fully participating in a NP activity, all of the video data for that NP activity were excluded. If only one student was absent, the rest of the students' videos remained as the data for this study. These rules for inclusion of interaction enabled 118 observations in total.

Data preparation

Embodied social presence

The ESP framework (Mennecke *et al.*, 2011) is used to guide coding of autistic youth's experience of embodied social presence in the game-based 3D CVLE.

The 118 video data observations were coded based on whether the participant showed evidence of (1) embodied presence which includes "sense of avatar representation of self" and "sense of virtual objects/context;" and (2) embodied co-presence which includes "sense of avatar representation of other people" and "sense of avatar proximity." Moreover, to the extent that these first 2 levels of presence could be shown, the author examined (3) embodied social presence which includes "sense of collaborative engagement," "sense and appraisal of others," and "sense of one's own actions as manifest in avatar-based social interaction." The coding definitions and examples are in Appendix A. The coding work was completed by three raters (research assistant doctoral students), 20% of each rater's work was coded by another rater. The overall inter-rater agreement was 97% (IRA% is calculated as (#agree)/(#agree + #disagree) $\times$ 100).

Results showed that only 61% of the time students achieved embodied social presence across the 13 NP activities. There are 3 out of 13 NP activities in which 100% of students achieved embodied social presence. However, in some NP activities, like NP #1, and #2, none of the students achieved embodied social presence. To achieve sense of collaborative engagement the student must show evidence of participating in the collaborative discussion in NP activities. Findings showed that on average, 83% of students in all NP activities achieved sense of collaborative engagement. Results showed that students achieved sense and appraisal of others by making evaluations or suggestions/comments on others' opinion during the collaborative learning activity. For example, saying "good job," "I agree" to other members in the group. On average, 76% of students in all NP activities achieved sense of one's own action and appraisal of others. Sense of one's own action as manifest in avatar-based interaction is achieved when students express their own opinion as a contribution to the collaborative discussion process in an embodied state, which is through the action of their own avatars in the 3D CVLE. Results showed that on average, only 67% of students in all NP activities demonstrated sense of one's own action as manifest in avatar-based interaction.

Social interactions

Social interaction includes two dimensions: verbal social interaction and non-verbal behaviors (Wang *et al.*, 2017). Verbal social interaction was further represented with two constructs: Appropriate Verbal (AV), Inappropriate Verbal (IV). Nonverbal social interaction behaviors were dependent upon verbal social interactions, and were observed if concurrent with the verbal. Nonverbal social interaction behaviors include "avatar Orientation (O)," "point or move to show joint attention (J)," "use of gesture (G)" and "avatar body proximity (P)." Autistic youth sometimes used a combination of concurrent verbal and nonverbal social interactions. For example: Roger appropriately answered Jason's question (AV), and Roger's avatar was oriented toward Jason, and kept a good avatar proximity. The code combination for Roger's concurrent verbal and nonverbal behavior can be represented as AVOP.

Eleven autistic young people's social interaction behaviors in the 118 video data observations were coded. The overall agreement for verbal social interaction coding is a Cohen's Kappa of .86. For nonverbal social interaction behaviors, the overall Cohen's Kappa is .92.

The frequency of each code in a certain NP activity for the individual youth was calculated after the coding. There are a total of 32 possible codes, which represents the combination of AV, IV and O, J, G, P.

The durations for the 13 NP game activities ranged from 6 minutes and 30 seconds to 30 minutes and 17 seconds. To allow comparison of the frequency of social interaction behaviors comparable across the 118 observations that have different durations in the following cluster analysis, the researchers preprocessed the data, and calculated each student's interaction frequency on a 10-minute duration period for each observation.

Cluster analysis

We conducted a K-means cluster analysis to identify distinct patterns of social interaction behaviors. The purpose of cluster analysis is to identify groups of objects that have similar properties or characteristics (Hair, Black, Babin, Anderson, & Tatham, 1998). These groups, or clusters, should exhibit high within-group homogeneity and high between-group heterogeneity (Xing, Guo, Petakovic, & Goggins, 2015). Using this technique, it is possible to identify groups that share common social interaction behavior characteristics. Once distinct groups of social interaction behaviors are identified, the embodied social presence levels (embodied presence, embodied co-presence and embodied social presence) across groups can be compared to evaluate how certain social interaction behaviors relate to levels of ESP.

The problem of data segmentation is addressed through the use of unsupervised machine learning methods. An assumption of Cluster analysis is that there is no "preferred" number of clusters. An important step of clustering is to define the cluster elements. In this study, the cluster elements are the social interaction codes (eg, AVOP, AVP, IVGP, etc.) for each autistic youth in each NP game-based learning activity. The data matrix, as a result, is high dimensional (32 dimensions) for analysis. However, clustering analysis is sensitive to the high-dimensional data, which can significantly influence the clustering performance (McCallum, Nigam, & Ungar, 2000). Therefore, before K-means clustering analysis is conducted, principle component analysis is performed for dimensionality reduction. Principle component analysis (PCA) is a typical linear technique for dimensionality reduction in the machine learning field (Hinton & Salakhutdinov, 2006). PCA performs a linear mapping of the data to a lower dimensional space where the variance of the data in the low-dimensional presentation is maximized.

On the other hand, as an unsupervised machine learning algorithm, K-means requires defining the preferred number of clusters in advance. In order to determine an optimal number of clusters used in this research, the Ball statistic is calculated. The Ball statistic is a classic measure to identify the best number for clustering. It is used to measure the dispersion of the data points within a cluster and between the clusters so that the data have the largest difference between clusters and smallest difference within clusters (Milligan & Gooper, 1985). Clustering analysis has the best performance when cluster K is set at the largest value of the successive difference of the Ball index values.

After the PCA for high dimension reduction and deciding the optimum number of clusters using the ball statistic, a K-means clustering algorithm was implemented. Since cluster elements are grouped according to their similarities, or the distances between them, Squared Euclidean distance (Dorling, Davies, & Pierce, 1992) was used in our study to calculate the distance between clusters.

Results

We have identified the combination of verbal and nonverbal codes as our elements (dimensions) of the cluster analysis, with 118 observations from which the clusters will emerge. Figure 3 (a) shows the results of principle component analysis to reduce the high dimensionality. It indicates that the first two components alone can already explain more than 99% of the variance in the dataset. Based on these components, the Ball statistic was calculated. It shows that the K-means clustering would have the best performance if the number of clusters is set to three as shown in Figure 3 (b). Then, the result of three clusters was generated using K-means clustering algorithm and the descriptive statistics are shown in Table 1.

Table 2 presents the cluster analysis results, including the mean values of frequency of the social interaction in the cluster. Cluster 1 has 54 observations, cluster 2 has 24 observations and cluster 3 has 40 observations.

Cluster 2 clearly has the highest frequency of social interaction types including: AVOP, IVOP, AVJP, AVOJP, AVOGP, AV and AVO. Among them, the most dominant behavior is AVOP (appropriate verbal with good proximity and orientation) 17.027 and IVOP (inappropriate verbal

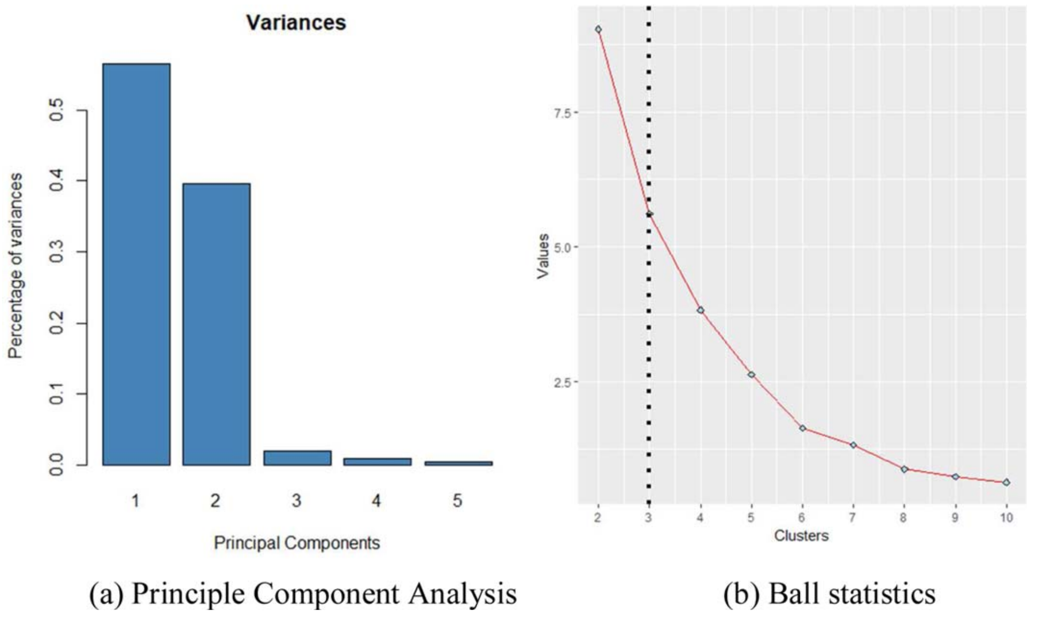

Figure 3: PCA and Ball statistics. *(a) Principle Component Analysis (b) Ball statistics [Colour figure can be viewed at wileyonlinelibrary.com]*

with good proximity and orientation) 1.035. These interaction behaviors all have good avatar orientation and avatar proximity. Besides AVOP and IVOP, the rest of the interaction behaviors (AVJP, AVOJP, AVOGP, AV and AVO) all have very low mean values (below 1.0).

Cluster 3 has the highest frequency of AVP (appropriate verbal with only proximity behavior) 15.463 and IVP (inappropriate verbal with only proximity behavior) 1.331 when compared to the other two clusters. Cluster 3 has the highest frequency of social interaction behaviors which only have the proximity behavior. This means that the most frequent interaction behaviors in observations in cluster 3 do not include orientation (facing) toward each other nor any other body gesture, or body movement to indicate joint attention.

Cluster 1, on the other hand, has the least frequency of almost all the social interaction types among the clusters according to Table 1. However, cluster 1 has slightly higher frequency of AVOP, AVO, AVOJP and AVOGP than cluster 3. This shows that observations in cluster 1 have lower frequency of social interactions in general, but observations in cluster 1 have higher frequency of interactions with both avatar proximity and avatar orientation (facing), or/and avatar body gesture, and/or body movement to indicate joint attention than observations in cluster 3.

From these findings, we learned that there is a set of NP activities as represented in cluster 2, where individuals were frequently using avatars for facing each other, for keeping good avatar distance and for using other nonverbal interaction behaviors such as avatar gesture and avatar movement to indicate joint attention during verbal interactions as they were undertaking an activity. However, in cluster 3 individuals verbally interact with each other while keeping a good avatar distance (proximity) but did much less facing (orientation) each other during the verbal interactions. In cluster 1, individuals have the least social interactions in general among these three clusters. But compare to cluster 3, individuals in cluster 1 interact verbally slightly more frequent with both good avatar proximity and avatar orientation, and/or other nonverbal interaction behaviors such as avatar gesture and avatar movement to indicate joint attention.

Table 2: Cluster analysis results

	Cluster Mean		
	1	2	3
Social interaction patterns			
AVOP (appropriate verbal with good avatar orientation and proximity)	8.02	17.027	3.962
AVP (appropriate verbal with only good proximity)	3.79	14.332	15.463
IVOP (inappropriate verbal with good avatar orientation and proximity)	0.481	1.035	0.401
IVP (inappropriate verbal with only proximity)	0.222	0.967	1.331
AVJP (appropriate verbal with avatar joint attention and proximity)	0.095	0.396	0.185
AV (appropriate verbal with inappropriate proximity and no avatar orientation, no gesture and no joint attention)	0.111	0.509	0.267
AVO (appropriate verbal with only good avatar orientation)	0.125	0.236	0.088
AVOJP (appropriate verbal with good avatar orientation and proximity and joint attention)	0.069	0.3196	0.010
AVOGP (appropriate verbal with good avatar orientation and proximity and joint attention and gesture)	0.064	0.2493	0.038
Cluster size	54	24	40

Having identified distinct patterns of social interaction behaviors in cluster 1, 2, and 3, we examined the relationship between the patterns of social interaction behaviors and the variable—embodied social presence level.

The results in Table 3 show that observations in cluster 2 have best experience of embodied social presence among the clusters. Because the mean value for embodied social presence of cluster 2 is 2.96, very close to 3, which represent embodied social presence. It is safe to say that almost all of the observations in cluster 2 achieved embodied social presence.

Cluster 3 has a mean score of 2.60 and cluster 1 mean score is only 2.37. These results indicated that individuals in cluster 3 and cluster 1 have most likely achieved embodied presence and embodied copresence. However, individuals in cluster 3 and especially cluster 1 were less likely to experience embodied social presence compared to cluster 2. The results of analysis of variance (ANOVA) test comparing the mean values of the ESP variable across clusters are shown in Table 2. The ANOVA test was statistically significant, and F-value was 0.035 ($p < .05$). This examination of the data revealed heterogeneity of variance between the three clusters with respect to embodied social presence variable.

Table 3: Experience of embodied social presence differences

	Cluster Mean			ANOVA F value
	1	2	3	
Embodied Social Presence (3 as embodied social presence, 2 as embodied copresence and 1 as embodied presence)	2.37	2.96	2.60	0.035*
Cluster size	54	24	40	

*$p < .05$.

Discussion and implications

3D game-based collaborative learning environments provide support for multiple learners in different physical locations to interact and collaborate synchronously in game-play learning activities (Barab *et al.*, 2005; Freitas & Neumann, 2009). While most educators understand the importance of social interactions in the learning process, studies have been needed for understanding more about the association between social interaction behavior patterns and the development of embodied social presence, and eventually lead to effective learning.

This study has been built upon research findings of two previous studies that analyzed autistic youth social interactions and embodied social presence in 13 game-based learning activities in iSocial (Wang *et al.*, 2016, 2017). In this article, we explored the link between social interaction behaviors and embodied social presence. We acknowledge that the two constructs share some overlapped social interaction facets, however, the association explored here is between the frequency of specific interaction patterns and sense of embodied social presence. We identified distinct patterns of verbal and nonverbal social interaction behaviors by cluster analysis of the 118 observations across 13 game-based learning activities in iSocial. The results showed that the frequency of these patterns was associated with different levels of embodied social presence. In doing so, we were able to show a deeper understanding of how do we support and structure certain patterns of verbal and nonverbal social interactions and design activities that foster the target form of avatar-mediated interactions for autistic youth learning in 3D game-based virtual environment.

Consistent with our expectations, our results show that the highest ESP score cluster (cluster 2) has the highest frequency of verbal interactions with good avatar proximity (stand not too close nor too far from each other) and avatar orientation (facing the avatar that she or he was interacting with) than the other two clusters with lower ESP scores. This means that when autistic youth have verbal interactions that include realistic nonverbal signals, they are more likely to achieve embodied social presence in the 3D game-based CVLE. We also discovered that cluster 3 has higher frequencies of verbal interactions with only avatar proximity than cluster 2, but cluster 3 did not have better embodied social presence than cluster 2. The reason for this might be that individuals who did not face each other's avatars during verbal social interactions in game-based activities perceived a lower sense of embodied social presence.

The embodied social presence theory identified factors that are fundamentally associated with rich, body-to-body interactions in the real world and applied these to interactions in the 3D virtual environment. This study further confirmed that the more autistic youth's social interactions resemble the real-life social interactions, the higher level of embodied social presence they achieve. The two most frequent nonverbal behaviors: avatar proximity and avatar orientation together helped the avatar-mediated social interactions in the collaborative game tasks in the 3D game-based CVLE. These findings echoed Yee, et al., and Ang *et al.*'s research on avatar-mediated communications (Ang *et al.*, 2013; Yee *et al.*, 2007). The findings also revealed that without avatar orientation (facing toward), which in real life is also very fundamental to social interactions, the experience of embodied social presence will be diminished in the 3D game-based CVLE.

Spatial social behaviors such as proximity and orientation in 3D virtual environments played an important role in social interactions (Friedman, Steed, & Slater, 2007). To facilitate the development of high levels of embodied social presence, we must pay attention to the design of the environment and the teacher/tutor role of how to assist desired patterns of verbal interactions with avatar proximity and avatar orientation. However, when designing game-based learning activities in a 3D CVLE, support for the spatial social behavior it is easily overlooked.

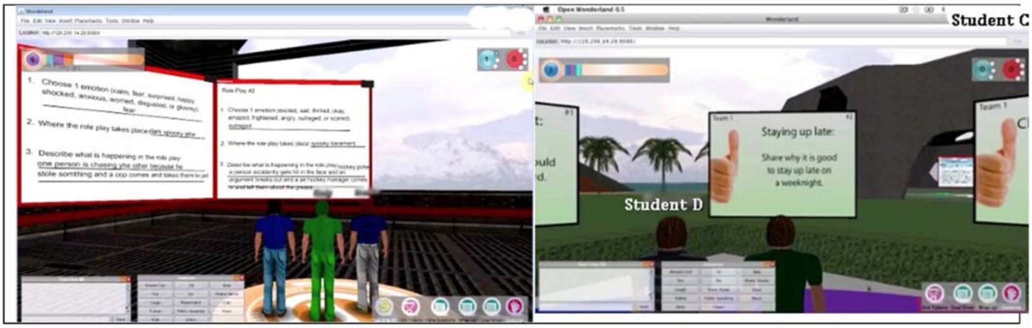

Figure 4: Example screenshot of students in front of a shared board in iSocial. [Colour figure can be viewed at wileyonlinelibrary.com]

In Figure 4, these two NP activity screenshots came from observations in cluster 3—the cluster that has the highest frequency of verbal interactions with only avatar proximity. Although it is good practice to put information inside the game-based environment for better accessing, this text-based, and question answering type of design undermined the opportunities for autistic youth to face their avatars toward the person they were interacting with. However, if they were in an open, less structured area (Figure 2), and the game-based task requires more exploratory effort collaboratively, they were more likely to orient avatars to face each other like they would do in face-to-face interaction.

Another characteristic of the game activities in cluster 2—the cluster that has the highest frequency of verbal with rich nonverbal signals, was that the students needed to both explore and interact with the game environment. For example, in the restaurant buffet activity, students planned their own restaurant buffet by interacting with the virtual restaurant and collaboratively "creating" these virtual items. In the King's Quest activity, students "collected" the virtual objects in the King's castle while discussing clues for finding the next item in order to meet the "King." However, game activities in cluster 1—the cluster that has the lowest frequency of verbal with nonverbal signals, are more traditional and structured activities that transform classroom-like tasks into the 3D virtual environment. Examples of cluster 1 activities are rating the video and sharing ideas based on a board with texts and instruction.

We need to re-conceptualize ideas around designing collaborative learning in the 3D game-based CVLE, in particular away from more traditional approaches and towards a notion of learning as more centered upon experience and exploration in order to support autistic youth's collaborative learning in the avatar embodied, game-embedded 3D virtual environment (De Freitas Rebolledo-Mendez, Liarokapis, Magoulas, & Poulovassilis, 2010; Dickey, 2007). In this way, we then could foster social interactions that are rich with verbal and nonverbal behaviors, and eventually lead to better experiences of embodied social presence and learning benefits.

Central to designing and facilitating social learning in 3D game-based CVLE is the idea of learning through experience. A structured undertaking and planning for learning in this way requires imagination and creativity on the part of the practitioner, as well as developing a different set of teaching skills with less emphasis upon curriculum and more upon arranging learning experiences that enriched with embodied social presence, peer social interaction and group work. The lessons from this study for practitioners and educators is the consideration of the characteristics of social interactions in the 3D virtual environment, and to design learning experiences that encourage the desired verbal and nonverbal interaction pattern. The role of the practitioner and learners is clearly being realigned in the light of more social modes and opportunities for learning.

Conclusion and future directions

This study focused on exploring the relationship of autistic youth's social interactions patterns with experience of embodied social presence in the 3D game-based CVLE iSocial. The findings provided insights into the use of game-based activities in 3D virtual environments to support collaborative learning that is based on combining verbal and nonverbal social interactions. Future empirical testing of different game-design elements of 3D collaborative learning is necessary to gain a deeper understanding of how the 3D game-based CVLE could become valuable platforms to transform collaborative learning.

Statements on open data, ethics and conflict of interest

Data from student experiences and about student accomplishment in iSocial were collected based upon approval for all procedures and assurances of anonymity from the University of Missouri-Columbia Institutional Review Board. Access to these data by other researchers is only available through specific requests for amendments to the current approvals and will be considered on a case by case basis.

Data from student experiences and about student accomplishment in iSocial were collected based upon approval for all procedures and assurances of anonymity from the University of Missouri-Columbia Institutional Review Board.

There is no conflict of interest with the current publication of the results of the research of iSocial.

Acknowledgements

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A090197 to the University of Missouri. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

Allmendinger, K. (2010). Social presence in synchronous virtual learning situations: the role of nonverbal signals displayed by avatars. *Educational Psychology Review*, *22*, 41–56. https://doi.org/10.1007/s10648-010-9117-8

Ang, C. S., Bobrowicz, A., Siriaraya, P., Trickey, J. & Winspear, K. (2013). Effects of gesture-based avatar-mediated communication on brainstorming and negotiation tasks among younger users. *Computers in Human Behavior*, *29*, 1204–1211. https://doi.org/10.1016/j.chb.2012.10.013

Aymerich-Franch, L. (2010). Presence and emotions in playing a group game in a virtual environment: the influence of body participation. *Cyberpsychology, Behavior, and Social Networking*, *13*, 649–654.

Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N. & Jin, M. (2008). The use of immersive virtual reality in the learning sciences: digital transformations of teachers, students, and social context. *Journal of the Learning Sciences*, *17*, 102–141. https://doi.org/10.1080/10508400701793141

Barab, S., Thomas, M., Dodge, T., Carteaux, R. & Tuzun, H. (2005). Making learning fun: quest Atlantis, a game without guns. *Educational Technology Research and Development*, *53*, 86–107.

Bauminger, N. (2002). The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: intervention outcomes. *Journal of Autism and Developmental Disorders*, *32*, 283–298.

Berns, A., Gonzalez-Pardo, A. & Camacho, D. (2013). Game-like language learning in 3-D virtual environments. *Computers & Education*, *60*, 210–220. https://doi.org/10.1016/j.compedu.2012.07.001

Biocca, F., Harms, C. & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: review and suggested criteria. *Presence: Teleoperators & Virtual Environments*, *12*, 456–480.

Burgess, M. L., Slate, J. R., Rojas-LeBouef, A. & LaPrairie, K. (2010). Teaching and learning in second life: using the community of inquiry (CoI) model to support online instruction with graduate students in instructional technology. *The Internet and Higher Education*, *13*, 84–88.

Cheng, Y., & Ye, J. (2010). Exploring the Social Competence of Students with Autism Spectrum Conditions in a Collaborative Virtual Learning Environment–The Pilot Study. *Computers & Education*, *54*, 1068–1077.

Cobb, S., Beardon, L., Eastgate, R., Glover, T., Kerr, S., Neale, H., *et al.* (2002). Applied virtual environments to support learning of social interaction skills in users with Asperger's syndrome. *Digital Creativity*, *13*, 11–22. https://doi.org/10.1076/digc.13.1.11.3208

Cole, H. & Griffiths, M. D. (2007). Social interactions in massively multiplayer online role-playing gamers. *CyberPsychology & Behavior*, *10*, 575–583. https://doi.org/10.1089/cpb.2007.9988

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*, 10–32. https://doi.org/10.1111/j.1467-8535.2009.01038.x

De Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G. & Poulovassilis, A. (2010). Learning as immersive experiences: using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world. *British Journal of Educational Technology*, *41*, 69–85. https://doi.org/10.1111/j.1467-8535.2009.01024.x

Dickey, M. D. (2007). Game design and learning: a conjectural analysis of how massively multiple online role-playing games (MMORPGs) foster intrinsic motivation. *Educational Technology Research and Development*, 55, 253–273. https://doi.org/10.1007/s11423-006-9004-7

Dorling, S. R., Davies, T. D., & Pierce, C. E. (1992). Cluster analysis: A technique for estimating the synoptic meteorological controls on air and precipitation chemistry—Results from Eskdalemuir, South Scotland. *Atmospheric Environment. Part A. General Topics*, *26*, 2583–2602.

DuCharme, R., & Gullotta, T. P. (2003). *Asperger Syndrome: A Guide for Professionals and Families*. Springer.

Esteves, M., Fonseca, B., Morgado, L. & Martins, P. (2011). Improving teaching and learning of computer programming through the use of the second life virtual world. *British Journal of Educational Technology*, *42*, 624–637. https://doi.org/10.1111/j.1467-8535.2010.01056.x

Fabri, M., Moore, D. J. (2005). The use of emotionally expressive avatars in collaborative virtual environments. In Proceedings of Symposium on Empathic Interaction with Synthetic Characters held at Artificial Intelligence and Social Behaviour Convention 2005 (AISB 2005), April 2005.

Freedman, B., & Silverman, W. (2008). Improving Social Skills for Children with High Functioning Autism. *Exceptional Parent*, *38*(5), 64–66.

Freitas, S. D. & Neumann, T. (2009). The use of 'exploratory learning' for supporting immersive learning in virtual environments. *Computers & Education*, *52*, 343–352. https://doi.org/10.1016/j.compedu.2008.09.010

Friedman, D., Steed, A. & Slater, M. (2007). Spatial Social Behavior in Second Life. In C. Pelachaud, J.-C. Martin, E. André, G. Chollet, K. Karpouzis, & D. Pelé (Eds.), *Intelligent virtual agents* (Vol. 4722, pp. 252–263). Berlin/Heidelberg: Springer. Retrieved from http://www.springerlink.com/content/g4175v2h4m3t0617/abstract/

Hair Jr., J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R. L. (1998). *Multivariate data analysis with readings*. Englewood Cliffs, NJ: Prentice-Hall.

Hamalainen, R. (2008). Designing and evaluating collaboration in a virtual game environment for vocational learning. *Computers & Education*, *50*, 98–109. https://doi.org/10.1016/j.compedu.2006.04.001

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507.

Laffey, J. M., Stichter, J. & Galyen, K. (2014). Distance learning for students with special needs through 3D virtual learning. *International Journal of Virtual and Personal Learning Environments*, *5*, 15–27. https://doi.org/10.4018/ijvple.2014040102

Le Couteur, A., Lord, C., & Rutter, M. (2003). *Autism Diagnostic Interview-Revised (ADI-R)*. Los Angeles: Western Psychological Services.

McGee, G. G., Almeida, M. C., Sulzer-Azaroff, B., & Feldman, R. S. (1992). Promoting reciprocal interactions via peer incidental teaching. *Journal of Applied Behavior Analysis*, *25*, 117–126. https://doi.org/10.1901/jaba.1992.25-117

Mennecke, B. E., Triplett, J. L., Hassall, L. M., Conde, Z. J. & Heer, R. (2011). An examination of a theory of embodied social presence in virtual worlds. *Decision Sciences*, *42*, 413–450. https://doi.org/10.1111/j.1540-5915.2011.00317.x

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W. & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: a meta-analysis. *Computers & Education*, *70*, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, *50*, 159–179.

Montoya, M. M., Massey, A. P. & Lockwood, N. S. (2011). 3D Collaborative virtual environments: exploring the link between collaborative behaviors and team performance. *Decision Sciences*, *42*, 451–476.

Odom, S. L., & Strain, P. S. (1986). A Comparison of Peer-Initiation and Teacher-Antecedent Interventions for Promoting Reciprocal Social Interaction of Autistic Preschoolers. *Journal of Applied Behavior Analysis*, *19*.

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.

Parsons, S., Mitchell, P., & Leonard, A. (2005). Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments? *Autism*, *9*, 95–117. https://doi.org/10.1177/1362361305049032

Riby, D., & Hancock, P. J. B. (2009). Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research*, *53*, 169–181.

Rutten, A., Cobb, S., Neale, H., Kerr, S., Leonard, A., Parsons, S., & Mitchell, P. (2003). The AS interactive project: single-user and collaborative virtual environments for people with high-functioning autistic spectrum disorders. *The Journal of Visualization and Computer Animation*, *14*, 233–241.

Schmidt, M., Laffey, J. M., Schmidt, C. T., Wang, X. & Stichter, J. (2012). Developing methods for understanding social behavior in a 3D virtual learning environment. *Computers in Human Behavior*, *28*, 405–413. https://doi.org/10.1016/j.chb.2011.10.011

Schroeder, R., Heldal, I., & Tromp, J. (2006). The usability of collaborative virtual environments and methods for the analysis of interaction. *Presence: Teleoperators & Virtual Environments*, *15*, 655–667.

Sigman M, Ruskin E. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. *Monographs of the Society for Research in Child Development*, *64*(1, Serial No. 256):v–114.

Standen, P. J. & Brown, D. J. (2006). Virtual reality and its role in removing the barriers that turn cognitive impairments into intellectual disability. *Virtual Reality*, *10*, 241–252.

Stichter, J. P., Laffey, J., Galyen, K. & Herzog, M. (2014). iSocial: delivering the social competence intervention for adolescents (SCI-A) in a 3D virtual learning environment for youth with high functioning autism. *Journal of Autism and Developmental Disorders*, *44*, 417–430. https://doi.org/10.1007/s10803-013-1881-0

Stichter, J. P., Randolph, J., Gage, N. & Schmidt, C. (2007). A review of recommended social competency programs for students with autism spectrum disorders. *Exceptionality: A Special Education Journal*, *15*, 219. https://doi.org/10.1080/09362830701655758

Stichter, J. P., Herzog, M. J., Visovsky, K., Schmidt, C., Randolph, J., Schultz, T., *et al*. (2010). Social competence intervention for youth with Asperger syndrome and high-functioning autism: an initial investigation. *Journal of Autism and Developmental Disorders*, *40*, 1067–1079.

Tromp, J. G., Steed, A. & Wilson, J. R. (2003). Systematic usability evaluation and design issues for collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *12*, 241–267. https://doi.org/10.1162/105474603765879512

Wallace, S., Parsons, S., Westbury, A., White, K., White, K., & Bailey, A. (2010). Sense of presence and atypical social judgments in immersive virtual environments Responses of adolescents with Autism Spectrum Disorders. *Autism*, *14*(3), 199–213.

Wang, X., Laffey, J., Xing, W., Galyen, K. & Stichter, J. (2017). Fostering verbal and non-verbal social interactions in a 3D collaborative virtual learning environment: a case study of youth with autism spectrum disorders learning social competence in iSocial. *Educational Technology Research and Development*, 1–25. https://doi.org/10.1007/s11423-017-9512-7

Wang, X., Laffey, J., Xing, W., Ma, Y. & Stichter, J. (2016). Exploring embodied social presence of youth with Autism in 3D collaborative virtual learning environment: A case study. *Computers in Human Behavior*, *55*, 310–321. https://doi.org/10.1016/j.chb.2015.09.006

Xing, W., Guo, R., Petakovic, E. & Goggins, S. (2015). Participation-based student final performance prediction model through interpretable genetic programming: integrating learning analytics, educational data mining and theory. *Computers in Human Behavior*, *47*, 168–181.
Yee, N., Bailenson, J. N., Urbanek, M., Chang, F. & Merget, D. (2007). The unbearable likeness of being digital: the persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior*, *10*, 115–121.

Appendix

Embodied presence, co-presence and social presence observation guide

ESP Level	*Sense within each level*	*Examining questions*	*Evidence*
Embodied presence	Sense of one's own avatar	Did the student regard his/her avatar as the representation of himself/herself in the environment?	Positive if always refer one's own avatar as "my avatar" or "I" Negative if refer to one's own avatar as this avatar or this character
	Sense of virtual object and space	Did the student aware of the spatial context and objects in the context?	Positive if one interacts with objects and indicate knowing the context of the environment Negative if no interaction between a student and the environment
Embodied copresence	Sense of other avatars	Did the student exhibit awareness of the existence of other avatars around?	Positive if one interacts with others via verbal or non-verbal communication Negative if one showed no interaction with others
	Sense of proximity of other avatars	Did the student have spatial proximity when using his/her own avatar in the environment?	Positive if one keeps his/her avatar an appropriate distance from others Negative if one shows no awareness of avatar proximity
Embodied social presence	Sense of one's own avatar	Did the student regard his/her avatar as the representation of himself/herself in the environment?	Positive if refer one's own avatar as my avatar or I Negative if refer to one's own avatar as this avatar or this character
	Sense of other user's avatar	Did the student have awareness of the existence of other avatar around them?	Positive if one interacts with others via verbal or non-verbal communication Negative if one showed no interaction with others
	Sense of collaborative engagement	Did the student's avatar engage in a goal-oriented collaborative/joint activity with the other social actors' avatar?	Positive if the student collaborates with others using verbal and non-verbal communication; interact with object in the environment during a goal-oriented joint activity

Appendix: Continued

ESP Level	*Sense within each level*	*Examining questions*	*Evidence*
			Negative if the student did not collaborate with the others or interact the object in the environment during a goal-oriented joint activity
	Sense and appraisal of others	Did the student show appraisal or reflection on the idea that "real person" is behind other avatar?	Positive if the student comment on or evaluate the opinions of others during collaboration Negative if the student shows no response to others' opinions and questions
	Sense of one's own actions as manifest in avatar-based social interaction	Did the student show appraisal or reflection on the idea or opinions of himself when embodied in one's own avatar?	Positive if refer one's own avatar actions as my avatar did... or I did... or I agree, or I think... during collaboration Negative if refer to one's own avatar actions as this avatar did... or this character did... or did not refer to oneself at all during collaboration