Technology-supported social skills training systems: A systematic literature review

Ding, Ding; Remeijsen, Pascal; Song, Zian; Neerincx, M.A.; Brinkman, W.P.

DOI
10.1109/CSCWD61410.2024.10580403

Publication date
2024

Document Version
Final published version

Published in
Proceedings of the 2024 27th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2024

Citation (APA)
Ding, D., Remeijsen, P., Song, Z., Neerincx, M. A., & Brinkman, W. P. (2024). Technology-supported social skills training systems: A systematic literature review. In W. Shen, W. Shen, J.-P. Barthes, J. Luo, T. Qiu, X. Zhou, J. Zhang, H. Zhu, K. Peng, T. Xu, & N. Chen (Eds.), *Proceedings of the 2024 27th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2024* (pp. 3249-3256). (Proceedings of the 2024 27th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2024). https://doi.org/10.1109/CSCWD61410.2024.10580403

Proceedings of the 2024 27th International Conference on Computer Supported Cooperative Work in Design

2024 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD) | 979-8-3503-4918-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/CSCWD61410.2024.10580403

Technology-supported social skills training systems: A systematic literature review

Ding Ding*[1], Pascal Remeijsen[2], Zian Song[1], Mark A. Neerincx[2,3], and Willem-Paul Brinkman[2]
[1]Southeast University, Nanjing, China
[2]Delft University of Technology, Delft, The Netherlands
[3]TNO Perceptual and Cognitive Systems, Soesterberg, The Netherlands
*Corresponding author. Email address: dingding-1@seu.edu.cn

***Abstract*—Social interactions form an essential aspect of people's life, however, it is quite challenging for individuals to handle a wide range of social situations. Therefore, a variety of training systems have been developed to improve their skills. This literature review seeks to give an overview of the state of the art of technology-supported systems for social skills training. The studies eligible for inclusion described a technology-supported system with the purpose of training social skills and included an experimental or observational study to evaluate the efficacy of the system. 225 studies (224 publications) with 216 systems were identified, characterized, and analyzed in this literature review. Using the taxonomy as put forward in this study, the analysis shows that the majority of these systems were screen-based applications, with virtual reality technology being the most frequently observed. The systems most often targeted communication skills that focus on transferring information to produce greater understanding, i.e. mending general communication impairments in children with autism. In terms of functions, support for learning-by-doing was the most observed function, while focusing on job interviews provided the largest number of functions. Finally, the studies reported overwhelmingly positively regarding the systems' impact, including 76 studies with a randomized controlled trial design. Still, most studies only used a quasi-experimental design based on self-report measures. We anticipate the proposed taxonomy to be a starting point for researchers to position their work and that the review will help them with gaining inspiration for the design and evaluation of social skills training systems.**

***Index Terms*—Social skills training, training system, technology supported, state of the art, systematic review**

I. Introduction

Social interactions permeate every aspect of our life as they occupy a considerable part of most people's waking life [1]. Unfortunately, for some, it is challenging to function well in situations such as public speaking [2], negotiations [3] and job interviews [4]. Social skills impediments can have various negative impacts. It can hamper peer acceptance and academic achievement for children, bring about vocational difficulties for adults, cause an economic loss for organizations, and even lead to life-threatening situations for soldiers [5]. Extreme examples of people with difficulties are those with social phobia or autism spectrum disorder. Given the impact of social interaction on everyone's daily life, self-help books [6], [7] and face to face training [8] have been developed. For clinical cases such as social anxiety disorder, various therapies have also been proposed, including social skills training (SST), cognitive-behavioural therapy (CBT) or even medication [9]. Typically a teacher or a therapist performs this conventional social skill training in person. The access is often mainly limited to people with serious disabilities or the elite, owing to their high cost [10]. Furthermore, creating adequate and controlled social interaction is difficult [11]. Since they carry the promise of cost-effectiveness and controllability, it is not surprising that researchers have flocked to study technology-supported systems for social skills training and therapy. In the past decades, a wide range of systems have been reported in the literature, applying a variety of technologies, including but not limited to immersive virtual reality (VR), augmented reality, robots, and screen-based software. Although the number of such systems and studies is considerable, state of the art about them seems still unclear.

This paper aims to present an overview of technology-supported social skills training systems from the scientific literature. Previous reviews can be categorized into three main types. First are the reviews that focus primarily on the treatment methods or theories underlying therapy for social skills [12]–[15]. They do not take a system or technology perspective. Second are the reviews that describe social skills training only applying a specific technology, such as virtual reality or robots [16]–[18]. Third are the reviews based on social skills training systems but targeting a specific group, such as children with autism [15], [19]–[21].

This study intends to cover various technologies, to target a wide range of audiences and skills, to consider both treatment- and therapy-oriented training systems, as well as both general teaching- and practice-oriented training systems. To the best of our knowledge, this is the first survey with this particular scope. There are several merits to be discussed. First, this study can serve as a state of the art overview of the field. Consequently, this paper could be a starting point for readers looking for future research directions or answers to questions regarding state of the art. For example, questions about which technologies, skills, and target audiences are most actively being studied? In addition to this, the paper outlines what functions are often included in these systems, and how these systems are evaluated.

To this end, this paper addresses the following three research questions: 1. How have these systems been developed? e.g., which types of technologies have been used? Which group of population or social skills do these systems target to most? 2.

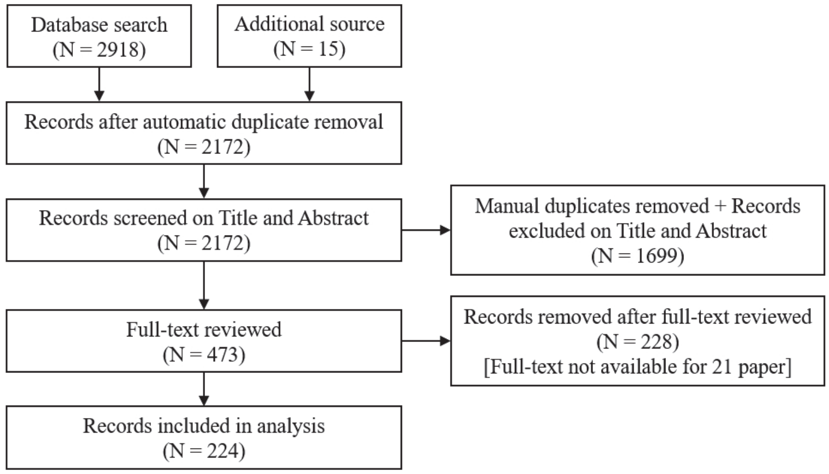

Fig. 1. PRISMA-Diagram for the filtering process

Which functions do these systems provide and how do they relate to the different applied technologies, target skills and target groups? 3. What is the overall perception of the efficacy of these training systems on social skills?

II. Methods

A. Search procedure

A systematic review was conducted using Scopus and Web of Science, two widely used abstract databases. The search query consisted of three major components: 1. targeting social skills, 2. describing a technology-supported system with the purpose of training, 3. reporting an empirical evaluation of the system. The list of search terms for the social skills concepts included synonyms for social skills, such as social capability, and social competence, as well as terms for specific common social skills, for example, public speaking, interview skills, or negotiation. The search terms for technology-supported systems included synonyms and common technologies, such as computer aided, computer based, virtual agents, and mobile applications. Appendix shows the detailed search queries. Reviews, surveys and meta-analysis papers were excluded as the last step. The search for English language publications was first conducted on October 02, 2017, updated on December 16, 2018 [22], and then final updated on June 30, 2022.

The search resulted in a total of 2918 conference papers, journal articles, and book chapters. Reference lists of the included papers were checked for potential papers; it resulted in 15 further papers. After automatically removing the duplicates, 2172 papers remained for further selection.

As shown in Fig. 1, a two-step review was carried out: screened based on the title and abstract and screened based on the full-text version of the papers. There were two coders active in this study. At each step, a coder performed the entire review while the other coder served as a control to determine inter-rater agreement. The control consisted of a random subsample of records that were double coded. The inter-observer agreement was calculated using a Kappa-metric [23]. For the selection based on title and abstract, a random sample of 600 papers was used for the double coding, resulting in an agreement of 0.90. For the second round, a random sample of 74 papers was used, resulting in an agreement of 0.99. Both showed acceptable levels of reliability and were considered sufficient to move on to the next step of selection.

B. Inclusion Criteria

We included articles that: (1) were available and accessible in a full-text version of the paper, written in English, (2) included a technology-supported system with the purpose of (3) training social skills, and (4) described an empirical study.

For criterion (2), all types of digital medium were included (i.e. iPads, Computers, Head-mounted displays, Robots). Systems could also have a human component. Regarding criterion (3), the training is not limited to therapy or treatment but also includes general teaching, education or practice. Furthermore, the training could target a behaviour, cognition or motivation related to social interactions [24].

C. Exclusion Criteria

Simply providing information in a digital format was not considered sufficient (e.g. a normal curriculum, but just online [25]). Also, papers about entirely human-based training were excluded (e.g., [26]). Furthermore, as a training system, they should offer something extra in terms of experience or interactions (e.g., excluding [27]). Moreover, systems that were solely used to measure or test skills were excluded (e.g., [28]). As for criterion (4), to ensure that systems were actually developed and functional at some point, papers that did not include an experimental or observational evaluation of the training systems were excluded (e.g., [29]).

D. Coding

Fig. 2 shows the taxonomy used for coding the data extracted from the papers. Each block represents a core concept related to the training system and the evaluation of it. The blocks have attributes, addressing the characteristics of each

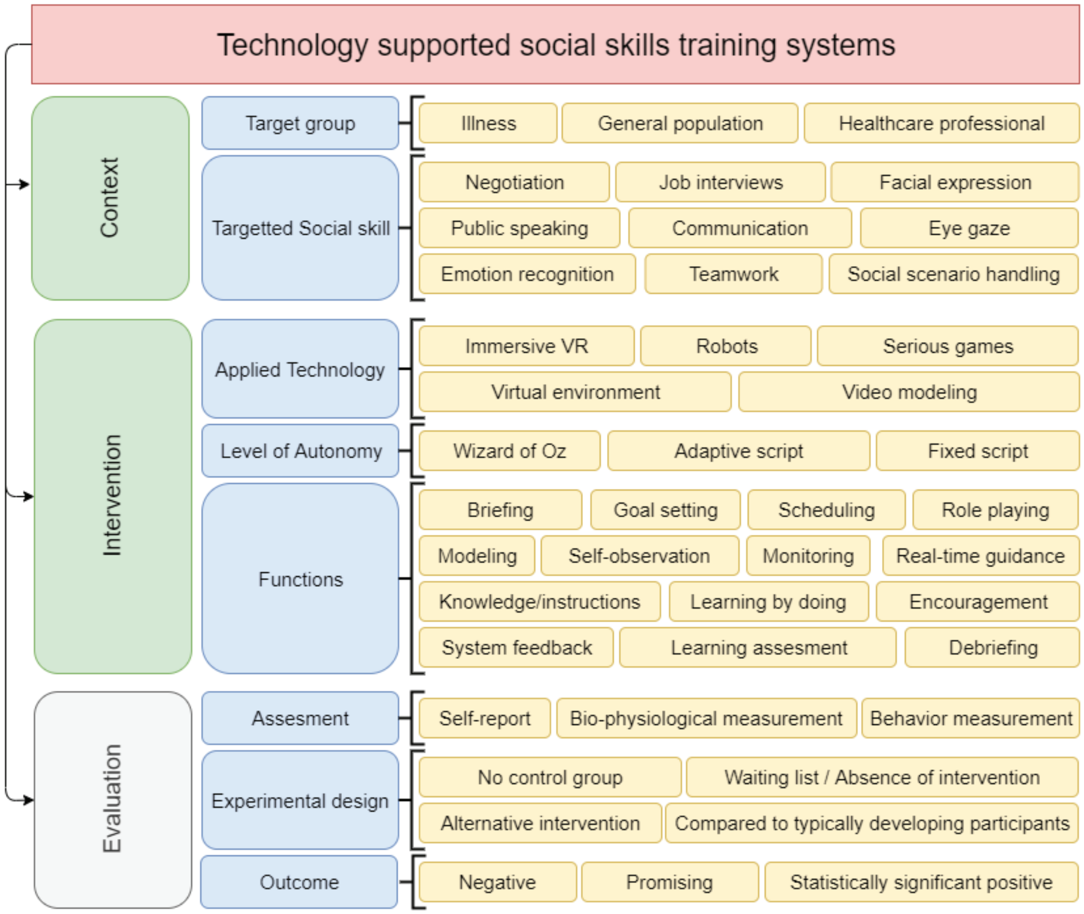

Fig. 2. Taxonomy of technology-supported social skills training systems

system and the evaluation. General information about the system such as system name, authors, year of publication, the reported location of the study etc. was collected directly from the paper.

As shown in Fig. 2, eight attributes were collected to categorize systems and to obtain a better comprehension of the social skills training systems. They are the technologies applied by the system, the target group of the system, the target skills of the system, level of autonomy of the system, functions provided by the system, the types of assessment used in the experiments, experimental design, and the results of the experiments. Within these eight attributes, the technologies applied by the system, the target group of the system, and level of autonomy of the system are single forced choice, while the left attributes allowed the coders to pick more than one. Most attributes are self-explanatory. Some need some further clarification.

Firstly, the systems were categorized into four types based on the technology applied. A distinction was made between immersive virtual reality, robots, screen-based applications and a remaining category for other technologies. The screen-based application category was again subdivided into three popular paradigms: serious games, virtual environments/agents, and video modelling. Secondly, a crucial task in the coding process was identifying the training system by functional elements, i.e., functions. Functions were the building blocks that made up the interventions of the training system, e.g. a function that allowed users to role-play an interviewee attending a job interview. The functions considered were chosen from both classical therapy and teaching methodologies [12], [30], [31]. Fig. 2 shows a list of the functions. Note that these functions were only considered when the technology-supported system provided them and not when the external environment, e.g. a human, embodied them. For instance, when a person at the start of training gave the user a thorough explanation, it was not regarded as a function of the system. Thirdly, the level of autonomy attribute illustrates the autonomous level of the system. It was categorized into three types: Wizard-of-Oz, Fixed scripts, and Adaptive scripts. Wizard-of-Oz is a popular paradigm for designing social skills training systems. When interacting with a Wizard-of-Oz prototype, people believe the system operates autonomous, where, in fact, an unseen human operator fully or partially controls the system. With fixed script systems, all users received a similar pre-defined system response. With an adaptive script, on the other hand, the system's response depended more on users'

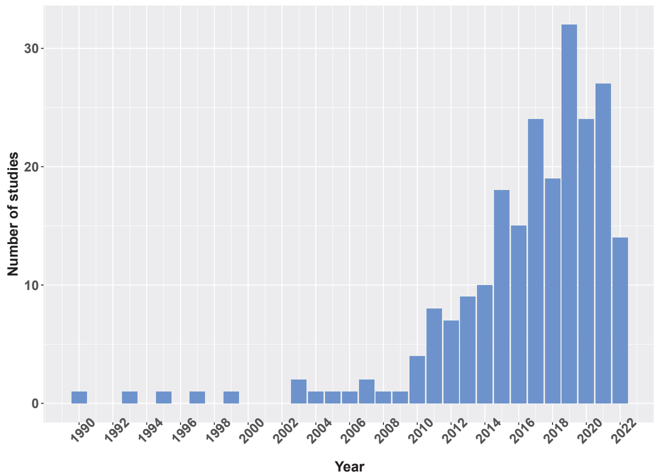

Fig. 3. Distribution of records over the years

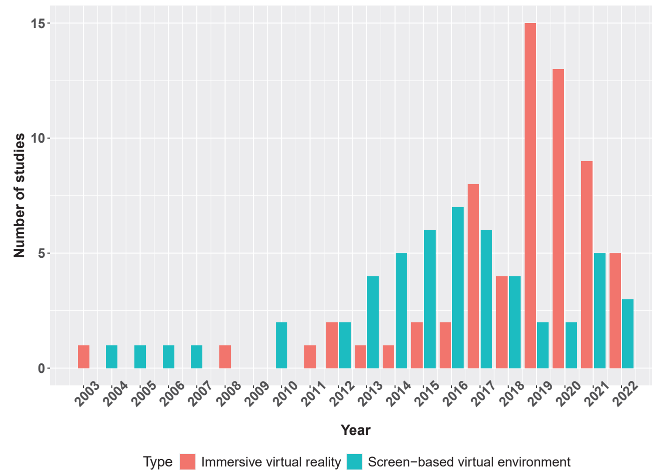

Fig. 4. Evolution of immersive VR vs Virtual Environments

input. Often this required the use of either some database or artificial intelligence algorithms. Fourthly, the experimental design attribute, four types were considered depending on whether the intervention group was compared with: a waiting list, an alternative intervention, typically developing (TD) person, or no control group.

E. Statistical Analysis

We conducted a descriptive analysis with R version 3.4.2. All the review data, the R scripts, and output files can be found online [1].

III. Results

Tables that show the characteristics of the technology-supported social skills training systems this review examined can be found online[1]. Sometimes the same version of a system was used in multiple studies, but with unique experiments [32]–[35]. In these cases, all studies were included in the table, while only one system was counted. If multiple papers presented the same version of a system with the identical experiment(s), only the latest record was included. In total, the review identified 225 studies (224 papers, one paper has two studies) describing 216 systems. Coding these systems on their key attributes led to the characterization presented in Table I.

A. Systems, versions and basic attributes

Systems had 1.00 versions on average, while more than 97.21% only had one. Therefore, most systems seem to have been one-offs without being part of a long-term scientific improvement cycle. Only six systems had two or more versions, for example, the "Automated social skills trainer" with three versions [36]–[38].

[1]During paper review phase, files can be accessed from the following link https://drive.google.com/drive/folders/1Hf3tPHfkPl99l3aJe3bM0Y9D05FXoHxu?usp=sharing. If the paper is published, these files will be moved for long-term storage. The DOI to this storage will be included in the paper.

The investigation also looked at where the research was carried out. If this was not explicitly stated in the paper, the first author's institution location was taken. North America was responsible for 39.56% of the studies, occupies the largest part. Moreover, the USA (36.89%), China (8.00%), The Netherlands (8.00%), Japan (4.00%) and the UK (3.56%) were the top 5 contributors.

Furthermore, the review examines how the number of published studies changed over time. As shown in Fig. 3, the number of research studies has exhibited fluctuations over the past several years. There was a consistent upward trend, peaking in 2019, followed by a small decline. The difficulty in conducting experiments due to the COVID-19 outbreak may have contributed to the decline. The figure shows a drop for 2022 because of the unaccounted research output from the latter half of this year. Moreover, The data also indicates varying trends in technology adoption. According to Fig. 4, prior to 2016, the number of studies using screen-based VR was greater than the number of studies using immersive VR and was increasing each year. Whereas, starting in 2017, the number of studies using immersive VR began to exceeds those using screen-based VR and reached a maximum in 2019. This may have happened due to the successive release of consumer-grade immersive VR devices after 2016. In recent years, the gap in the number of studies using the two technologies narrows. It might be that the challenges posed by the COVID-19 pandemic have made experiments related to the use of immersive VR difficult to conduct, while screen-based VR is relatively more convenient. However, overall, the former still exceeds the latter in quantity.

B. Technologies applied and target group

As shown in Table I, the most common type of system for social skills training is a screen-based system (61.57%, 133/216). Among them, the virtual avatar was most often used (21.76%, 47/216). Nevertheless, in respect of a specific technology, a broader definition of virtual reality technology that covers both immersive virtual reality (29.17%, 63/216) and virtual avatar (21.76%, 47/216), also accounted for half of the systems. Other types of technology were serious games (15.74%, 34/216), robots (9.72%, 21/216), and video modelling (4.63%, 10/216). Most systems belong to only one type, with one exception [39], this system has both screen-based part and immersive virtual reality part.

Of the 216 systems, 56.02% (121/216) were aimed at people with a specific medical or psychological disorder. Among these, people on the autism spectrum were the most common (66.12%; 80/121), and 61.16% (74/121) of these systems target children specifically. Besides this, about a quarter of the studies were targeting the general population (26.39%; 47/216). Finally, another popular target audience was health-care professionals. Understandably, a large number of health care jobs rely on patient interaction. As can be seen in Table I, 14.35% (31/216) of the systems targeted health care professionals.

C. Target skills

The most common targeted skill of the training systems is communication (42.59%, 92/216), followed by social scenario handling (25.00%, 54/216). Skills like negotiation (9.26%, 20/216), job interviews (7.41%, 16/216) and teamwork/shared decision making (6.48%, 14/216), their proportion is relatively low. Another finding is that 54.17% (117/216) of systems focused on macro skills, compared to 16.20% (35/216) and 29.63% (64/216) for micro and a combination of both respectively. Evidently, systems often targeted skills like social scenario handling and public speaking, also choosing to focus on specific elements such as emotion recognition and eye contact.

D. Size and Functionality

The average system offered 3.07 functions ($SD = 1.09$) with a range from one to seven. There is a considerable difference between the most and the least frequently provided functions. Four functions that are provided often in classical therapy or conventional therapy-oriented social skills training [40] seemed to appear less often in the technology-supported systems. These are briefing (4.63%, 10/216), scheduling (0.93%, 3/216), goal setting (9.26%, 20/216) and debriefing (0.46%, 1/216). For most cases, the whole training program still might have offered these functions; however, a person might have carried them out, e.g., a trainer that provided the debriefing. Consequently, the review did not count them as system functions. On the other side of the spectrum, functions like learning by doing (82.41%, 178/216), role-playing (58.80%, 127/216) as well as imparting knowledge (45.83%, 99/216) seem much more popular.

Moreover, by looking at the number of functions used when targeting a certain social skill, we found the systems focusing on job interviews to provide the largest number of functions ($M = 3.56$, $SD = 1.03$), while the systems targeting teamwork or shared decision making skills provided the fewest functions ($M = 2.86$, $SD = 0.86$). Besides, from the perspective of the target group, the average number of functions provided by systems targeting clinical population, health-care professional, and the general population was 3.04, 2.94, 3.26, which was quite close to each other. Nevertheless, except these three typical population, the systems targeting the other population offered much fewer functions ($M = 2.00$, $SD = 0.63$). When investigating the number of functions provided by the systems employing different technologies, the systems employing video modelling technology were designed with the largest number of functions ($M = 3.50$, $SD = 1.43$).

E. Evaluation

1) Assessment: The most used form of assessment for the studies included in this review was self-reporting (66.22%, 149/225). This included reports about their experience, anxiety, self-efficacy, and social skills. Another common form of assessment was behavioural observations (60.00%, 135/225). Bio-physiological measurements (7.56%, 17/225) were less frequently reported. They included measurements like blood pressure, heart rate, and stress-levels through skin conductance.

2) Experimental design: The majority of the experiments described in the study were quasi-experiments (66.22%, 149/225). Often the evaluations did not include a follow-up assessment (80.89%, 182/225). Still, 19.11% (43/225) of the studies had a follow-up assessment. Among them, 48.84% (21/43) were done within four weeks; while 18.60% (8/43) of the studies measured after more than half a year. The average sample size of the studies was 51.31 participants, ranging from 1 to 1178 participants; respectively, 30.85 participants for studies with a single-group design, 38.48 participants per group for the studies with a comparison or control group. A considerable number of these studies might have been underpowered when considering a 5% level of significance, 80% statistical power to detect at least a large effect size (0.80 Cohen'd) [41], [42]. Only 24.03% of the studies with a single-group design, and 31.25% of the studies with a multi-group design seems to have had an adequate sample size, taking a group size of 28 [1] and 26 [2] as the cutoff respectively.

3) Results of experiments: Table I shows that 98.67% of the studies suggested that the system had a positive impact. 77.33% (174/225) reporting a statistically significant positive or improvement result, while 21.33% (48/225) indicated the system shows promise without underlying hypothesis testing support. Only 1.33% of studies reported negative results, from which one study based on statistical hypothesis testing. Of the studies that had a randomized controlled trial design, 93.42% (71/76) reported statistically significant positive or improvement results on measures such as conversational skills rating scale [43] and liebowitz social anxiety scale [44].

[1] Group size is based on a significant test for a product-moment correlation coefficient with a large effect size [41].

[2] Group size is based on a significant test for difference between two independent sample means with a large effect size [41].

TABLE I
DISTRIBUTION OF KEY ATTRIBUTES

	%		%		%		%
Applied technology		Job Interviews	7.41	Scheduling	0.93	Biophysiological measurement	7.56
Immersive VR	29.17	Communication	42.59	Role playing	58.80	Other	4.89
Robot	9.72	Eye gaze movement/eye contact	19.44	Modeling	17.59		
Screen Based		Facial expression	10.65	Self-Observation	8.33	**Experimental control**	
Serious Game	15.74	Emotion/facial/body language recognition	25.93	Monitoring	6.02	Yes(absence of intervention)	15.11
Virtual Avatar	21.76	Teamwork/shared decision making	6.48	Real-time guidance	9.26	Yes(compared to alternative)	22.67
Video Modeling	4.63	Social scenario handling	25.00	Imparting knowledge/Instructions	45.83	Yes(compared to TD person*)	3.11
Other	19.44	Other	30.09	Learning by doing	82.41	No (single group)	57.33
				Encouragement, praise and rewarding	17.13		
Target group		**Level of Autonomy**		Performance feedback and reflection	36.11	**Results of experiment**	
Illness	56.02	Wizard of Oz	10.65	Learning assessment	10.65	Positive (Statistically significant)	77.33
Healthcare professional	14.35	Adaptive interaction	20.83	Debriefing	0.46	Positive (without statistically support)	21.33
General Population	26.39	Fixed interaction	60.19			Negative (Statistically significant)	0.44
Other	2.78	Other	7.41			Negative (without statistically support)	0.89
Targeted Social Skills		**Functions**		**Assessment**			
Negotiation	9.26	Briefing	4.63	Self-report	66.22		
Public speaking	16.20	Goal setting	9.26	Behavior observation	60.00		

*TD person: Typically developing person.

IV. Discussion and Conclusion

This study proposed a taxonomy, which researchers can use to position their work. Furthermore, in line with the research questions, the findings allow for three main conclusions. Firstly, although the review found that social skill training systems use a variety of technologies, two-thirds of the systems were screen-based applications; from which, virtual reality was studied most frequently. The most targeted social skill for these training systems was communication skills, with more than one-third of the systems developed to train it. Secondly, the function learning-by-doing was provided by most of the systems. Besides, the systems focusing on job interviews provided the largest number of functions, while the systems employing video modelling technology were designed with the most functions. This indicates that these systems were the most extensive. Finally, 98.67% studies reported the systems to have a positive impact, such as improving people's feeling, cognitions, emotions or behaviour. Among these studies, there were 71 studies with a randomized controlled trial design that reported statistically significant improvements.

Besides the above, there are some other interesting findings. The review found a continuous growth in the number of systems developed each year. Evidently, research into these systems is still ongoing and attracting more momentum. The application of technologies such as robots [45], augmented reality [46], and combination of virtual reality and brain-computer interfaces [47] seem the latest to attract attention in this field. In recent years, studies on immersive virtual reality has experienced a significant increase followed by a gradual decrease compared to screen-based virtual environment. Besides technology diversification, a broadening of the target audience is also possible, as most systems in the review focused on a clinical population targeting people with social anxiety and autism. While the review found systems targeting health care professionals, it also found systems that target other professions such as software engineers [48], law enforcers [49], and crisis managers [50]. Potentially, more jobs that depend on social skills could benefit likewise.

Additionally, developers can also extend on the system's functionality. Only one-third of the systems provided feedback functionality, despite meta-analysis research [51] having identified that receiving feedback on one's performance as one of the essential components in social skills training, next to practising. Similarly, the function, "positive reinforcement", which includes encouragement, praise and rewarding is critical for achieving a positive impact on the improvement of social skills [52]. However, less than one-fifth of the systems offered this and therefore is another opportunity for extending current systems.

For the design of the evaluations, the review shows that most studies used a quasi-experiment design, almost two-thirds of the research only conducted a single group study, without a comparison or control group. Furthermore, roughly 70% to 80% of the evaluations had an insufficient sample size making them underpowered. Still, well-powered studies with a true experimental design are essential for studying causal effects, and therefore, for extending scientific understanding about the impact of these systems. Likewise, more long term follow-up studies, as they were rarely reported, would also help understanding lasting effects. Furthermore, despite the popularity of self-reported measures, they received criticism for their measurement bias potential [53]. Therefore, researchers should consider including other types of measures as well. As for assessments, studies that target general population prefers to use self-report (80.00%, 48/60), while that target people with illness are more inclined to use behavior observation (68.60%, 83/121). This might because that among the latter, 61.16% (74/121) of the participants are children with illness, and it is difficult for them to give accurate self-reports.

Of course, the review also has several limitations that could be noticed when considering its implications and generalisation. First, although the search included several synonyms, some authors might still have used other terms, making it impossible to claim the review to be exhaustive. Despite the possibility that the review ignored some papers, the broad search query, and the subsequently large number of systems identified makes the study's general observations likely to be reliable and representative for peer-reviewed literature in this area. Second, the review only considered systems reported in the literature, ignoring potentially commercial systems not reported in the searched literature. Third, the often reported publication towards positive results [54]–[56], makes that overwhelmingly positive evaluation reports should be considered with some caution. Finally, the review has only used descriptive statistics and refrained itself from apply inferential statistics. Still, the review with the papers and systems identified could form the basis for future meta-analysis.

To sum up, this review presents a comprehensive overview of the state of the art of technology-supported social skills training systems and identifies some of the characteristics, challenges, and trends in this field. Taken together, it offers inspirations for developing new social skills training systems and serve a starting point for further research.

V. Acknowledgement

This work was supported by the National Natural Science Foundation of China (NSFC) [grant number: 62102081].

References

[1] M. R. Mehl and J. W. Pennebaker, "The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations." *Journal of personality and social psychology*, vol. 84, no. 4, p. 857, 2003.

[2] M. B. Stein, J. R. Walker, and D. R. Forde, "Public-speaking fears in a community sample: Prevalence, impact on functioning, and diagnostic classification," *Archives of General Psychiatry*, vol. 53, no. 2, pp. 169–174, 1996.

[3] R. Fisher, W. L. Ury, and B. Patton, *Getting to yes: Negotiating agreement without giving in*. Penguin, 2011.

[4] K. Anderson, E. André, T. Baur, S. Bernardini, M. Chollet, E. Chryssafidou, I. Damian, C. Ennis, A. Egges, P. Gebhard *et al.*, "The tardis framework: intelligent virtual agents for social coaching in job interviews," in *International Conference on Advances in Computer Entertainment Technology*. Springer, 2013, pp. 476–491.

[5] J. M. Kim, R. W. Hill Jr, P. J. Durlach, H. C. Lane, E. Forbell, M. Core, S. Marsella, D. Pynadath, and J. Hart, "Bilat: A game-based environment for practicing negotiation in a cultural context," *International Journal of Artificial Intelligence in Education*, vol. 19, no. 3, pp. 289–308, 2009.

[6] D. Carnegie, *How to win friends & influence people*. Musaicum Books, 2017.

[7] M. L. Silberman, M. Silberman, and F. Hansburg, *People smart: Developing your interpersonal intelligence*. Berrett-Koehler Publishers, 2000.

[8] R. Bulkeley and D. Cramer, "Social skills training with young adolescents," *Journal of Youth and Adolescence*, vol. 19, no. 5, pp. 451–463, 1990.

[9] R. A. Gould, S. Buckminster, M. H. Pollack, M. W. Otto, and L. Y. Massachusetts, "Cognitive-behavioral and pharmacological treatment for social phobia: A meta-analysis," *Clinical Psychology: Science and Practice*, vol. 4, no. 4, pp. 291–306, 1997.

[10] J. Gratch, D. DeVault, and G. Lucas, "The benefits of virtual humans for teaching negotiation," in *International Conference on Intelligent Virtual Agents*. Springer, 2016, pp. 283–294.

[11] G. Robillard, S. Bouchard, S. Dumoulin, T. Guitard, and E. Klinger, "Using virtual humans to alleviate social anxiety: preliminary report from a comparative outcome study," *Stud Health Technol Inform*, vol. 154, pp. 57–60, 2010.

[12] A. Kopelowicz, R. P. Liberman, and R. Zarate, "Recent advances in social skills training for schizophrenia," *Schizophrenia bulletin*, vol. 32, no. suppl_1, pp. S12–S23, 2006.

[13] J. W. Maag, "Social skills training for students with emotional and behavioral disorders: A review of reviews," *Behavioral Disorders*, vol. 32, no. 1, pp. 4–17, 2006.

[14] C. R. Cook, F. M. Gresham, L. Kern, R. B. Barreras, S. Thornton, and S. D. Crews, "Social skills training for secondary students with emotional and/or behavioral disorders: A review and analysis of the meta-analytic literature," *Journal of Emotional and Behavioral Disorders*, vol. 16, no. 3, pp. 131–144, 2008.

[15] P. A. Rao, D. C. Beidel, and M. J. Murray, "Social skills interventions for children with asperger's syndrome or high-functioning autism: A review and recommendations," *Journal of autism and developmental disorders*, vol. 38, no. 2, pp. 353–361, 2008.

[16] O. Mubin, C. J. Stevens, S. Shahid, A. Al Mahmud, and J.-J. Dong, "A review of the applicability of robots in education," *Journal of Technology in Education and Learning*, vol. 1, no. 209-0015, p. 13, 2013.

[17] L. Freina and M. Ott, "A literature review on immersive virtual reality in education: state of the art and perspectives," in *The International Scientific Conference eLearning and Software for Education*, vol. 1. " Carol I" National Defence University, 2015, p. 133.

[18] M. C. Howard and M. B. Gutworth, "A meta-analysis of virtual reality training programs for social skill development," *Computers & Education*, p. 103707, 2019.

[19] S. Parsons and P. Mitchell, "The potential of virtual reality in social skills training for people with autistic spectrum disorders," *Journal of intellectual disability research*, vol. 46, no. 5, pp. 430–443, 2002.

[20] S. Parsons and S. Cobb, "State-of-the-art of virtual reality technologies for children on the autism spectrum," *European Journal of Special Needs Education*, vol. 26, no. 3, pp. 355–366, 2011.

[21] S. Ramdoss, W. Machalicek, M. Rispoli, A. Mulloy, R. Lang, and M. O'Reilly, "Computer-based interventions to improve social and emotional skills in individuals with autism spectrum disorders: A systematic review," *Developmental neurorehabilitation*, vol. 15, no. 2, pp. 119–135, 2012.

[22] D. Ding, "Design and evaluation of simulated reflective thoughts in virtual reality exposure training," 2020.

[23] M. B. Miles and A. Huberman, "a.(1994). qualitative data analysis: An expanded sourcebook," *Evaluation and Program Planning*, pp. 88 232–2, 1984.

[24] S. G. Little, J. Swangler, and A. Akin-Little, "Defining social skills," in *Handbook of Social Behavior and Skills in Children*. Springer, 2017, pp. 9–17.

[25] E. Wittenberg-Lyles, J. Goldsmith, B. Ferrell, and M. Burchett, "Assessment of an interprofessional online curriculum for palliative care communication training," *Journal of palliative medicine*, vol. 17, no. 4, pp. 400–406, 2014.

[26] L. F. Gooding, "The effect of a music therapy social skills training program on improving social competence in children and adolescents with social skills deficits," *Journal of music therapy*, vol. 48, no. 4, pp. 440–462, 2011.

[27] T. Köhler, I. Fischlmayr, T. Lainema, and E. Saarinen, "Bringing the world into our classrooms: The benefits of engaging students in an international business simulation," in *Increasing student engagement and retention using classroom technologies: Classroom response systems and mediated discourse technologies*. Emerald Group Publishing Limited, 2013, pp. 163–198.

[28] J. McGrath, N. Kman, D. Danforth, D. P. Bahner, S. Khandelwal, D. R. Martin, R. Nagel, N. Verbeck, D. P. Way, and R. Nelson, "Virtual alternative to the oral examination for emergency medicine residents," *Western Journal of Emergency Medicine*, vol. 16, no. 2, p. 336, 2015.

[29] O. Miglino, A. Di Ferdinando, A. Rega, and B. Benincasa, "Sisine: Teaching negotiation through a multiplayer online role playing game," in *The 6th European Conference on E-Learning, Copenhague, Danemark*, 2007.

[30] F. M. Gresham, "Social skills training with handicapped children: A review," *Review of educational research*, vol. 51, no. 1, pp. 139–176, 1981.

[31] H. P. Sims Jr and C. C. Manz, "Social learning theory: The role of modeling in the exercise of leadership," *Journal of Organizational Behavior Management*, vol. 3, no. 4, pp. 55–63, 1982.

[32] M. R. Kandalaft, N. Didehbani, D. C. Krawczyk, T. T. Allen, and S. B. Chapman, "Virtual reality social cognition training for young adults with high-functioning autism," *Journal of autism and developmental disorders*, vol. 43, no. 1, pp. 34–44, 2013.

[33] N. Didehbani, T. Allen, M. Kandalaft, D. Krawczyk, and S. Chapman, "Virtual reality social cognition training for children with high functioning autism," *Computers in Human Behavior*, vol. 62, pp. 703–711, 2016.

[34] R. P. Sanchez, C. M. Bartel, E. Brown, and M. DeRosier, "The acceptability and efficacy of an intelligent social tutoring system," *Computers & Education*, vol. 78, pp. 321–332, 2014.
[35] R. Sanchez, E. Brown, K. Kocher, and M. DeRosier, "Improving children's mental health with a digital social skills development game: a randomized controlled efficacy trial of adventures aboard the ss grin," *Games for health journal*, vol. 6, no. 1, pp. 19–27, 2017.
[36] H. Tanaka, S. Sakriani, G. Neubig, T. Toda, H. Negoro, H. Iwasaka, and S. Nakamura, "Teaching social communication skills through human-agent interaction," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 6, no. 2, p. 18, 2016.
[37] H. Tanaka, S. Sakti, G. Neubig, H. Negoro, H. Iwasaka, and S. Nakamura, "Automated social skills training with audiovisual information," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2016, pp. 2262–2265.
[38] H. Tanaka, H. Negoro, H. Iwasaka, and S. Nakamura, "Embodied conversational agents for multimodal automated social skills training in people with autism spectrum disorders," *PloS one*, vol. 12, no. 8, p. e0182151, 2017.
[39] A. Kolk, M. Saard, L. Pertens, T. Kallakas, K. Sepp, and K. Kornet, "Structured model of neurorehab: a pilot study of modern multitouch technology and virtual reality platforms for training sociocognitive deficit in children with acquired brain injury," *Applied Neuropsychology: Child*, vol. 8, no. 4, pp. 326–332, 2019.
[40] J. A. Cully and A. L. Teten, "A therapist's guide to brief cognitive behavioral therapy," *Houston: Department of Veterans Affairs South Central MIRECC*, 2008.
[41] J. Cohen, "Quantitative methods in psychology: A power primer," *Psychol. Bull.*, vol. 112, pp. 1155–1159, 1992.
[42] J. Eng, "Sample size estimation: how many individuals should be studied?" *Radiology*, vol. 227, no. 2, pp. 309–313, 2003.
[43] M. R. Ali, D. Crasta, L. Jin, A. Baretto, J. Pachter, R. D. Rogge, and M. E. Hoque, "Lissa—live interactive social skill assistance," in *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2015, pp. 173–179.
[44] S. Bouchard, S. Dumoulin, G. Robillard, T. Guitard, E. Klinger, H. Forget, C. Loranger, and F. X. Roucaut, "Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial," *The British Journal of Psychiatry*, vol. 210, no. 4, pp. 276–283, 2017.
[45] B. Scassellati, L. Boccanfuso, C.-M. Huang, M. Mademtzi, M. Qin, N. Salomons, P. Ventola, and F. Shic, "Improving social skills in children with asd using a long-term, in-home social robot," *Science Robotics*, vol. 3, no. 21, p. eaat7544, 2018.
[46] G. Lorenzo, M. Gómez-Puerta, G. Arráez-Vera, and A. Lorenzo-Lledó, "Preliminary study of augmented reality as an instrument for improvement of social skills in children with autism spectrum disorder," *Education and Information Technologies*, vol. 24, no. 1, pp. 181–204, 2019.
[47] C. Amaral, S. Mouga, M. Simões, H. C. Pereira, I. Bernardino, H. Quental, R. Playle, R. McNamara, G. Oliveira, and M. Castelo-Branco, "A feasibility clinical trial to improve social attention in autistic spectrum disorder (asd) using a brain computer interface," *Frontiers in neuroscience*, vol. 12, p. 477, 2018.
[48] M. D. Nazligul, M. Yilmaz, A. E. Yilmaz, V. Isler, R. O'Connor, M. Gozcu, and P. Clarke, "An interactive 3d virtual environment to reduce the public speaking anxiety levels of novice software engineers," *IET Control Theory and Applications*, 2018.
[49] K. Mykoniatis, A. Angelopoulou, M. D. Proctor, and W. Karwowski, "Virtual humans for interpersonal and communication skills' training in crime investigations," in *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 2014, pp. 282–292.
[50] N. Haferkamp, N. C. Kraemer, C. Linehan, and M. Schembri, "Training disaster communication by means of serious games in virtual environments," *Entertainment Computing*, vol. 2, no. 2, pp. 81–88, 2011.
[51] D. Blanch-Hartigan, S. A. Andrzejewski, and K. M. Hill, "The effectiveness of training to improve person perception accuracy: a meta-analysis," *Basic and Applied Social Psychology*, vol. 34, no. 6, pp. 483–498, 2012.
[52] L. J. Pfiffner and K. McBurnett, "Social skills training with parent generalization: Treatment effects for children with attention deficit disorder." *Journal of Consulting and Clinical Psychology*, vol. 65, no. 5, p. 749, 1997.
[53] R. F. Baumeister, K. D. Vohs, and D. C. Funder, "Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?" *Perspectives on Psychological Science*, vol. 2, no. 4, pp. 396–403, 2007.
[54] K. Dickersin, "The existence of publication bias and risk factors for its occurrence," *Jama*, vol. 263, no. 10, pp. 1385–1389, 1990.
[55] M. Callaham, R. L. Wears, and E. Weber, "Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals," *Jama*, vol. 287, no. 21, pp. 2847–2850, 2002.
[56] C. B. Begg, "Publication bias," *The handbook of research synthesis*, vol. 25, pp. 299–409, 1994.