Recommending Video Games to Adults with Autism Spectrum Disorder for Social-Skill Enhancement

Alisha Banskota
baskotaalisha@gmail.com
Computer Science Dept., Brigham Young University
Provo, Utah, 84602, USA

Yiu-Kai Ng
ng@compsci.byu.edu
Computer Science Dept., Brigham Young University
Provo, Utah, 84602, USA

ABSTRACT

Autism spectrum disorder (ASD) is a long-standing mental condition characterized by hindered mental growth and development and is a lifelong disability for the majority of affected individuals. In 2018, 2-3% of children in the USA have been diagnosed with autism. As these children move to adulthood, they have difficulty in developing a well-functioning motor skill. Some of these abnormalities, however, can be gradually improved if they are treated appropriately during their adulthood. Studies have shown that people with ASD enjoy playing video games. Educational games, however, have been primarily developed for children with ASD, which are too primitive for adults with ASD. We have developed a gaming and personalized recommender system that suggests therapeutic games to adults with ASD which can improve their social-interactive skills. The gaming system maintains the entertainment value of the games to ensure that adults are interested in playing them, whereas the recommender system suggests appropriate games for adults with ASD to play. The effectiveness and merit of our gaming and recommender system is backed up by an empirical study.

CCS CONCEPTS

• **Information systems → Information retrieval**; • **Retrieval tasks and goals** → *Recommender systems.*

KEYWORDS

Autism, Adults, Game Development, Recommender System

ACM Reference Format:
Alisha Banskota and Yiu-Kai Ng. 2020. Recommending Video Games to Adults with Autism Spectrum Disorder for Social-Skill Enhancement. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20), July 14–17, 2020, Genoa, Italy.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3340631.3394867

1 INTRODUCTION

Autism Spectrum Disorder (ASD) is a developmental disorder that affects social, behavioral, and communication skills of an individual. People with ASD face different challenges that could lie in cognitive area [24], in addition to developing appropriate relationship with their peer at different deficiency levels [19]. Despite the limitation in social activity, individuals with ASD are highly interested in video games as shown in different studies [14, 15, 18]. Children and adolescents with ASD spend more than two hours a day playing video games which is more than by typically developing children or children with other disabilities [17]. Not just for children, adults with ASD have also reported to be engaged in video games. These video games have positive impact on them as the adults find video games to be entertaining, stress relieving, and creative [17].

Regular video games, however, are not designed especially for adults with autism in mind, even though games have been previously developed for children with ASD to enhance their social skills. There are specific games [9] developed that are both fun and acting as a therapeutic agent for children with ASD. These systems, however, are not suitable for adults with ASD as these games are too basic for adults and they do not enjoy or learn from them. Also, adults with ASD face different challenges than children with ASD which means that there is a demand to design therapeutic games especially for adults with ASD. Unfortunately, very little has been explored in this area. In this paper, we propose a gaming and recommender system for adults with ASD to enhance their social-interactive skills in order to connect with others.

Developing well-designed therapeutic games for adults with ASD is essential, since they can provide entertainment value to players and better engage adults with ASD to learn social skills. If players are not attracted to a game, they will not spend time playing the game. Games that can be both entertaining and educational give rise to the concept of 'edutainment' [23]. Such edutainment games first started appearing in the early 80's [7] and today they are estimated to be an $8.4 billion market in the USA [12].

We have developed a gaming and recommender system with 'edutainment' values. Among a myriad of options available as base to develop our games, we have chosen Minecraft[1], since people with ASD enjoy Minecraft and it is quick and easy to develop them.

We have created therapeutic games for adults with ASD and personalized the most appropriate recommendations to the players based on their areas of weakness. In order for our recommender system to work, we relied on a finite set of areas of weakness established by ScenicView Academy[2], a school in Provo, Utah for adults with ASD. ScenicView Academy assists their adult students, who are between 18 and 30 years of age, in maintaining social relationships. Based on the various principles of deficiencies developed by ScenicView, we have derived the most prominent challenges faced by adults with ASD which comprises of (i) *developing audio communication skills*, (ii) *recognizing facial expressions*, (iii) *maintaining*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
UMAP '20, July 14–17, 2020, Genoa, Italy
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-6861-2/20/07…$15.00
https://doi.org/10.1145/3340631.3394867

[1]https://www.minecraft.net/en-us/
[2]https://www.svacademy.org/

eye contact, (iv) *showing empathy*, and (v) *engaging in speech therapy*. We start with studying the weakness areas of autistic adults based on their *gaming profiles* and a *questionnaire* containing questions that address to different weakness areas. Hereafter, we recommend appropriate games to them to work on. Conducted empirical studies reflect the changes in social behavior observed after the participants have played the recommended therapeutic games.

2 RELATED WORK

A large number of gaming systems have been developed in the past, since 97% of teens between the age of 12-17 play some kind of video games [5]. There are games developed specifically for children with ASD, which include Go Go Games [9] to teach them "multiple cue responding", Autcraft [25], which is a Minecraft server, and the Flexibility Learning Project (FLOW) [22], which works on rigidity of autistic children. These games, however, are designed solely for children with ASD and too primitive for adults with ASD.

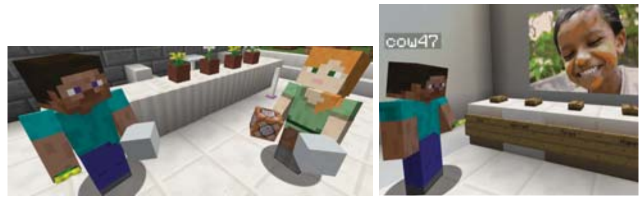
To enhance the social skills of autistic adults, both gaming and non-gaming approaches have been used in the past. Social skills and reciprocal social behavior show improvement with age, even though to a much more modest degree than overall verbal and speaking skills [1, 10, 28, 29]. Neural imaging shows that adults with ASD are more interested in social interaction than children with ASD, and show greater brain activity in social tasks [4]. As they enter adolescence, those with ASD tend to become less withdrawn and self-focused socially. By the time they reach adulthood, many of them will have been able to establish friendships and (in some cases) even intimate relationships such as marriage, though this is more common among "high-functioning" individuals [2, 6, 10, 21]. However, even in high-functioning cases, non-verbal skills and overall social fluency deficits often persist into adulthood, which include abnormal vocalics, proxemics, chronemics, haptics, and over-literalness. Non-verbal skills remain a primary challenge regardless of cognitive ability or functional level [1, 21, 27, 29].

There are a few gaming systems designed especially for adult's cognitive disabilities. The authors of [33] design two games for people with cognitive disabilities and discuss the potential improvements that could be made for adults with cognitive disabilities. The authors of [20] recommend social-interactive games to adults with ASD, suggesting games that are familiar to the user and unfamiliar to them using features of video games, such as complexity of games. However, games recommended by [20] are neither personalized nor consider 'edutainment' at all. They separately show that their approaches perform better than other existing approaches.

The authors of [13] argue that games with 'edutainment' purpose have failed, since they are incapable of increasing the endogenous value of a game without compromising the fun aspect. We have balanced the development of games that have both endogenous value of enhancing the skills lacked by autistic adults and entertainment values to keep them engaged.

3 OUR GAMING SYSTEM

Games developed by us have entertaining objectives where social expertise is a natural part of succeeding in the game. This approach builds on the *endogenous value* already inherent in Minecraft, including value that comes from ongoing discovery of information, collecting increasing types and levels of materials, and achieving goals of increasing complexity along the way to a difficult end goal. We build on all of the principles that make Minecraft one of the most successful video games ever, while expanding this world to target the five specific *weakness areas* (listed below) based on which different therapeutic games in Minecraft are designed and implemented.



(a) Quiz game for developing audio communication skills

(b) Jim's World for recognizing facial expression

Figure 1: Minecraft Video games for skill development

Developing Audio Communication Skills. Adults with ASD face challenges in maintaining communication, listening to others, or understanding others' perspectives. As a result, they have difficulty in forming relationships, expressing themselves to others, etc. Our *communication room* games enforce people who are not very effective in communicating to find new people, initiate a conversation, learn to keep a conversation going, and make friends. Figure 1(a) shows a snapshot of the quiz-and-answering game.

Recognizing Facial Expression. Studies have shown that in typical development, even an infant can distinguish facial expressions. However, for adults with ASD, understanding the situation or what the other person is saying through their expressions can be confusing. This kind of games is a good exercise for the users to learn basic expressions, such as happy, sad, angry, surprised, shocked, and tired. Figure 1(b) shows the screenshot of a facial recognition game.

Engaging in Speech Therapy. Adults with ASD can have problems paying attention to sounds of others, decipher them, or produce sounds themselves. These are some challenges that are unique to people with ASD. Challenges in speech again impedes them from expressing themselves or understanding others.

Showing Empathy. Some adults with ASD have problems showing empathy or expressing how they feel in a sensitive situation which is a major underlying cause behind deficiencies in social-interaction or relationship building. To address this issue, we have developed games where players communicate with a fictional character who interacts with the player and shares his thoughts.

Maintaining Eye Contact. Maintaining eye contact is a non-verbal communication behavior that is essential to know if the other person is interested in the conversation or to pick up important cues in the conversation. For some adults with ASD, maintaining eye contact during a conversation makes the situation even more stressful and they become more distracted during the conversation.

4 OUR RECOMMENDER SYSTEM

Our recommender system suggests games to users based on their weakness areas so that they can enhance their social-interactive skills in the respective weakness areas. Since it is difficult to obtain user's clinical data as it is confidential, we infer the weakness areas faced by adults with ASD through the VideoGameGeek[3], denoted

[3] https://videogamegeek.com/

Table 1: Some of the questions in questionnaire with their weakness areas: Developing Audio Communication Skills, Maintaining Eye Contact, Recognizing Facial Expression, Showing Empathy, and Engaging in Speech Therapy

Questions	Areas
I am often the last to understand the point of a joke	ST
In a social group, I can easily keep track of several different people's conversations	AC
I am quick to comfort someone in distress	SE
When I'm reading a story, I can easily imagine what the characters might look like	FE
I become overwhelmed when looking into others' eyes	EC

VideoGG, games they have played previously. VideoGG archives thousands of games and provides different metadata such as the themes of each game, number of players in that game, description of the game, etc. These games are *not* therapeutic for adults with ASD, since they are only developed for entertainment or educational purpose, but not both, whereas our therapeutic games developed in Minecraft do. We use *metadata* about games from VideoGG to rank our Minecraft games for recommendation. We begin the recommendation process by using a *questionnaire*[4] and analyze the answers to its questions for game recommendation.

4.1 Questionnaire-Based Recommendation

To recommend the most appropriate therapeutic game developed by us to each user individually, we create a questionnaire and ask the user to fill out the questionnaire so that we can study (the changes in terms of) the deficiency areas of the user. Questions in the questionnaire co-relate to each of the *five* weakness areas and the answers given to these questions indicate the severity of deficiency of the user in those areas. The users choose if they 'Strongly Disagree', 'Disagree', 'Slightly Disagree', 'Neither Agree Nor Disagree', 'Slightly Agree', 'Agree', and 'Strongly Agree' to each question. A subset of the 34 questions in the questionnaire with their corresponding weakness areas is shown in Table 1.

Prior to recommending any therapeutic games to an adult with ASD, we require the adult to answer the questions in the questionnaire. Since the questions are associated with the five weakness areas, after the adult has filled out the questionnaire, we take the *average score* of the answers (between 1 to 7 inclusively) to the questions belonged to each weakness area. The average score computed for each weakness area yields the *baseline score*. We set **5.5** as a *threshold* so that a value 5.5 (close to *Agree*) or higher indicates positivity, i.e., absence of weakness, and below 5.5 implies negativity. The threshold value 5.5 is used to determine whether to continue recommending games in a particular weakness area to be played by an adult or draw the conclusion that the adult has no deficiency in the corresponding weakness area (anymore).

The *baseline scores* of all the weakness areas computed for a user before playing any therapeutic games is called the *initial baseline scores* of the user. If the initial baseline score of a weakness area indicates the presence of a particular weakness, we recommend a therapeutic game in that area one at a time. However, since there is

[4]The questionnaire includes questions gleaned from psychometrically evaluated tools, such as the Autism Quotient [30].

abstract, adult, agriculture, aliens, amusement park, animals, anime, art, board game, business, card game, cartoon, children, comedy, comics, cooking, crime, cyberpunk, dancing, detective, driving, Egyptian, fairy tale, fantasy, fishing, flight, folklore, history, horror, hunting, martial arts, medical, medieval, military, minigame compilation, mining, monster collecting, moral choices, morbid, movie and book and tv show, music, mythology, nature, ninja, occult, ocean, other, physics, pirates, politics, public transport, puzzle, relationship, robot, samurai, school, science and technology, science fiction, space, sports, spy, steampunk, superhero, surreal, talking animals, Viking, war, western, words, zombies

Figure 2: Themes of games extracted from VideoGG

a number of therapeutic games that can be recommended in each of these areas, we develop a *ranking strategy* of those games to decide the sequence in which we recommend them to the user.

4.2 Ranking Therapeutic Games

We have developed a number of therapeutic games in each of the weakness areas and rely on the top-10 VideoGG games that a user has enjoyed playing in the past specified in the user's profile to rank therapeutic games using *attribute values* of the VideoGG games.

We determine the closeness of the attribute values between the VideoGG games in a user profile with the attribute values of the therapeutic games in a weakness area to rank the latter in order to generate a sequence of recommended games for the area. The therapeutic game, whose attribute values that are *closest* to the attribute values of the VideoGG games specified in a user profile, is ranked the *highest* followed by the second closest game and so on. Attributes of VideoGG games considered by our recommender are (i) *theme*, (ii) *topic*, (iii) *user rating*, and (iv) *number of players*. Values of these attributes are made available through VideoGG, which are either provided by VideoGG or the end users. They are valuable for determining the closeness of various video games.

4.2.1 Themes. A theme of a VideoGG game is the *category* to which the game belongs. There are 70 unique themes among all the games in VideoGG (shown in Figure 2) and each game can have one or more themes associated to it. We manually assign one or more of these themes to each therapeutic game as well. In order to rank the therapeutic games to be recommended based on their themes with respect to the themes of VideoGG games specified in a user's profile, we create a *co-relation matrix* of *themes* to determine their *frequencies of co-occurrence*. A therapeutic game whose themes have *highest co-relation values* with the themes of the VideoGG games in a user's profile affects its ranking among other therapeutic games.

To compute a co-relation matrix of themes, we first create a matrix of size 70×70 as there are 70 distinct themes in a collection of 100,000 VideoGG games extracted from VideoGG and initialize each entry as zero. For each theme of a VideoGG game, we create an index to be used in the matrix. The matrix is then filled with the number of times each theme is co-occurrent with one another in the same game. In other words, it keeps track of how many times *two* themes appear in the same game, which is computed below.

$$CF(V,W) = \frac{P_{V,W}}{P_{V,V} + P_{W,W} - P_{V,W}} \tag{1}$$

where $CF(V,W)$ is the ratio of *co-occurrence frequency* between themes V and W among all the VideoGG games, $P_{V,W}$ is the *frequency of co-occurrence* of V and W together in the VideoGG games, $P_{V,V}$ and $P_{W,W}$ are the number of times V and W appear individually among all the games in VideoGG, respectively.

Having calculated the ratio of the CF value between two themes, we can compute the *category similarity score* among each therapeutic and the VideoGG games in a user profile as follows:

$$CTS(TG) = \sum_{j \in VG} \frac{\sum_{(V,W) \in S(j,TG)} CF(V,W)}{|TG| \times |VG_j|}, S(j,TG) = \sum_{i \in VG_j} \sum_{k \in TG} (i,k) \quad (2)$$

where $CTS(TG)$ is the *theme similarity score* of a particular therapeutic game, denoted TG, $|TG|$ is the number of themes in TG, VG stands for the set of VideoGG games specified in a user profile, $|VG_j|$ is the number of themes in the j^{th} VideoGG game, denoted VG_j, in the user profile, $S(j, TG)$ is the set of any two themes created using the cross product of themes in the j^{th} VideoGG game in the user profile and TG, and $CF(V, W)$ is as defined in Equation 1.

4.2.2 Topic Analysis. There is a *textual description* for each game in VideoGG explaining what the game is, and we have created our own description for each therapeutic game as well. Different *topics* can be assigned to these descriptions, and determining games with the *same* topics is a feature adapted by us in *ranking* our therapeutic games. The therapeutic games whose descriptions are *closely related* to that of the VideoGG games in a user profile are ranked *higher*. We train a *Latent Dirichlet Allocation (LDA)* model to generate a list of unique topics. Hereafter, we assign the topics that match our therapeutic games with the VideoGG games played by the user to develop a ranking based on *topics*.

The LDA model consists of two sets of metrices. The first matrix determines the probability of selecting a *word* given a *topic* and the second one finds out the probability of selecting a particular *topic* given a *document*. Based on these metrices, the model determines the probability for each topic given a document. During the training, LDA determines the probability of a word w given a topic z, i.e., $P(w|z)$, and the probability of a topic z given a document d, i.e., $P(z|d)$. We apply Gibb's sampling [8] to estimate $P(w_i|z_j)$ and $P(z_j|d_k)$ by iterating over each word w_i in each document d_k and assigns a cluster for w_i based on its probability.

$$P(z_i = j|w_i, d_k, z_{-i}) \propto \frac{C_{WZ}(w_i, j) + \beta}{\sum_w C_{WZ}(w_i, j) + W\beta} \times \frac{C_{DZ}(d_k, j) + \alpha}{\sum_z C_{DZ}(d_k, z) + Z\alpha} \quad (3)$$

where $P(z_i = j|w_i, d_k, z_{-i})$ is the probability in which *topic* z_i is assigned to word w_i and z_{-i} denotes all topic-and-word and document-and-topic assignments excluding the current assignment z_i for w_i, the first multiplicative factor in the equation computes $P(w_i|z_i = j)$, i.e., the probability of a word given a topic, and the second factor calculates $P(z_i = j|d_k)$, i.e., the probability of a topic given a document.

The Training Process. To train the LDA using game descriptions, we extract the descriptions of around 15,000 games from VideoGG and preprocess the description of each game. We first remove the stop words, stem the words, and construct the trained LDA on the pre-processed description of VideoGG games based on the 100 keywords used to distinguish each topic along with their weights. We consider LDA for different number of topics starting from 20 to 50 and manually study the results of LDA for different number of topics. Eventually, we conclude that *twenty-seven* is the optimal number of topics as we were able to associate the descriptions of the therapeutic games to the topics, and the topic assigned to each description captured the important keywords of the description.

arcade 0.0168, version 0.0164, fighter 0.0153, super 0.0124, street 0.0117, characters 0.0110, fighting 0.0075, time 0.0069, turtles 0.0061

Teenage Mutant Ninja Turtles: Turtles in Time is a video game produced by Konami. A sequel to the original Teenage Mutant Ninja Turtles (TMNT) arcade game, it is a scrolling beat 'em up based on the 1987 TMNT animated series. Originally an arcade game, Turtles in Time was ported to the Super Nintendo ... Years later, the arcade version of Turtles ...

Figure 3: Topic assigned to the game "Teenage Mutant Ninja Turtles: Turtles in Time" and its description with the matching keywords highlighted

4X Strategy, Action, Action Adventure, Action RPG, Adventure, Arcade, Augmented Reality, Beat 'em up, Classic Games, Clicker, Dating sim, Dungeon Crawler, Educational, Endless runner, Fighting, First person shooter, Fitness, Flight simulator, Hidden object, Interactive movie, Life simulation, Light gun shooter, Mana gement, Maze, MMO, MOBA, Other, Party, Pinball, Platform, Point-and-click, Puzzle, Racing, Real time strategy, Rhythm, Rouguelike, RPG, Run-n-gun, Sandbox, Scrolling, Shoot 'em up, Shooter, Simulation, Sports, Stealth, Strategy, Text adventure, Tower defense, Trivia, Visual novel, Walking simulator

Figure 4: Genres in VideoGG games

Topic Modeling on Game Descriptions. After we have trained an LDA for topic analysis of our recommender system, we can determine which topic is to be assigned to a game using Equation 4, which is given as an input a game description D. Based on the distribution of each word w_i in D and the probability of occurrence of each word w_i in each topic z_j, i.e., $P(w_i|z_j)$, Equation 4 determines the topic with the*highest* probability for the game description.

$$Topic(D_G) = max_{j=1}^{N} \sum_{i=1}^{K} P(w_i|z_j) \quad (4)$$

where $Topic(D_G)$ is the topic assigned to the description D of game G, N is the total number of latent topics, i.e., 27 in our case, and K is the total number of words in D_G.

Figure 3 shows the topic assigned to the VideoGG game "Teenage Mutant Ninja Turtles: Turtles in Time" along with its description.

4.2.3 Predicted Ratings of Therapeutic Games. VideoGG posts *average user rating* of each game. Based on these ratings, we can predict the user rating for each therapeutic game. The predicted rating is then used to partially rank the therapeutic games.

After extracting the *average ratings* of the ten games in VideoGG provided by an adult with ASD, we apply our *rating prediction algorithm* to rate therapeutic games to be recommended to the adult based on those ratings. Our rating prediction algorithm relies on a *weighing* mechanism to decide the *closeness* between a therapeutic game and each of the ten VideoGG games based on their *genres*. VideoGG comes with 51 *unique* genres. These genres describe the nature of the games and are shown in Figure 4.

Ratings of VideoGG. To predict user ratings, we extracted 45,000 games from VideoGG with their *average user ratings*. Players of a VideoGG game G rank G from 1 to 10, with 10 being the highest, and an average is computed to obtain a single rating for G.

The Genre-Based Rating Prediction approach. Having obtained the *genres* of the 45,000 games from VideoGG, we use the *collaborative filtering* technique to *rate* each therapeutic game based on the common genres shared by the VideoGG games with respect to the therapeutic game and the *average ratings* of the VideoGG games.

We first compute the *weight* between a VideoGG game and a therapeutic game as the fraction of the number of *common genres* over the total number of genres of the two games. The genres common to a VideoGG game and a therapeutic game indicates the *closeness* of the games. We obtain the predicted therapeutic *game rating* using the *weight* between a VideoGG game and a therapeutic game and the average rating for the VideoGG game as follows:

$$r_{u,k} = \frac{\sum_{j \in N_u} W_{j,k} \times r_j}{\sum_{j \in N_u} W_{j,k}} \tag{5}$$

where $r_{u,k}$ is the *predicted rating* of the therapeutic game k for user u, $w_{j,k}$ is the *weight* of the therapeutic game k and one of the VideoGG games j specified in u's profile based on their genres, r_j is the *average rating* for the VideoGG game j, N_u is the number of VideoGG games specified in u's profile, which is 10.

After the *rating* for each of the therapeutic games is predicted for a particular user, we can *partially rank* those games to be recommended to the user.

4.2.4 Number of players. The number of players of a game is a *range of potential players* who can play the game. Some of the VideoGG and therapeutic games involve multiple players while others are just single-player games. We compare the recommended number of players in each VideoGG game specified in a user profile with that of a therapeutic game to rank the latter as follows:

$$NPlayers(TG) = \begin{cases} \sum_{n_{TG}} P_{TG} + \sum_{n_{VG}} -1 \times P_{VG} & \text{if } n_{TG} = 1 \\ -(\sum_{n_{TG}} P_{TG} + \sum_{n_{VG}} -1 \times P_{VG}) & \text{if } n_{TG} > 1 \end{cases} \tag{6}$$

where $NPlayers(TG)$ is the score based on the potential number of players for the therapeutic game TG and the one specified in each of the 10 VideoGG game included in a user profile, n_{TG} (n_{VG}, respectively) is the number of player options in TG (in VideoGG game VG, respectively), P_{TG} (P_{VG}, respectively) is (one of) the possible number(s) of players indicated in TG (VG, respectively).

Based on the score computed for each therapeutic game using Equation 6, we rank the therapeutic games in *decreasing order* such that the therapeutic game that receives the highest score with respect to the VideoGG games in a user profile is ranked the highest.

4.2.5 Final Ranking of Therapeutic Games. After computing the partial rankings for each therapeutic game based on each one of its four different attribute values (shown in Sections 4.2.1 through 4.2.4), we combine them to obtain a *single ranking* for the therapeutic game. The single ranking of a therapeutic game determines the cumulative effect of the four attributes, and we rely on the CombMNZ model to obtain the final ranking of the game. CombMNZ is a commonly-used fusion method [3] that combines multiple ranking lists on an item I to determine a *joint ranking* of I.

$$CombMNZ_I = \sum_{K=1}^{N} I^K \times |I^K > 0|, \quad I^K = \frac{S^I - I^K_{min}}{I^K_{max} - I^K_{min}} \tag{7}$$

where N is the number of ranked lists to be fused, which is *four* in our case, I^K is the normalized score of item I in the ranked list K, $|I^K > 0|$ is the number of non-zero, normalized scores of I in the ranked list K, S^I is the score of I in the ranked list K to be normalized, I^K_{max} (I^K_{min}, respectively) is the maximum (minimum, respectively) score available in K.

5 RECOMMENDATION OF GAMES USING ACCUMULATIVE VALUES

After sequencing therapeutic games using VideoGG attribute values, we ask the users to play those games, starting with the top-ranked game, and continue until certain criterion is satisfied in each weakness area prevalent in them.

5.1 Baseline Scores

Recall that before playing any of the therapeutic games, each user is asked to fill up the questionnaire, which determines the *initial baseline scores* of the user. If the initial baseline score corresponding to a weakness area indicates deficiency in the weakness area, the user is recommended to play the top-ranked therapeutic game in the area. After the user plays a game belonged to a particular weakness area, (s)he is asked to answer the questionnaire again and the *initial baseline score* of that weakness area is modified to become the *updated baseline score* for that area. The updated baseline score shows the impact of playing the game with respect to the user. If the updated baseline score for an area *increases* compared with the initial baseline score, we claim that the corresponding game has *affinity* to that user. If the updated baseline score for that area is at least 5.5, which is the threshold, we assert that the user has made significant improvement in that weakness area and the area is *removed* from the list of deficiencies designated for the user. Otherwise, we record the initial baseline score, called *original initial baseline score* (as it will be used later) and assign the *initial baseline score* of that weakness area as *updated baseline score.* Hereafter, we recommend the next therapeutic game from the ranked list in that weakness area to the user. We continue this process for the user until either the threshold of 5.5 for each weakness area is reached or we exhaust all the games in each area.

5.2 Accumulative Values

While a user is playing therapeutic games recommended to them, we maintain a key metric, called *accumulative values*, which capture the change in the baseline scores of each weakness area, i.e., after the user plays a therapeutic game, the accumulative value of that game is set to be the *difference* between the *updated baseline score* and the *initial baseline score* of the weakness area. The accumulative values of a user is a vector of size $1 \times N$, where N is the total number of therapeutic games and each component of the vector stores an accumulative value corresponding to a therapeutic game (in a particular weakness area), which is applicable to all the users. (The layout of therapeutic games in each weakness area is shown below.) For therapeutic games that are *never* played or that do *not* show any change between the two scores, the accumulative value for the game remains *zero.*

Developing Audio Communication Skills	... ...	**Engaging in Speech Therapy**
$G_{1_1}, \ldots, G_{1_{M_1}}$	...	$G_{5_1}, \ldots, G_{5_{M_5}}$

5.3 Affinity Values

After all the weakness areas prevalent in a user have been addressed, i.e., either the user has played each therapeutic game in the weakness area or (s)he has made significant improvement (crossed the threshold) in that area, we assign **affinity values** for the user, which consists of the *latest accumulative value* of each weakness area. After the *affinity values* is assigned to a user, his profile is called an *updated profile* that is used in the 2^{nd} phase for recommendation. The layout of updated profile for a user is shown below, where $BS_1, \ldots, BS_5$ represent the *original initial baseline scores* in all the weakness areas, $G_1A_1, \ldots, G_{10}A_4$ is a matrix of size 10×4 representing the *attribute values* of each of the top-10 VideoGG games preferred by the user, and $AS_1, \ldots, AS_n$ are the *affinity values*, where n is the total number of therapeutic games.

Original Initial Baseline Scores			Attribute Values of VideoGG			Affinity Values		
BS_1	...	BS_5	G_1A_1	...	$G_{10}A_4$	AS_1	...	AS_n

6 AN AUTOMATED STRATEGY

During the 2^{nd} phase of our recommendation process, we develop an automated recommending mechanism in which, given the profile of each new user, we use the data from updated profiles of former users to rank the list of therapeutic games to be recommended to the user. By assigning affinity values to users in Phase 1, we collected the required data for this automated recommendation approach. This recommendation strategy is *fully automated*, since we do not require users to fill up the questionnaire again and again, but only for the first time, to determine the therapeutic games recommended for the user to play. In order to automate the process of recommendation, we determine the *affinity values* for a new user based on the KNN (K-Nearest Neighbor) regression strategy.

6.1 KNN Regression

KNN regression is a supervised learning algorithm [16] which is used to find the output label or class for a new data point in the vector space based on its closest neighbors. We adopt *regression*, since we are supposed to estimate the value of output label for each new user by calculating the mean of output labels of its neighboring data points. (The output label for each data point is further discussed in Section 6.2.) For each new user U, KNN regression calculates the average of the output labels of the K neighbors who are closest to U and assigns the average as the class to U, and K is determined empirically. KNN is preferred over other supervised learning approaches, such as multilayer perceptron [32], Support Vector Machine [11], linear regression [26], etc., since KNN is an instance-based learning algorithm which means it adapts to new addition of data and does not require training to build any model.

6.2 Calculating Affinity Values with KNN

KNN regression assigns output label to a data point, i.e., a new user in our case, by averaging the labels of its neighbors, which are the *affinity values* of the updated profiles that are closest to the new user profile. We consider the *initial baseline scores* of a new user and the *attribute values* of the top-10 VideoGG games that the user prefers as *features*. Using these *features*, if the initial baseline scores of two users are close, and if they have played same types of games in the past, they could be facing challenges in the same weakness areas. We change the representation of the features to numeric and normalize them so that we can calculate the *distance* between them as vectors in the vector space. We find the distance between feature values of updated profiles with that of the new user profile and sort them in *increasing order* using the Euclidean distance measure. We then take the *average* of affinity values for the first K updated profiles and assign it as the *affinity values* for the new user.

6.3 Recommendation Using Affinity Values

When a new user U fills up the questionnaire, we can determine which weakness areas are prevalent in the user according to the *initial baseline scores*. The user is recommended therapeutic games in those areas based on the decreasing order of affinity value of the games assigned to U. There can be a discrepancy between the weakness areas shown by the *initial baseline scores* and the *affinity values* of the user, and there are different cases to consider in recommending games to the user due to this discrepancy.

- Case 1. If the weakness areas shown by both initial baseline scores and the affinity values are the *same*, we simply rely on the affinity value of each game in these weakness areas to recommend games based on the *descending order* of the *affinity values*. However, sometimes two or more therapeutic games in the same weakness area may have the same affinity value. In this case, we break the *tie* to rank these games using the *ranking* of the games as computed in Phase 1.
- Case 2. The initial baseline scores may indicate the presence of *larger* number of weakness areas than the affinity values. For the weakness areas not shown by affinity values, we rely on the ranking of therapeutic games generated by comparing them with the VideoGG games the user likes to play as computed in Phase 1.
- Case 3. The affinity values can indicate *larger* number of weakness areas than that shown by the initial baseline scores. In this case, we do not recommend games in the weakness areas that are not indicated by the initial baseline scores.

7 EXPERIMENTAL RESULTS

In this section, we present the performance evaluation of our gaming and recommender system.

7.1 The Empirical Study

We collaborate with ScenicView Academy, a non-profit school for adults with ASD located in Provo, Utah for conducting our empirical study. Since our gaming and recommender system focuses on the five weakness areas that are vital aspects of lifelong educational skills, we coordinate with ScenicView with the desire to provide the support we can offer through our system.

During the first phase of the study, i.e., development of therapeutic games for adults with ASD, we coordinated with the teachers and therapists at ScenicView who offered their feedbacks on each of the therapeutic games designed for adults with ASD. We completed the development of the games in all five weakness areas in a 2-year period which lasted from April 2017 to March 2019.

Upon developing the Minecraft therapeutic games, we began the testing phase by recommending therapeutic games and making the games accessible to the ScenicView students. The entire study was carried out from April 19, 2019 to December 19, 2019. We first installed the therapeutic games on the computer systems at ScenicView. With the help of teachers at ScenicView, we recruited 78 students who were willing to volunteer for the study. These students either played the games and filled up the questionnaire or simply filled up the questionnaire, which yield the baseline scores.

Table 2: Samples of baseline score pair for the Test and Control group, where BLS stands for Baseline Score

Test Group		Control Group	
Initial BLS	Updated BLS	Initial BLS	Updated BLS
4.6	5.0	5.0	5.0
4.5	4.5	3.0	2.6
3.4	4.2	4.2	5.5

After the training session, the volunteer students started playing the games recommended to them based on the students' initial baseline scores. A month later, the student was asked to fill up the questionnaire again. If the student had not crossed the threshold, i.e., a baseline score < 5.5, in a weakness area, we recommended another game to the student in that area to be played according to our recommender strategy. This process lasts for eight months.

7.2 The Data

The data for our performance evaluation includes the scores of the answers to the questionnaire provided by the volunteers at ScenicView Academy, in addition to the affinity values of therapeutic games played by the volunteers. We collected our data for two different groups. The first group of students, i.e., the *test group*, are volunteers who played the games recommended to them, whereas the second group of students, i.e., the *control group*, did not play any of the therapeutic game but simply filled up the questionnaire at an interval of a 3- to 4-week period. The students were divided into the control and test group in order to study the effectiveness of our gaming and recommendation system by determining whether there were any differences in terms of enhancing their social-interactive skills through playing therapeutic games or not. Out of the 78 volunteers, 65 were in the test group, and the remaining ones in the control group.

Throughout the 8-month period, we observed the change in baseline scores achieved by the students in each weakness area with respect to the previous baseline scores and obtained *pairs* of baseline scores. We collected a total of 352 pairs of baseline scores in all weakness areas for the test group. We followed the same steps for the control group and obtained a total of 50 pairs. These pairs (see samples in Table 2) yield the data to evaluate the performance of our gaming and recommender system using the significant test.

7.3 Performance Evaluation of Our Gaming and Recommender System

We ran the Wilcoxon test on the 352 pairs of baseline scores (and their distribution are shown in Figure 5(a)) achieved by different students and the result of the test shows that the difference is *statistically significant* ($p < 0.02$), which proves that the therapeutic games recommended to them have accomplished the design goal.

In order to further strengthen the claim that the design goal of our gaming and recommender system is achieved, we ran the Wilcoxon test again, but this time it was on the baseline score pairs achieved by the control group students (see the distribution of scores in Figure 5(b)). The p-value for the test is 0.76, which indicates that the difference between the data pairs is *not statistically significant*. This further strengthens our belief that the change of baseline scores for the test group is due to the recommended games, and our gaming and recommender system is effective.

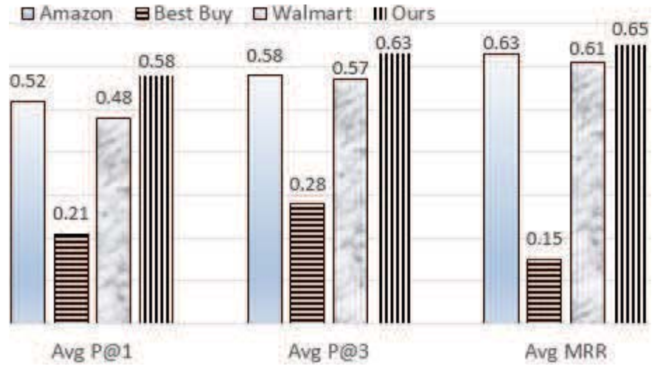


Figure 6: Performance evaluation on the ranking strategy of our recommendation system over others

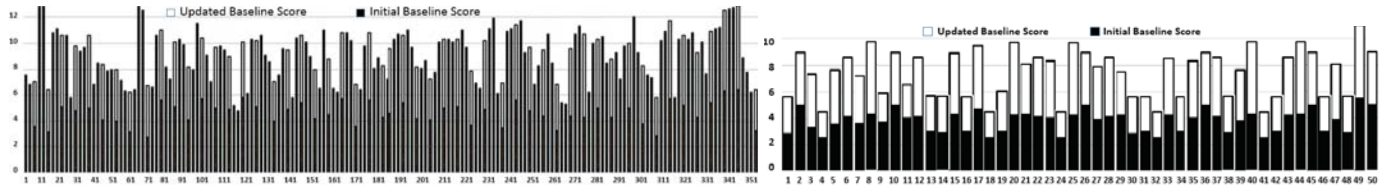
7.4 Performance Evaluation of Our Therapeutic Game Ranking Strategy

Given a user profile and a set of therapeutic games in each particular weakness area, we evaluate the performance of our recommender system in terms of its *ranking accuracy* of therapeutic games and comparing its suggestions with three widely-used video games recommenders, *Amazon*[5], *Best Buy*[6], and *Walmart*[7]. As there are no benchmark datasets that can be used to assess the performance of a video game recommender system, we constructed our own dataset. The dataset consists of over 90,000 VideoGG games, in addition to our therapeutic games, to measure the usefulness of a video game recommended by either Amazon, Best Buy, Walmart, or our recommender that is determined by a number of independent appraisers, who are the 78 volunteers from ScenicView Academy, which serve as the *gold standard* of our evaluation. Based on the decisions on (non-)relevant recommended games made by the individual appraisers, the accuracy of the corresponding ranking strategy is computed.

The appraisers who participated in the evaluation were asked to determine which ones of the 12 recommendations[8], if there were any, were relevant games with respective to the corresponding designated game, which is one of the VideoGG games. The three games marked as *relevant most often* by the appraisers were considered the *gold standard* for the designated game (and the corresponding profile in the case of our recommender system).

Besides marking which recommendation was relevant, the appraisers were also asked to *rank* the recommendations in terms of their degrees of relevance with respect to the corresponding designated game. After the gold standard for each one of the 500 test cases provided by the 78 appraisers were recorded, we computed the *average precision* at rank 1 (i.e., average P@1), at rank 3 (i.e., average P@3), and the *normalized Discounted Cumulative Gain* (nDCG) of the results. As shown in Figure 6, our recommender is significantly better than Best Buy, and outperforms Amazon and Walmart, which are statistically significant ($p < 0.05$).

[5] https://www.amazon.com
[6] https://www.bestbuy.com/
[7] https://www.walmart.com/
[8] Three each from Amazon, Best Buy, Walmart, and our recommender, which were the top-3 recommendations made by each of the four recommenders, respectively. The appraisers had no idea which recommendation was made by which recommender.



(a) Baseline score pairs in the test group

(b) Baseline score pairs in the control group

Figure 5: Distribution of baseline score pairs in the test and control group, respectively

7.5 Performance Evaluation of Our KNN Recommendation Strategy

During the 2^{nd} phase of our game recommendation methodology, we run KNN on the updated profile of a new user to obtain the affinity value for each therapeutic game. We employ KNN on the data obtained through the volunteered students at ScenicView academy. We used the baseline scores and attribute values of games, set the number of neighbors, i.e., K, as **five**, and ran the KNN algorithm to obtain the affinity values. In order to experimentally verify that KNN is the appropriate choice among different existing regression and classification approaches, we ran the same dataset through other approaches as well. We calculate the RMSE values for multilayer perceptron (0.38), linear regression (0.29), KNN (0.12) and SVM (0.25), and KNN has the least RMSE value.

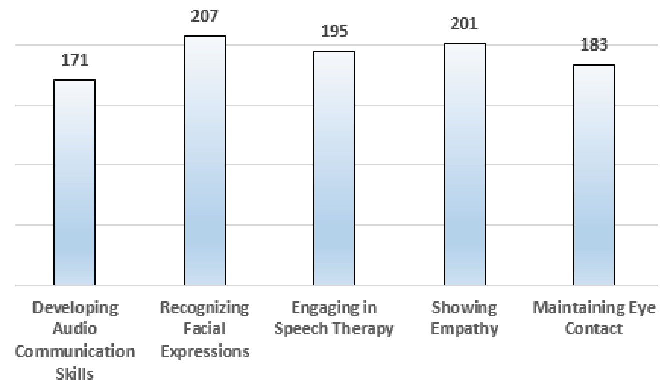


Figure 7: Distribution of games by weakness areas

7.6 Statistical Data of the Empirical Study

Apart from studying the effectiveness of our gaming and recommender system, the data obtained from the test and control group also provide useful demographic information. Statistical data, such as the number of games that were played, number of players, types of game played, etc., convey important information. We studied these statistical data based on different metrices: (i) the time period when the games were played, (ii) the number of responses received in each weakness areas, and (iii) the gender of students.

7.6.1 Distribution of Weakness Areas by Months. We studied the distribution of games played between April and December 2019 by the students. According to the collected data, the greatest number of games were played in the month of May as participants spent April in getting accustomed to playing the games and caught the pace in the following month.

7.6.2 Distribution of Weakness Areas by Volunteer's Responses. We also computed the distribution of responses of the students in each weakness area to determine the most prominent weakness area. Figure 7 shows that the greatest number of responses were given in 'Recognizing Facial Expressions', making it the most prominent weakness area among the participating students, whereas 'Developing Audio Communication Skills' is the least prominent as there are least number of responses.

7.6.3 Distribution of Weakness Areas by genders. Out of the 65 participating students in the test group, there were 37 males and 28 females. Figure 8 shows the distribution of weakness areas prevalent on the participants based on genders. The figure reflects that greater number of females struggle in 'Engaging in Speech Therapy' and 'Showing Empathy'. On the other hand, majority of males

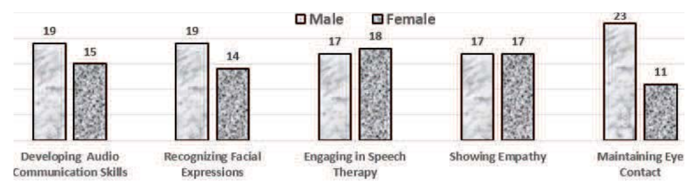


Figure 8: Distribution of weakness areas based on genders

have deficiency in 'Maintaining Eye Contact'. Based on the statistical data by combining the number of males and female students, we draw the conclusion that 'Engaging in Speech Therapy' is the most common weakness areas irrespective of gender. Although Figure 7 shows that the number of responses in 'Recognizing Facial Expression' is the most among all the weakness areas, the number of both male and female participants, whose baseline scores are less than 5.5 in this area, is greater than or equal to other areas, making it one of the most common weakness areas.

8 CONCLUSIONS

Adults with autism (ASD) have difficulty communicating with others, expressing themselves, and making friends that forbid them to live an independent, productive, and fulfilling life [31]. In this paper, we propose a recommender system that suggests therapeutic games to adults with ASD based on the detected weakness areas without relying on user's medical profiles, which are confidential. The effectiveness of our gaming and recommender system has been verified through an empirical study which shows that the social-interactive skills of adults with ASD can be improved by playing the therapeutic games recommended by our system.

REFERENCES

[1] E. Billstedt, I. Gillberg, and C. Gillberg. Autism after Adolescence: Population-based 13- to 22-Year Follow-Up Study of 120 Individuals with Autism Diagnosed in Childhood. *Autism and Developmental Disorders*, 35(3):351–360, 2007.

[2] M. Cederlund, B. Hagberg, E. Billstedt, I. Gillberg, and C. Gillberg. Asperger Syndrome and Autism: A Comparative Longitudinal Follow-up Study More than 5 Years after Original Diagnosis. *Autism and Developmental Disorders*, 38(1):72–85, January 2008.

[3] G. Comrack, C. Clarke, and S. Buttcher. Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods. In *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 758–759, 2009.

[4] D. Dickstein, M. Pescosolido, B. Reidy, T. Galvan, K. Kim, K. Seymour, A. Laird, A. Di Martino, and R. Barrett. Developmental Meta-Analysis of the Functional Neural Correlates of Autism Spectrum Disorders. *The American Academy of Child and Adolescent Psychiatry*, 52(3):279–289, March 2013.

[5] R. Dunifon and L. Gill. Games and Children's Brains: What is the Latest Research? Department of Policy Analysis and Management, College of Human Ecology, 2013. https://www.human.cornell.edu/sites/default/files/PAM/Parenting/FINAL-Video-Game-Research-Brief-5.pdf.

[6] M. Farley, W. McMahon, E. Fombonne, W. Jenson, J. Miller, M. Gardner, H. Block, C. Pingree, E. Ritvo, R. Ritvo, and H. Coon. Twenty-Year Outcomes for Individuals with Autism and Average or Near-Average Cognitive Abilities. *Autism Research*, 2(2):109–118, April 2009.

[7] Edutainment Games. https://tvtropes.org/pmwiki/pmwiki.php/Main/ EdutainmentGame.

[8] T. Griffiths and M. Steyvers. Finding Scientific Topics. In *PNAS*, pages 5228–5235, Boston, 2004. The National Academy of Sciences.

[9] A. Hiniker, D. Wong, and H. Williamson. Go Go Games: Therapeutic Video Games for Children with Autism Spectrum Disorders. In *Proceedings of ACM International Conference on Interaction Design and Children (IDC)*, pages 463–466, 2013.

[10] P. Howlin, S. Goode, J. Hutton, and M. Rutter. Adult Outcome for Children with Autism. *Child Psychology and Psychiatry*, 45(2):212–229, 2004.

[11] C. Hsu and C. Lin. A Comparison of Methods for MulticlassSupport Vector Machines. *IEEE Transactions on Neural Networks*, 13(2):415–425, March 2002.

[12] E. Johnson. Where Are All the Educational Video Games for Adults?, April 2015. https://www.recode.net/2015/4/24/11561860/where-are-all-the-educational-video-games-for-adults.

[13] J. Kirriemuir and A. Mcfarlane. Literature Review in Games and Learning. Technical report, A NESTA Futurelab, 2004. Report 8.

[14] A. Knutas, J. Ikonen, U. Nikula, and J. Porras. Increasing Collaborative Communications in a Programming Course with Gamification: A Case Study. In *Proceedings of the 15th International Conference on Computer Systems and Technologies*, pages 370–377. ACM, 2014.

[15] C. Lister, J. West, B. Cannon, T. Sax, and D. Brodegard. Just a Fad? Gamification in Health and Fitness Apps. *JMIR Serious Games*, 2(2), August 2014.

[16] C. Manning, P. Raghavan, and H. Schutze. *Introduction to Information Retrieval*. Cambridge, 2008.

[17] M. Mazurek, C. Engelhardt, and K. Clark. Video Games from the Perspective of Adults with Autism Spectrum Disorder. *Computers in Human Behavior*, 51(Part A):122–130, October 2015.

[18] M. Mazurek, P. Shattuck, M. Wagner, and B. Cooper. Prevalence and Correlates of Screen-based Media Use Among Youths with Autism Spectrum Disorders. *Autism and Developmental Disorders*, 42(8):1757–1767, August 2012.

[19] A. Memari, N. Panahi, E. Ranjbar, P. Moshayedi, M. Shafei, R. Kordi, and V. Ziaee. Children with Autism Spectrum Disorder and Patterns of Participation in Daily Physical and Play Activities. *Neurology Research International*, pages 1–7, 2015.

[20] Y.-K. Ng and M. Pera. Recommending Social-Interactive Games for Adults with Autism Spectrum Disorders (ASD). In *Proceedings of ACM Conference on Recommender Systems (RecSys)*, pages 209–213, 2018.

[21] V. Nordin and C. Gillberg. The Long-Term Course of Autistic Disorders: Update on Follow-Up Studies. *Acta Psychiatrica Scandinavica*, 97(2):99–108, February 1998.

[22] P. Ohring. Web-Based Multi-Player Games to Encourage Flexibility and Social Interaction in High-Functioning Children with Autism Spectrum Disorder. In *Proceedings of International Conference on Interaction Design and Children, Vol. 7*, pages 171–172, 2008.

[23] N. Pe-Miguel and S. Mximo. Educational Games for Learning. *Universal Journal of Educational Research*, 2(3):230–238, 2014.

[24] Research Autism. Introduction to the Issues Facing People on the Autism Spectrum, June 2018. http://www.researchautism.net/autism-issues/autism-issues-introduction.

[25] K. Ringland, C. Wolf, L. Boyd, M. Baldwin, and G. Hayes. Would You Be Mine: Appropriating Minecraft as an Assistive Technology for Youth with Autism. In *Proceedings of International ACM SIGACCESS Conference*, pages 33–41, 2016.

[26] G. Seber and A. Lee. *Linear Regression Analysis, 2nd Edition*. Wiley, 2012.

[27] M. Seltzer, M. Krauss, P. Shattuck, G. Orsmond, A. Swe, and C. Lord. The Symptoms of Autism Spectrum Disorders in Adolescence and Adulthood. *Autism and Developmental Disorders*, 33(6):565–581, 2003.

[28] M. Seltzer, P. Shattuck, L. Abbeduto, and J. Greenberg. Trajectory of Development in Adolescents and Adults with Autism. *Mental Retardation and Developmental Disabilities Research Reviews*, 10:234–247, 2004.

[29] P. Shattuck, M. Seltzer, J. Greenberg, G. Orsmond, D. Bolt, S. Kring, and C. Lord. Change in Autism Symptoms and Maladaptive Behaviors in Adolescents and Adults with an Autism Spectrum Disorder. *Autism and Developmental Disorders*, 37(9):1735–1747, 2007.

[30] J. Stevenson and K. Hart. Psychometric Properties of the Autism-Spectrum Quotient for Assessing Low and High Levels of Autistic Traits in College Students. *Autism and Developmental Disorders*, 47(6), 1838-1853 2017.

[31] J. Stewart. *Bridges We Build: The Art of Making Friends, A Guide for Connecting, Relating, and Socializing with Others.* ScenicView Academy, 2016.

[32] J. Tang, C. Deng, and G. Huang. Extreme Learning Machine for Multilayer Perceptron. *IEEE Transactions on Neural Networks and Learning Systems*, 27(4):809–821, May 2015.

[33] J. Torrente, A. del Blanco, P. Moreno-Ger, and B. Fernandez-Manjon. Designing Serious Games for Adult Students with Cognitive Disabilities. In *Proceedings of International Conference on Neural Information Processing*, pages 603–610, 2012.