The Efficacy of Head-Mounted-Display Virtual Reality Intervention to Improve Life Skills of Individuals with Autism Spectrum Disorders: A Systematic Review

Nikki M Skjoldborg, Patrick K Bender, Kristine M Jensen de López

Center for Developmental & Applied Psychological Science, Department of Communication and Psychology, Aalborg University, Aalborg, Denmark

Correspondence: Kristine M Jensen de López, Email kristine@ikp.aau.dk

Abstract: Challenges in life skills in individuals with autism spectrum disorders (ASD) are associated with dependency on others and increased isolation from peers. In recent years, interventions using virtual reality (VR) technology have been proposed to improve life skills in ASD populations. This systematic review seeks to evaluate the efficacy of employing VR interventions mediated via head-mounted displays (HMD) for the improvement of life skills in individuals with ASD. Several databases were searched and a narrative synthesis was conducted to examine the findings of the included studies. Eight studies including a total of 58 participants were deemed relevant for this systematic review. The methodological quality of the included studies was assessed via the use of critical appraisal tools. Results were generally positive, with one study reporting statistically significant results, and one study not reporting any change in abilities. The remaining six studies reported varying degrees of life skill improvement. The studies were characterized by methodological issues, such as very low sample sizes. The findings of this systematic review indicate some potential for HMD VR interventions in the improvement of life skills in individuals with ASD. However, this review also highlights the current lack of methodologically strong study designs, which prohibits any firm conclusions. Findings are discussed regarding methodological recommendations for further research as well as practical implications for life skills interventions for individuals with ASD.

Keywords: virtual reality, autism spectrum disorder, life skills, social skills, head-mounted displays, intervention

Introduction
Virtual reality (VR) technology offers a controllable and realistic environment in which customized scenarios can be adjusted to meet the needs of the individual user. For this reason, VR has been used in a wide variety of interventions, ranging from the treatment of phobias to interventions aimed at improving false belief understanding.\(^1\) VR interventions have also been used in an array of interventions aimed at individuals with autism spectrum disorders (ASD), such as improving the understanding of social rules and norms.\(^1\) Even though VR interventions constitute a relatively new field of research, current studies already emphasize the value of such interventions by referencing previous studies and their positive results.\(^2\) This indicates a need to systematically investigate the efficacy of these types of interventions.

Autism Spectrum Disorder and Life Skills
ASD is a life-long neurodevelopmental condition with a large amount of variance within the spectrum and, as such, individuals with ASD can have different developmental challenges and characteristics.\(^3\) One key characteristic of individuals with ASD are weaknesses in social skills (eg, engaging in social interaction), and practical skills (eg, handling money).\(^3\) Thus, individuals with ASD often show weaknesses or challenges with regard to life skills, which can be defined as a combination of social and practical skills, that allow an individual to successfully interact with their environment in everyday life.\(^4\) Life skill challenges are associated with great economical and emotional dependence on...
family, heightened reliance on professionals, and isolation from peers. Thus, improving the life skills of individuals with ASD is of vital importance.

Currently, a number of interventions aimed at improving life skills in individuals with ASD exist. One approach is Applied Behaviour Analysis (ABA), which involves the training of routines and abilities related to the individual’s difficulties. Generally, in ABA, the tasks related to a specific challenge are broken up into smaller segments, which are easier to address for the individual and can be varied in complexity. Another approach used to improve social skills in individuals with ASD are social stories. These social stories are framed in such a way that they can help the individual understand the social situation in question, as well as the relevant contextual cues. Given the large amount of variability in the ASD population, Lahiri suggests that interventions need to be somewhat tailored to the individual for them to be effective.

When considering interventions to address life skill challenges, one should also be aware that individuals with ASD often experience varying degrees of increased sensitivity to sensory inputs. This increased sensitivity to sensory stimuli may in some cases impede their learning process in natural settings by distracting the individual and may therefore limit their ability to engage in learning situations in everyday life.

**Virtual Reality Interventions**

Psychological interventions mediated via technology are not a new phenomenon. As early as the mid-1990's, instant messaging or email therapy was available. The arrival of virtual reality interventions has brought with it a large array of possibilities for new types of interventions. Following Sherman and Craig, VR can be defined as:

A medium composed of interactive computer simulations that sense the participant’s position and actions and replace or augment the feedback to one or more senses, giving the feeling of being mentally immersed or present in the simulation (a virtual world).

In the last 20 years, a variety of different virtual reality systems have been developed to treat phobias, anxiety disorders, and post-traumatic stress disorders, amongst others. A newer addition to the field of VR interventions, are interventions aimed at helping the ASD population. According to Gal and et al, the VR approach to interventions for individuals with ASD allows for a number of external sensory stimuli, which may be experienced as arduous, to be removed or ignored. Additionally, the nature of the intervention, where the task is often clearly defined and consistent, and where there are fewer social demands than in real-life situations, seems to help individuals with ASD to focus and concentrate on the task at hand.

**Head-Mounted Displays in VR Interventions**

During the last years, a number of reviews have addressed the use of VR systems to improve skills in individuals with ASD. However, these reviews have investigated the general use of a variety of VR systems, which include a range of technologies such as computer screens, tablets, and phones to display the virtual environment. Compared to other VR technologies, Head-mounted displays (HMDs) allow for the most immersive and realistic VR experience and therefore minimize the gap between the virtual training environment and the real life environments where life skills are needed. This is thought to maximize the transfer of skills from the virtual to the real world. Compared to using other screen technology, HMD technology also allows for the most effective manipulation of sensory stimuli since there are no external visual stimuli apart from the simulation when using this kind of technology. Therefore, this systematic review seeks to evaluate the efficacy of VR interventions using HMDs. HMDs are a type of display that utilises one or two small screens, which are held in a headset and positioned in front of the user’s eyes, to create the illusion of 3-dimensional images. The headset tracks the user's head-movements and allows the program to change the point-of-view presented in the HMD to match that, which the user should see. Often, the system is combined with other technologies, such as full motion tracking (allowing the user to move around the virtual environment), controllers (to interact with the virtual environment), and surround sound (to include the sense of hearing). These additions further increase the sense of immersion experienced by the user.

As mentioned, HMD technology differs from receiving VR interventions via eg, a computer screen in important ways. The virtual environment presented on a traditional computer screen and in a head-mounted display may be the
same. However, the user experiences a higher sense of immersion when the environment is presented utilizing a head-mounted display and there are no additional visual inputs besides the simulation displayed.14 As such, HMDs have an advantage over more traditional technologies, as they produce a more realistic experience for the user. Following Miller and Bugnariu, this more realistic experience may increase the chance of transfer of the learned life skills from the VR intervention to the real world.13 Given these potential advantages of HMD VR, over other types of VR technology, we focus exclusively on the potential benefits and research findings regarding HMDs in this systematic review.

The Potential of HMD VR Interventions

As outlined, VR interventions have been proposed to have several advantages over traditional interventions. One advantage is the increased security that comes with this type of life skill intervention, as the virtual environment constitutes a controlled, and thus, safer environment for the individual to practice certain skills.15 Another potential of HMD VR interventions is the fact that they allow for a great deal of control concerning the number and intensity of stimuli that the user is subjected to. This aspect is relevant for individuals with ASD, as they often experience sensory overload, ie, a sensitivity to stimuli, which individuals without ASD would not be as sensitive to.16 Thus, being able to individually tailor the stimuli an individual is presented with during the intervention may avoid potential sensory overload and distress by partly or fully isolating the individual from certain stimuli of the surrounding environment, which may improve learning outcomes for individuals with ASD. In addition, the tasks within the virtual environment and their level of complexity can be scaled to fit the needs of the individual, which could be particularly useful in the initial stages of life skill training and is fully compatible with an Applied Behaviour Analysis approach.1

Furthermore, in comparison to practising social skills by interacting with peers, the virtual environment may offer a conceptually safer space in the sense that there are no real-life repercussions for not mastering a given skill initially.1 Individuals with ASD may isolate themselves from social interactions out of fear of negative consequences, which in turn may hinder their social skill development further. Thus, as pointed out by Lahiri, the virtual environment may offer a space in which the individual can test and improve their own social skills without fear of negative consequences.1

Method

We conducted a systematic review to analyse the efficacy of virtual reality interventions, utilizing HMD technology, in improving the life skills of individuals with ASD, as well as the quality of the studies that have investigated this issue.

To identify relevant studies, four online databases were searched for relevant studies (PsychNet, ACM digital library, Scopus, and ProQuest) without constraints regarding time of publication. A number of relevant search terms were identified utilizing the population, intervention, comparison, and outcome (PICO) model.17 The search terms are presented in Table 1.

Search terms regarding specific comparison groups were not specified as all types of comparisons were included in this review.

Inclusion and Exclusion Criteria

The following inclusion and exclusion criteria were applied in the screening of studies.

Population: the population receiving intervention in the included studies were individuals diagnosed with ASD. For studies including a comparison group, this group could consist of individuals without ASD. As mentioned in the introduction, ASD is a life-long condition. We therefore included all age groups. Intervention: the types of intervention included in this systematic review were limited to virtual reality interventions that utilized head-mounted displays and were aimed at improving life skills. Based on the previously mentioned definition of life skills, studies were thus included if they addressed either social or practical skills, that are necessary for individuals to successfully interact with their environment in everyday life.4 Interventions that utilized other VR technologies were excluded. Interventions aimed at improving skills other than life skills were also excluded. Design: All types of design were included in the systematic review, including single case designs and designs with and without control groups. Studies which were not empirical in nature, such as literature reviews, were excluded. Outcome: Outcome measures related to life skills were included in the analysis and the narrative synthesis.
The initial searches yielded a total of 1037 results once duplicates were removed. The online reference management tool RefWorks was used to keep track of studies throughout the screening process, and to remove duplicates. See Table 1 for an overview of the search strings used within each database, as well as the number of results found in each database. See Figure 1 for a PRISMA flow chart of the screening process. 1037 papers were screened on title and abstract by the first author, and this resulted in the inclusion of 58 papers for full-text screening. The first author was responsible for the full-text screening of these papers, and she consulted the last author regarding unclearness and discussion about inclusion. A total of eight studies were deemed relevant for this systematic review.

### Table 1: Overview of Searches Conducted for This Review

<table>
<thead>
<tr>
<th>Source</th>
<th>Search Query</th>
<th>Number of Results</th>
<th>Date of Performed Search</th>
</tr>
</thead>
<tbody>
<tr>
<td>PsychNet</td>
<td>((Any Field: (&quot;Life skills&quot;)) OR (Any Field: (Competence)) OR (Any Field: (&quot;Self-Care Skills&quot;)) OR (Any Field: (&quot;Socioemotional functioning&quot;)) OR (Any Field: (Ability*)) OR (Any Field: (Skill*))) AND (Any Field: (&quot;Virtual reality intervention&quot;)) OR (Any Field: (&quot;Virtual Reality&quot;)) OR (Any Field: (&quot;VR&quot;) OR (Any Field: (&quot;Head mounted display&quot;)) OR (Any Field: (&quot;Head-mounted display&quot;)) OR (Any Field: (&quot;Virtual reality training&quot;)) AND (Any Field: (&quot;Autism Spectrum Disorder&quot;)) OR (Any Field: (Autism)) OR (Any Field: (Autistic*)) OR (Any Field: (ASD)) OR (Any Field: (&quot;Autistic traits&quot;)))</td>
<td>158</td>
<td>February 2021 (updated April 2022)</td>
</tr>
<tr>
<td>ProQuest</td>
<td>(noft(&quot;Autism spectrum disorder&quot;) OR noft(Autism) OR noft(Autist*) OR noft(ASD) OR noft (&quot;Autistic traits&quot;) AND (noft(&quot;Virtual Reality intervention&quot;) OR noft(&quot;Virtual Reality&quot;) OR noft(&quot;Virtual Reality&quot;) OR noft(&quot;VR&quot;) OR noft(&quot;Head mounted display&quot;) OR noft(&quot;Head-mounted display&quot;) OR noft (&quot;Virtual reality training&quot;) AND (noft(&quot;Life skills&quot;) OR noft(Competence) OR noft(&quot;Self-Care Skills&quot;) OR noft(&quot;Socioemotional functioning&quot;) OR noft(“Ability&quot;) OR noft(&quot;Skill&quot;))))</td>
<td>312</td>
<td>February 2021 (updated April 2022)</td>
</tr>
</tbody>
</table>

### Search Results

The initial searches yielded a total of 1037 results once duplicates were removed. The online reference management tool RefWorks was used to keep track of studies throughout the screening process, and to remove duplicates. See Table 1 for an overview of the search strings used within each database, as well as the number of results found in each database. See Figure 1 for a PRISMA flow chart of the screening process. 1037 papers were screened on title and abstract by the first author, and this resulted in the inclusion of 58 papers for full-text screening. The first author was responsible for the full-text screening of these papers, and she consulted the last author regarding unclearness and discussion about inclusion. A total of eight studies were deemed relevant for this systematic review.

### Synthesis and Critical Appraisals

The 8 included studies were analysed using a narrative synthesis given the inclusion of a variety of study designs. The aim of a narrative synthesis is to summarize the findings of the identified studies. Also, in order to assess the methodological quality and the risk of bias in the identified studies, a number of quality assessment tools were utilized to evaluate the different study designs utilized in the included studies. Quantitative studies were rated using the Quality Assessment Tool for Quantitative Studies (QATQS) developed by the Effective Public Health Practice Project. Single subject designs were rated using The Single Case Experimental Design (SCED) scale, developed by Tate et al. Finally, single case study designs were evaluated using the Critical Appraisal of Case Reports (CACR) tool. Critical appraisals
were conducted on each study and the results were used to evaluate the potential risks of bias when considering the findings.

**Results**

**Overview of the Included Studies**

The countries of origin of the studies were USA (3), Denmark (2), Taiwan (2) and Spain (1). Table 2 provides an overview of the eight included studies. Consistent with the novelty of the technology, the included studies consisted of recently published papers between 2015 and 2020. All of the studies dealt with a population of solely individuals with previously diagnosed ASD. The number of participants included varied but was generally low with only three studies including 10 or more participants.\(^\text{21-23}\) The five other studies varied between two and nine participants. Participants were mostly children or teenagers, with only one study examining young adults aged 18–22.\(^\text{2}\) Gender-wise the studies were either all male or included a single or two females. Few studies considered the severity of participants’ ASD symptoms; Herrero and Lorenzo included participants with level one and two ASD with regard to support need, according to the DSM-V.\(^\text{21}\) Dixon et al presented participant scores on the Pervasive Developmental Disorder Behaviour Inventory (PDDBI), on which the participants’ scores were distributed somewhat close to the average for individuals with ASD.
with a few scores slightly elevated. Yeh and Meng and Cheng and et al reported participants’ IQ scores, ranging from 70 to 105, and 80 to 84 for full-scale IQ, respectively.

### Study Designs

Table 3 shows the research designs employed in the eight studies. Two of the studies included a control group. One of these was a randomized controlled trial, and the other a controlled clinical trial. Two of the studies utilized a cohort (one group) design, two studies were single subject designs, one study was a case study, and one study used a time series approach, where measures were recorded at all training sessions, and the post-intervention measure was recorded during the last intervention session. None of the included studies conducted follow-up assessments of the participants. However, Cheng et al conducted the maintenance session/post measures during the 20 days after the last intervention session and Ravindran et al conducted the post-measure 4 weeks after the last intervention session.

### Overall Findings of the Included Studies

As presented in Table 4 most of the studies showed generally positive results, demonstrating some level of improvement of the life skill in question. Improvement in this case refers to an increased score on the measures utilized, or a perceived improvement observed by the researchers or others in the abilities of participants. As such, in some cases improvement indicates that the abilities of study participants moved closer to those of individuals without ASD, in other cases it entails an improvement on a specific skill, such as paying the correct amount of money for groceries. Only one of the studies showed no improvement in life skills, which may be attributed to a ceiling effect. Two of the eight studies conducted inferential statistical analyses. Yeh & Meng reported p-values that showed significant levels of change, whereas the

---

**Table 2 Study Characteristics of Empirical Studies Investigating the Use of HMD in Life Skill Improvements for Individuals with ASD**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Country</th>
<th>Aim</th>
<th>Participants</th>
<th>Life Skill</th>
<th>HMD Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjorlu &amp; Serafin (2019)</td>
<td>Denmark</td>
<td>To investigate the effectiveness and feasibility of VR for the improvement of money skills</td>
<td>5 18-22 All M</td>
<td>Money skills</td>
<td>HTC Vive</td>
</tr>
<tr>
<td>Adorlu et al (2017)</td>
<td>Denmark</td>
<td>To investigate whether VR is effective and feasible for training of daily living skills (DLS) for individuals diagnosed with AD</td>
<td>9 12-15 8 M 1 F</td>
<td>Shopping</td>
<td>HTC Vive</td>
</tr>
<tr>
<td>Beach &amp; Wende (2016)</td>
<td>USA</td>
<td>To investigate whether students with ASD could improve their social interaction skills by practicing in VR</td>
<td>2 15-18 All M</td>
<td>Social interaction skills</td>
<td>Oculus Rift</td>
</tr>
<tr>
<td>Cheng et al (2015)</td>
<td>Taiwan</td>
<td>To investigate the effectiveness of a VR social understanding system in improving the social understanding and skills in children with ASD</td>
<td>3 10-13 All M</td>
<td>Social understanding and skills</td>
<td>I-Glasses PC 3D Pro</td>
</tr>
<tr>
<td>Dixon et al (2020)</td>
<td>USA</td>
<td>To assess the effectiveness of VR training for improvement of pedestrian skills in children with ASD</td>
<td>3 4-10 All M</td>
<td>Pedestrian skills</td>
<td>Oculus Rift</td>
</tr>
<tr>
<td>Herrero &amp; Lorenzo (2020)</td>
<td>Spain</td>
<td>To investigate whether VR could be satisfactory as an educational tool in improving communication skills and problem solving for ASD children</td>
<td>14 8-15 13 M 1 F</td>
<td>Communication and problem solving</td>
<td>Oculus Rift</td>
</tr>
<tr>
<td>Ravindran et al (2019)</td>
<td>USA</td>
<td>To explore the feasibility of the use of a VR joint attention program to improve joint attention skills in children with ASD</td>
<td>12 9-16 10 M 2 F</td>
<td>Joint attention</td>
<td>Smartphone and Google Cardboard</td>
</tr>
<tr>
<td>Yeh &amp; Meng (2020)</td>
<td>Taiwan</td>
<td>To investigate the effectiveness of virtual reality teaching methods for social skills interventions for students</td>
<td>10 7-14 9 M 1 F</td>
<td>Social skills</td>
<td>HTC Vive</td>
</tr>
</tbody>
</table>

---

Abbreviations: M, male; F, female.

---

Neuropsychiatric Disease and Treatment 2022:18

2300

https://doi.org/10.2147/NDT.S331990

DovePress
results of the study by Adjorlu et al\textsuperscript{26} did not show significant improvement as a result of the intervention. The study by Cheng et al\textsuperscript{25} was based on a single subject design and hence the authors reported the percentage of non-overlapping data (PND),\textsuperscript{27} which indicated that the intervention had been effective with 100\% of non-overlapping data points. Herrero and Lorenzo utilized probabilistic analysis to calculate changes in life skills.\textsuperscript{21} This indicated an improvement, which,
according to the authors, was significant. The remaining four studies did not report inferential statistical analyses or PND calculations but reported solely descriptive statistics.2,12,22,24

Narrative Synthesis
In this section the findings of the included studies are presented and examined via a narrative synthesis.28 In the synthesis we illuminate relevant aspects of the studies in order to examine the efficacy of utilizing virtual reality interventions with head-mounted displays to improve the life skills of individuals with autism spectrum disorders. The definition of life skills utilized in this systematic review allows the inclusion of a broad variety of abilities and competencies. In the following, we present our analysis of the included studies based on the two components of life skills that were suggested by Hodge et al4 We first present results for the studies addressing practical skills followed by results for the studies addressing social skills.

Studies of Practical Skills
Three of the included studies investigated the efficacy of VR intervention on individuals’ practical skills. McDougle and Accordino refer to practical skills as the skills needed to function independently in everyday life, such as riding the bus.3 In the study by Adjorlu et al, the authors sought to evaluate the feasibility and effectiveness of VR to train shopping skills in individuals with ASD.26 To investigate this, they recruited nine students with ASD from a school in Denmark, five of which constituted a passive control group. The remaining four students received seven sessions of VR training administered by a teacher, over 10 days. Training consisted of the participants retrieving four items as specified on a shopping list from a virtual supermarket. Outcome measures were recorded before and after the training sessions, and were conducted similarly to the VR training, but in a real supermarket. The treatment group reached 100% effectiveness in both the pre and post assessments. Thus, no change in skill was detected. The control group decreased slightly in their performance in the post measure, from 97% to 91%. The authors reported a small difference between pre and post assessments in terms of task completion time and participants’ confidence for the treatment group. They argue that the continuous training of the skill in VR may have helped the treatment group maintain their skills.

Adjorlu and Serafin investigated whether VR is a feasible and effective way of improving money skills in individuals with ASD.2 They recruited five students from a special education school in Denmark and administered a VR intervention over a period of two weeks through a series of five sessions. Sessions consisted of three levels of training, with increasing complexity, all centred around paying the correct amount of virtual money. The outcome measures were measured in real life, pre and post intervention, by presenting students with price cards and asking them to pay the correct amount of real money as indicated on the price card. The authors found that four out of the five participants improved their money skills, although with some degree of variability.

Dixon et al investigated the effectiveness of a VR environment in teaching street-crossing skills to children with ASD.24 Three participants were recruited from a clinic in the USA and were given between 13 and 16 training sessions. Training sessions included several trials, which varied in length, and in which participants would be placed in the virtual environment of a street. The environment had been created using recordings from the real world, and participants answered questions regarding whether it would be safe to cross the street. Baseline and post measurements were conducted in the real environment, where participants were asked whether it was safe to cross the street. All participants eventually reached mastery criteria, which was defined as 100% correct answers in three consecutive session, for the natural environment, however the authors reported that this was only achieved once longer trials were utilized.

As described above, two of the studies that examined practical skills showed improvements in real-life outcome measures.2,24 However, only Adjorlu et al utilized inferential statistics, and they did not find a significant difference between the control and treatment groups.26 Due to a pre-intervention ceiling effect, they were not able to report any potential improvement in the abilities of participants. The ceiling effect might indicate that the task was not sensitive to the challenges of the participants, or that individuals with ASD do not have difficulties with this type of task.
Outcome Measures
All three studies utilized a real-life outcome measure when evaluating the effectiveness of the applied intervention. However, the similarity of the outcome measures to daily life situations varied somewhat. Two of the studies assessed outcome measures in natural settings, whereas one study created a semi-realistic situation, in which part of the context was natural (i.e., using real money, but in a classroom). Thus, all three studies aimed to assess their outcome measures in an environment which closely mimics that in which participants would be required to use the various practical skills in real life.

Design of the Interventions
All three practical skill studies attempted to create virtual environments similar to the real-life environments, in which the outcome measures were assessed. The degree of realism varied somewhat between the three studies with the VR environment utilized by Dixon et al being the most realistic, as they used real recordings of a street, thus creating a VR environment that closely resembles the real world.

As well as creating a virtual environment similar to that of the real world, all three studies created scenarios, which followed a sequence of events similar to the real world. Dixon et al combined this with asking additional questions during training. Adjorlu and Serafin differed slightly by utilizing different scenarios in the intervention compared to the post-measure. During the intervention, supermarket items with a price attached were used in the final level of intervention. When assessing the outcome measure, they relied solely on presenting price cards, which the participants had to match money to. Adjorlu and Serafin and Adjorlu et al both allowed, and relied on, the participants interacting with the environment as part of the training. In contrast, Dixon et al allowed the participant to look around, but not to otherwise interact with the environment.

Studies of Social Skills
Social skills may be understood as the life skills that concern themselves with the social abilities needed to function in everyday life. A total of five of the included studies investigated whether VR interventions could improve social skills.

Beach and Wendt investigated whether a VR environment could be used to improve social interaction skills in individuals with ASD. Two participants were recruited from a group of students attending a summer camp in the US that aimed to allow students to practice their social skills, and the study was conducted while the students were attending the summer camp. Participants were given tailored training in VR, which was adjusted to meet their specific social challenges. These scenarios were later re-enacted by camp teachers, and a change in both participants’ abilities was noted. This study relied on observations from both the teachers and the researchers as well as the monitoring of a specific trait that needed improvement. Generally, the authors reported positive results, with both participants improving their specific trait in need of improvement, as well as their general social interaction abilities, as observed by the camp teachers.

Cheng et al investigated the effectiveness of a VR social understanding system in improving the social understanding and skills of three children with ASD. Participants were recruited from a special education school in Taiwan and were subjected to five intervention sessions each. These training sessions consisted of virtual classroom and bus stop environments, in which participants were asked problem-based questions related to non-verbal communication, social initiation, and social cognition. Baseline and maintenance measures were observed by presenting participants with a scenario and asking them 12 questions from the social behaviour scale developed by the authors. The authors report that data points were 100% non-overlapping, which indicates that the treatment was effective in the improvement of the skills examined.

Herrero and Lorenzo evaluated the design and application of a VR system for the improvement of social and emotional skills of individuals with ASD. 14 participants were recruited and divided equally into a treatment and a control group. The treatment group was presented with 10 training sessions, in which participants could practice both verbal and non-verbal communication skills. This took place in a virtual classroom and in a garden environment. Outcome measures were assessed within a VR setting and within a real life setting and included a series of questionnaires given to the parents/tutors and teachers before and after treatment sessions. Additionally, session questionnaires and
automatic registration of participant performance in VR was used to measure participants’ performance in the VR setting. To evaluate the findings in relation to the effectiveness of the treatment, the authors utilized a probabilistic approach rather than traditional statistical analysis. Findings indicated a significant improvement in all four areas investigated: social and emotional reciprocity, non-verbal communication, inflexibility to changes, and stereotypes and sensorial reactivity.

Ravindran et al evaluated the efficacy of both a VR joint attention module in improving joint attention skills in the ASD population. The 12 participants were recruited from a local special education school in the US and received 14 intervention sessions. Sessions consisted of a safari themed virtual environment, in which the participant would practice joint attention with a virtual avatar following the set goals specified by different learning cards. A joint attention measure, a behavioural play-based assessment, was conducted pre and post intervention. Findings indicate that there was an improvement in at least one area of joint attention skills for 10 out of the 12 participants. The authors did not conduct inferential statistics and were therefore not able to comment on the statistical significance of their results.

The study by Yeh and Meng represents the only included study that relied solely on a virtual reality outcome measure. In the study, the authors evaluated the effectiveness of VR in teaching social skills to students with ASD. 10 participants were recruited through a monthly magazine published by an autism association and were each given eight courses in the VR intervention. The VR intervention consisted of a sequence of scenarios all part of a virtual tour created to resemble an art museum with a school class. The scenarios allowed for training different social skills ranging from talking to classmates to interacting with ticket vendors. The outcome measure consisted of an evaluation sheet. Outcome measures were recorded during all eight sessions, and statistical significance levels (p < 0.05) indicated significant improvement in all four areas: answer speed, conversation and expression, the completeness of sentence structure, and conversation etiquette.

The five studies that investigated social skills all reported varying degrees of efficacy. However, only Yeh & Meng utilized inferential statistics and reported statistically significant results for their outcome measure. The remaining four studies all reported some increase in the specific skill under investigation. Cheng et al and Herrero and Lorenzo reported, based on their analysis of the results, that the intervention had been highly effective in improving the respective life skills. The final two studies reported some improvement of specific social skills.

Outcome Measures
The studies investigating social skills measured their outcome variables differently. One study used a questionnaire presented to parents/tutors and educators, while other studies attempted to create a situation in which they could observe the behaviour of participants in a laboratory or in a natural environment. One study asked participants questions based on a chosen scenario. Yeh and Meng exclusively evaluated participants’ performances in the virtual environment.

As described, most of the studies that sought to improve social skills did not assess the respective skills in a natural, real-life setting following the VR intervention. Beach and Wendt observed students in a somewhat natural environment, but this was still an uncommon setting for their participants, as it took place at a summer camp.

Design of the Interventions
The types of interventions employed in the social skills studies varied greatly in content. Beach and Wendt tailored the intervention to the needs of their participants, thereby providing a participant-centred intervention. In contrast, most of the other studies followed a set protocol for the intervention, which barely varied between participants. Another difference between the studies is the level of interaction that was possible between the participant and the virtual environment. While all of the social skills studies allowed for participants’ actions to influence the virtual environment in varying degrees, the impact this had on the environment varied.

When constructing the intervention environments, all of the social skills studies attempted to create a situation similar to the one in which participants might need the skill. Three of the studies created everyday situations, like a classroom. Two studies created environments that the students might encounter occasionally, but probably not regularly in everyday
life, ie, a zoo and an art museum.\textsuperscript{22,23} Social interactions within the virtual environments varied across the five studies, with three studies reporting that participants were able to interact with virtual human characters in the virtual environment.\textsuperscript{12,21,23} Two of these employed a combination of spoken conversations and non-verbal cues.\textsuperscript{12,21} In the remaining two studies, participants were prompted with questions, but this was done by a person outside of the virtual environment, and as such did not entail interaction with a virtual character.\textsuperscript{22,25}

**Summary of Efficacy for All Included Studies**

To evaluate the efficacy of VR interventions using head-mounted displays, we analysed what characterized the studies showing statistically significant improvements. Three of the studies that either reported statistically significant improvement\textsuperscript{23} or reported their intervention to be highly effective were based on third party or VR-based evaluations.\textsuperscript{21,25} Both Cheng et al\textsuperscript{25} and Herrero and Lorenzo\textsuperscript{21} utilized other-reporting (a student, parent, or researcher) as opposed to a self-reporting questionnaire as outcome measures. Yeh and Meng assessed the actions of participants using evaluation sheets and observing the participants in the virtual environment.\textsuperscript{23} As such, none of these studies relied on actually testing participants’ skills in real-life settings.

The remaining five studies (apart from Adjorlu et al\textsuperscript{26}) showed some improvement in outcome measures following interventions. Interestingly, these five studies all assessed outcomes in natural or semi-natural environments.\textsuperscript{2,12,22,24} Based on this, an important pattern arises. Studies that assessed outcomes in natural settings, in which participants would need the respective skills, reported less improvement than studies that utilized a non-natural setting for the evaluation of participants’ life skill improvement. It is also important to note that none of the studies that assessed outcomes in a natural environment conducted inferential statistics and only reported subjective improvements, making it difficult to generalize the reported improvements.

**Critical Appraisal**

To evaluate the methodological soundness and quality of the included studies, three different quality assessment tools were utilized, addressing the various study design. The five quantitative studies were rated using the Quality Assessment Tool for Quantitative Studies (QATQS).\textsuperscript{18} The two studies utilizing a single subject design were rated using The Single Case Experimental Design (SCED) scale.\textsuperscript{19} The single case study was evaluated using the Critical Appraisal of Case Reports (CACR) tool.\textsuperscript{20} Table 5 provides an overview of all ratings.

The five quantitative studies varied somewhat regarding the quality appraisal, with three achieving a moderate rating,\textsuperscript{2,22,23} one a weak rating,\textsuperscript{26} and one a strong rating.\textsuperscript{21} Most of these studies shared similar weaknesses, such as not describing the validity and reliability of their data collection tools (section E), with Herrero & Lorenzo\textsuperscript{21} being the only exception. Additionally, blinding was poorly described in most, and somewhat lacking in all of the studies (section D).

The two single subject design studies, which were rated using the SCED scale, scored in the higher end of the scale, with similar weaknesses. Both studies lacked statistical analyses. Cheng et al\textsuperscript{25} did analyse the percentage of non-overlapping data through comparison of the data points across study phases. This makes it difficult to evaluate the significance and magnitude of the reported improvements in this study. Additionally, Cheng et al\textsuperscript{25} did not utilize independent assessors, and Dixon et al\textsuperscript{24} did not clarify whether they did or not.

The included case study\textsuperscript{12} was evaluated using the CACR tool, and obtained a score on the higher end (8/12). This study lacked methodological descriptions, eg clear descriptions of the data collection methods.

Overall, the included studies are characterised by moderate to low quality with a lack of statistical analyses and active control groups. Two studies included a control group,\textsuperscript{21,26} however, these acted as passive control groups. The choice to not employ a control group puts a study at risk of bias by potentially overestimating the effectiveness of the intervention, as it makes it difficult to assess the efficacy of the intervention compared to other types of interventions. It also makes it difficult to determine whether observed improvements are the result of the intervention or other factors influencing participants, such as simply participating in the study.\textsuperscript{17}

Another general methodological weakness across the studies was either a lack of blinding, or a failure to report blinding procedures. The only study to report blinding was conducted by Adjorlu et al.\textsuperscript{26} The remaining studies either did
not blind assessors,\(^2,^{12,25}\) or did not report on this aspect. None of the included studies reported whether participants were aware of the research questions at the time of participation.

**Discussion**

The present systematic review assessed the efficacy of interventions designed in virtual reality and mediated via head-mounted displays to improve the life skills of individuals with autism spectrum disorder. We also evaluated the quality of the included studies. The eight studies showed some promising results with three studies reporting significant or substantial improvement and five studies reporting some improvement following interventions. However, whether HMD VR interventions reliably improve the life skills of individuals with autism spectrum disorder has yet to be determined. The findings and quality of the included studies are somewhat mixed, with few studies utilizing inferential statistics or control groups. Furthermore, studies that assessed participants’ life skill improvements in natural settings, where the skills are requested reported less improvement than studies that chose to rely on third party reports or assessments within the virtual environment. Therefore, the current knowledge does not allow to draw firm conclusions,
and several methodological issues need to be addressed in future research before we can confidently answer the question addressed in this review.

Generally, studies were characterized by lack of statistical analysis to evaluate the significance of their results. This makes it difficult to evaluate whether improvements reported in individual studies are indicative of actual changes in life skills or merely the effect of random variation within samples. Additionally, many of the studies were carried out by researchers evaluating interventions developed by the authors themselves. This is not inherently a problem, but it may be a potential source of bias. In the future, it will be important for VR interventions to be evaluated by independent research groups.

Another methodological concern is the representativeness of the participants included in the studies. The term external validity refers to the extent to which one can generalize the findings of a study to a different context. Firstly, only a few of the included studies addressed the level of severity of ASD experienced by participants. This is a potential threat to the external validity of the studies regarding generalisability. This is especially of concern considering the well acknowledged vast differences in abilities and difficulties between individuals with ASD. Further, the included studies did generally not follow the gender-ratio found amongst individuals with ASD, which is estimated to be one female to every three or four males. As a result, females were generally underrepresented in the studies, and therefore, the findings become difficult to generalize to the female ASD population. These issues, combined with small sample sizes, which were not justified statistically, call into question whether we can draw any firm conclusions concerning the ASD population as a whole, based on the studies included in this review.

The internal validity of a study can be described as the accurateness of any conclusions about a potentially causal relationship between two variables: in this case, the HMD VR intervention and a change in life skills. A variety of factors influence the internal validity of a study. Especially relevant for this systematic review is the lack of an active control group and follow-up assessments.

The general lack of a control group influences the internal validity of the included studies due to the increased risk of bias that comes with it. When a study does not include some type of control group it is impossible to know what might have happened, had the intervention not taken place. Two of the included studies utilized control groups, however these were passive control groups only. The lack of active control groups receiving standard interventions makes it impossible to compare HMD VR interventions to the effects of more traditional or treatment-as-usual interventions, which is important when considering substituting tried approaches with novel ones such as VR interventions.

A further methodological concern is that none of the included studies conducted follow-up assessments; however, some conducted the post-assessments after a set amount of time had passed. For instance, Cheng et al put the participants through a maintenance phase of 20 days, during which the last post-intervention assessments were conducted.

The included studies were also generally characterized by a lack of blinding of both assessors and participants. Blinding is useful in studies as it decreases the risk of bias caused by conscious or unconscious influences upon the research. However, given the relative infancy of this line of research, it may be argued that blinded studies are somewhat premature at this point in time. Once a more substantial evidence base for the effect of HMD VR interventions on life skills in individuals with ASD has been accumulated appropriate blinding should be implemented to decrease the risk of overestimating intervention effects.

Finally, one of the major potential advantages of VR interventions utilizing HMDs is the ability to adjust the amount of sensory input that the user is presented with. None of the included studies described adjusting the sensory input to the need of their participants, and as such it is not possible to conclude on whether or not such a more individualized approach may enhance the outcomes of interventions aimed at improving life skills in individuals with ASD. On a similar note, only one study took advantage of the ability to tailor the virtual environment and training to the specific difficulties of individual participants. Once again, this means that studies did not generally capitalize on another potential benefit of VR interventions; the ability to customize the scenario to the individual. Additionally only one study utilized scenarios of increasing complexity.

As such, most of the intervention designs remained at the same difficulty level throughout the intervention program.
Ecological Validity
An additional important methodological issue relevant to the discussion of the efficacy of HMD VR interventions for the improvement of life skills in individuals with ASD, is ecological validity. Ecological validity refers to the degree to which the context and procedures, as well as measures, of a given study are artificial, compared to the real world. One aspect relevant to the ecological validity of studies are the outcome measures utilised. This systematic review investigated the efficacy of HMD VR interventions, and thus, the ecological validity of outcome measures is important to draw conclusions regarding the ability of an intervention to improve the actual life skills of participants. The ecological validity of outcome measures employed in the included studies varied greatly. As described previously, the three studies reporting the most solid evidence of efficacy did not assess participants’ life skills in a natural environment. Rather, these studies employed questionnaires or observed participant skills in a virtual environment. This methodology lacks ecological validity and constrains the possibility of drawing any conclusion concerning participants’ real-life skills. Any improvement may be due to participants’ knowledge of correct responses (without actually being able to implement such behaviour in real life), or due to increased abilities in the virtual environment (again, without being able to implement such behaviour in real life circumstances). This suggests that the use of questionnaires to assess life skills, both practical and social, may be problematic and lack validity as outcomes measures. As such, the ecological validity of this type of outcome measure is somewhat weak, and further investigation may be needed to confirm the effectiveness of the administered interventions using more ecologically valid measures. Thus, it may be argued that the studies that did not use an ecologically valid outcome measure risk overestimating the effect of the intervention on the actual life skills of participants. In fact, considering that the dependent variable of the included studies was life skills stresses the importance of striving for a minimum level of ecological validity in the outcome measures.

Strengths and Limitations
This systematic review aimed to evaluate the efficacy of HMD VR interventions in improving the life skills of the ASD population. Since this is a relatively new field of research, it was deemed important to evaluate the status of the field, especially given that some studies present this technology as a viable alternative to traditional therapies. The current review contributes to the growing field of research concerning technology mediated interventions.

Given their customizability HMD VR interventions may have an advantage over traditional types of interventions that take place in the real world, in that they may be tailored to the specific needs of an individual. For example, clients with high levels of anxiety or clients who experience sensory overload may benefit from the customizability of this type of intervention. Additionally, the included studies reported little to no drop out, which could indicate that individuals with ASD tolerate this type of intervention and might seem motivated to participate in interventions of this type.

A limitation of this systematic review is the wide inclusion criteria employed. Due to our broad focus on life skills, none of the included studies assessed the same interventions or used the same outcome measures. Some were similar in nature, for example, Beach and Wendt and Herrero and Lorenzo both created virtual environments, in which students could practice social skills. However, the employed interventions were characterized by differences, seen in dosage, and in the specific life skill that were in focus making it difficult to compare the studies directly. Hence the only common factor for the included studies is that of the technology mediating them, ie, HMD VR. Another limitation is that we did not attempt to obtain any grey literature regarding the effects of HMD VR interventions on life skills in individuals with ASD. Thus, while we did perform an extensive search of the literature using the specified databases, there may exist relevant research findings that have not been published via traditional channels and which therefore are not part of this review.

Future Research
Future research within this field should seek to address the methodological issues identified in the current review, such as the assessment of life skill improvements in real world settings along with follow-up assessments. This would allow for
firmer conclusions regarding the ability of ASD individuals to successfully transfer skills acquired within the virtual learning environment to the real world, as well as the longer-term effects of HMD VR interventions on participant life skills. Additionally, further research into this field of research might benefit from taking advantage of and exploring the effect of some of the potential benefits of HMD VR interventions. For instance through the use of sensory input adjusted scenarios based on the individual’s needs, or through the gradual increase of difficulty in the learning scenario. Such research would help illucidate the potential benefit to learning outcomes as a result of modulating sensory stimuli and task difficulty. Another potential advantage of VR interventions is the ability to replay the same scenario repeatedly and have participants implement small variations in behavior in response to the scenario to observe the effects of such behavioral variations. In short, it may be beneficial for future research to generally adopt a more personalized approach to designing and implementing HMD VR interventions.

Future research should also seek to implement control groups, include larger sample sizes and a wider range of the ASD population, thereby increasing the generalizability of their results. Thus, the next step for this field of research must be methodologically sound and rigorous studies, which allow for the careful evaluation of potential life skill improvements using this type of technology. Finally, future research needs to compare HMD VR interventions to the interventions currently in use, such as ABA, in order to properly evaluate this type of intervention against treatment-as-usual.

Concluding Remarks
The studies included in this systematic review generally reported improvements in specific life skills, apart from the study by Adjorlu et al26 where the authors encountered a ceiling effect and thus, were not able to detect a change in participants’ abilities. However, a number of methodological concerns have been raised in this review, such as the risk that some studies may suffer from a lack of ecological validity or control groups. A central aspect of the efficacy of HMD VR interventions is the question of whether such interventions can actually improve the life skills of participants in a real-world scenario. Showing improvements within a virtual environment or at answering questions concerning a given life skill does not necessarily imply increased real-life functionality. Some of the included studies did assess real-life skills as their outcome measures; however, due to a lack of statistical testing and weak study designs, it is difficult to draw any firm conclusions based on their findings. This is not to say that this field of research does not show potential. However, research of higher methodological quality is needed to be able to draw firm conclusions about the efficacy of employing HMD VR interventions to improve the life skills of individuals with ASD. Based on the above discussion, therapists and policymakers should be cautious before choosing to replace existing interventions to improve life skills in individuals with ASD with HMD VR interventions.

Disclosure
The authors do not have any conflicts of interest to report in this work.

References


Neuropsychiatric Disease and Treatment

Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the ‘PsycINFO’ database and CAS, and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.