INTERACTIVE LEARNING ENVIRONMENTS
https://doi.org/10.1080/10494820.2020.1811733

# A systematic review: the application of virtual reality on the skill-specific performance in people with ASD

Grace Mak and Lea Zhao

School of Health Sciences, University of South Australia, Adelaide, Australia

**ABSTRACT**
Given the prevalence of Autism Spectrum Disorder (ASD) and the demand for treatment, there is a continuous seeking and uncertainty of effective interventions for people with ASD. As technology continues to advance, the application of Virtual Reality is emerging in clinical settings. This systematic review summarised findings to evaluate the application of virtual reality (VR) on the skill-specific performance in people with ASD. The purpose is to determine (1) if VR is an effective treatment for people with ASD in skill-specific performance and (2) can Occupational Therapists employ VR in their practice. Eight databases were systematically searched for peer-reviewed articles that were published from January 2012 to February 2018. Eight articles met the inclusion criteria. The measurements of specific skills were categorised into three main domains: job interviewing, driving, and other ADLs. A diverse range of outcome measures were utilised and provided various results. Despite the consistent positive results reported in the studies, the current evidence base lacks justification of sample sizes, reliability and validity of the findings. Although VR shows potential as an effective intervention, limitations and bias of studies should be considered. Results of studies must be interpreted with caution if Occupational Therapists are interested in employing VR in their practice.

**ARTICLE HISTORY**
Received 22 July 2019
Accepted 14 August 2020

**KEYWORDS**
Occupational therapy; virtual reality; Autism Spectrum Disorder

## Introduction

### *Autism spectrum disorder*

Autism Spectrum Disorder (ASD) is a lifelong neurodevelopmental disability characterised by impairments in social communication and social interaction, as well as restricted, repetitive patterns of behaviour, interests, or activities (Roley et al., 2008). These symptoms can range on a spectrum from mild to severe based on the level of support the person needs. Symptoms might not be recognised until social demands such as schooling become greater. According to the Australian Bureau of Statistics (Australian Bureau of Statistics, 2018) and the Australian Institute of Health and Welfare (AIHW, 2017), an estimated 164,000 people were diagnosed with ASD in 2015, representing 1 in 150 Australians. Worldwide, the incidence of ASD was 1 per 110 in the United States, 1 per 64 in the United Kingdom, and 83% of them were under the aged of 25 (Ratajczak, 2011). According to Wikipedia, more than 21.7 million people worldwide are directly affected by ASD. Autism Spectrum Disorder, once considered rare, accounted for 31% of NDIS participants in 2015. However, the reason for the increasing prevalence of ASD is not clear, heightened awareness and efficient diagnosis may have contributed to the increased number (Australian Bureau of Statistics, 2018). There is no single diagnostic test for ASD, the diagnostic process consists of developmental assessments and

**CONTACT** Grace Mak makgt001@mymail.unisa.edu.au

behavioural observations. In order to diagnose a person, two specialists need to conduct assessments separately. The cause of ASD is uncertain however, it is now almost universally accepted that there is a genetic cause for ASD. Also, research is being conducted into environmental triggers for ASD, such as viral infections, complications during pregnancy and air pollutants. Males are more likely to become autistic. It occurs four to five times more often in boys. However, the ratio varies depending on studies and countries. A systematic review suggests that among people who are diagnosed with ASD, the male to female ratio is close to 3:1 (Loomes et al., 2017).

ASD can negatively impact a person's daily life in many ways due to sensory processing difficulties (Ausderau et al., 2014). They may show hypersensitive behaviours, and their basic living skills are likely to be affected, for example, cooking, street-crossing, handwriting, driving, and job interviewing skills. Moreover, crossing streets can be dangerous and challenging due to the person's sensory behaviours. In addition, controlling a vehicle can be a daunting multi-task challenge for novice drivers with autism (Ross et al., 2018). Furthermore, people with autism also encounter employment restrictions; about 3 in 10 (29%) people with ASD were permanently unable to work due to their condition or disability (AIHW, 2017). Hence, ASD can affect individuals across different aspects over the course of development.

### *Occupational therapy and autism spectrum disorder*

Occupational therapy focuses on the therapeutic use of everyday life activities to enhance independence, in addition, enable participation in roles, habits, and routines in homes, schools, communities, and other settings (Roley et al., 2008). Occupational therapists can play an important role in children's development, particularly children with ASD and their families. The American Occupational Therapy Association (AOTA) recommends that interventions should involve a family-centred approach and begin during the early years of a child's life in order to have successful outcomes (Roley et al., 2008). Early diagnosis of ASD enables families to learn about their child's behaviours and coping strategies, to seek appropriate services, and to set life goals (Shattuck et al., 2009). Occupational therapists work on different skills with children with ASD, which include Activities of Daily Living (ADL), self-care, sensory profiles, emotional regulation, observation of play, school performance, etc. There are many evidence-based ASD-related interventions that occupational therapists use in practice. For instance, coaching is a specific intervention used within a family-centred practice (Simpson, 2015); cognitive behavioural therapy is suggested to be effective (Ho et al., 2015); relationship-based approaches promote development by encouraging children to interact with parents and others (Di Renzo et al., 2015); the play-based approaches are very crucial to children with ASD because play is the main occupation for children (Roley et al., 2008).

### *Virtual reality*

Virtual reality (VR) is a computer-programmed technology that represents a projected simulation of real-world environments using images, sound, and/or tactile feedback. It includes interactive video gaming, virtual environments, multi-sensory experiences, interesting interactions, and situations for participants to immerse themselves in. The history of VR was not uniform and verified in any journal articles. However, the Oculus Rift that was launched on Kickstarter in 2012 drew broad interest from the public. It is considered "the first truly immersive VR for video games" (Barnard, 2019). In addition, the effectiveness of VR is shown to be effective as a learning tool for children with autism (Strickland, 1997). Furthermore, it has several useful features to master interactions with the real world. For example, VR offers controllable input stimuli to adjust the level of stimulation and interaction, provides safer learning environments so mistakes are less catastrophic, and also allows generalisation and repetitive practices (Strickland, 1997). It is very important for people with ASD to have many opportunities to overcome mistakes due to repetitive and restricted behaviours which would be difficult to practice in the real world (Didehbani et al., 2016). VR can also provide various scenarios that may not be applicable in a real-life therapeutic setting. A large number of children with ASD are visual learners who learn better through visual inputs (Quill, 1995). As such, VR technology appears to be suitable to increase skill-specific performance in people with ASD.

### *Our review*

The reason skill-specific performance was chosen as the area of interest was because a previous systematic review had already been done for VR and social skills (Thai & Nathan-Roberts, 2018). The study stated that there was a myriad of VR systems which targeted a variety of different social skills. Thus, they could not confidently conclude that VR training could replace traditional face-to-face therapy to enhance social skills. As a result, we wanted to explore new skills and areas of research in regard to ASD. Moreover, as Occupational Therapy students, we are interested in specific tasks and activities of daily living that are important to people with ASD.

Although the use of VR technology is becoming more popular in clinical settings, critical appraisal of the research evidence is lacking. No review exists to systematically evaluate the quality of evidence presented for the application of VR interventions that targets different types of skill performances. Therefore, the goals of this systematic review are to determine (1) if VR is effective as a treatment for people with ASD in skill-specific performance and (2) if Occupational Therapists can employ VR in their practice.

## Methods

### *Search strategy*

The method for searching journal articles followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) format (Moher et al., 2009). Eight databases have been searched based on MeSH headings and keywords from 2012 to 2019 including: Embase, Ovid Emcare, Eric, Medline, PubMed, OTSeeker, Cochrane, and Google Scholar. These databases were chosen due to their relevance to the topic and diversity of the studies. The limit from the year 2012 was set to include more relevant publications, as the first development kit for Oculus Rift - the first truly immersive virtual reality headset for video games was launched as a Kickstarter campaign and VR technology became commerically available in 2012 (Barnard, 2019). The MeSH headings were: Autism OR Autism Spectrum Disorder OR Asperger syndrome AND Virtual Reality. The search was limited to English and human. The result of the database, Medline is displayed in Appendix 1.

### *Study design*

This review included all levels of studies that met the eligibility criteria, including randomised controlled trials (RCTs), clinical control trials (CCTs), quasi-experimental, pre–post studies, cohort studies, and case studies. We selected population-intervention-comparator-outcome (PICO) model as the eligibility criteria to assist in the screening process. The four PICO elements are outlined below.

### *Population*

As mentioned before, ASD is a lifelong condition that impacts people of all age groups. Therefore, this review included all populations who have been diagnosed with ASD by DSM-5 (American Psychiatric Association, 2013) and Asperger's syndrome by DSM-4 (American Psychiatric Association, 1994). Populations with other conditions were excluded, such as visual and hearing disabilities, and Attention Deficit Hyperactivity Disorder (ADHD).

### *Intervention*

Studies were included if virtual reality technology was developed to improve specific living skills, such as driving, job interviewing. Studies were excluded if the VR software was specifically designed for assessing social skills. Other computer-based interventions were excluded, such as augmented reality (AR), video watching, and computer simulation (non-immersive VR).

### *Comparison*

Comparators were not limited to the search, desirable comparators included but not limited to non-VR-based treatment, participants condition before VR-based treatment, and no comparator.

### *Outcome*

As a variety of outcomes can be related to the effect of VR, the search was not limited to any specific outcome. Outcome interests included but not limited to occupational functioning, performance, living skills, quality of life, attitude, self-confidence, body functions, and independence levels.

### *Literature search*

The databases have been searched after the development of search strategy and eligibility criteria. All results were imported to EndNote™ to remove duplicates then exported to Covidence™ for screening. Titles and abstracts were screened independently by two reviewers (GM & LZ) to determine their eligibility before full-text screening. Any disagreements were resolved by discussion between the two reviewers.

### *Methodological quality*

The Modified McMaster Quantitative Studies Appraisal Tool (Law et al., 1998) and the PEDro scale (Physiotherapy Evidence Database, 1999) were employed as they are both widely recognised in systematic reviews, free to use, and include an easy scoring system that facilitates the users. The appraisal tool from the Modified McMaster Quantitative Studies Appraisal Tool consists of 15 questions about the study purpose, design, sample size, intervention, results, and clinical implications. 14 of the questions allows users to score accordingly. The answer of "Yes" grants 1 point, while the answer of "No" or "Not applicable" apply zero point into the raw score. All studies were systematically assessed under the Modified McMaster Quantitative Studies Appraisal Tool to facilitate comparison. Three RCT were additionally evaluated by the PEDro scale, which are outlined in Appendix 2. The PEDro scale is a valid measure of the methodological quality of clinical trials specifically for RCT. There are 11 answerable questions, while 10 of them are scoring questions. The scoring system is the same as the Modified McMaster Quantitative Studies Appraisal Tool, which the answer "Yes" gives 1 point to the raw score.

Two reviewers (GM and LZ) independently analysed the methodological quality of selected studies, and disputes were resolved through discussion. To determine the level of evidence of selected studies, the hierarchy of evidence for each study was assessed according to the NHMRC Designation of Levels of Evidence (National Health and Medical Research Council, 2009).

## Results

### *Study selection*

The initial search generated a result of 436 studies for screening. After the elimination of duplicates, 249 studies were screened for titles and abstract. Thirty-eight studies were reviewed in full text and eight met the inclusion criteria. Thirty studies were excluded due to irrelevant outcome measures ($n$ = 9), wrong intervention ($n$ = 5), wrong study designs ($n$ = 5), duplication (5), reviews ($n$ = 4), ongoing studies ($n = 1$) and wrong keywords of intervention ($n = 1$). The literature selection process is presented in PRISMA format in Figure 1.

## *Risk of bias*

An overview of the levels of evidence and critical appraisal scores of the included studies were shown in Table 1 (Modified McMaster Quantitative Studies) and Table 2 (PEDro Scale). As per NHMRC levels of evidence (National Health and Medical Research Council, 2009), three studies we rated as Level II RCTs (Cox et al., 2017; Ross et al., 2018; Smith et al., 2014) and five studies were rated as level IV case-series (Burke et al., 2018; Lu et al., 2018; Muneer et al., 2015; Simões et al., 2018; Smith et al., 2014). The main methodological concerns were: lack of justification of the sample size -only one study completed power calculation (Lu et al., 2018); lack of validity and reliability recording in outcome measures -two studies reported reliability (Ross et al., 2018; Smith et al., 2014) and only one study reported validity (Lu et al., 2018); and lack of addressing clinical significance -none of the studies reported. In addition, none of the therapists was blinded, and allocation was concealed within

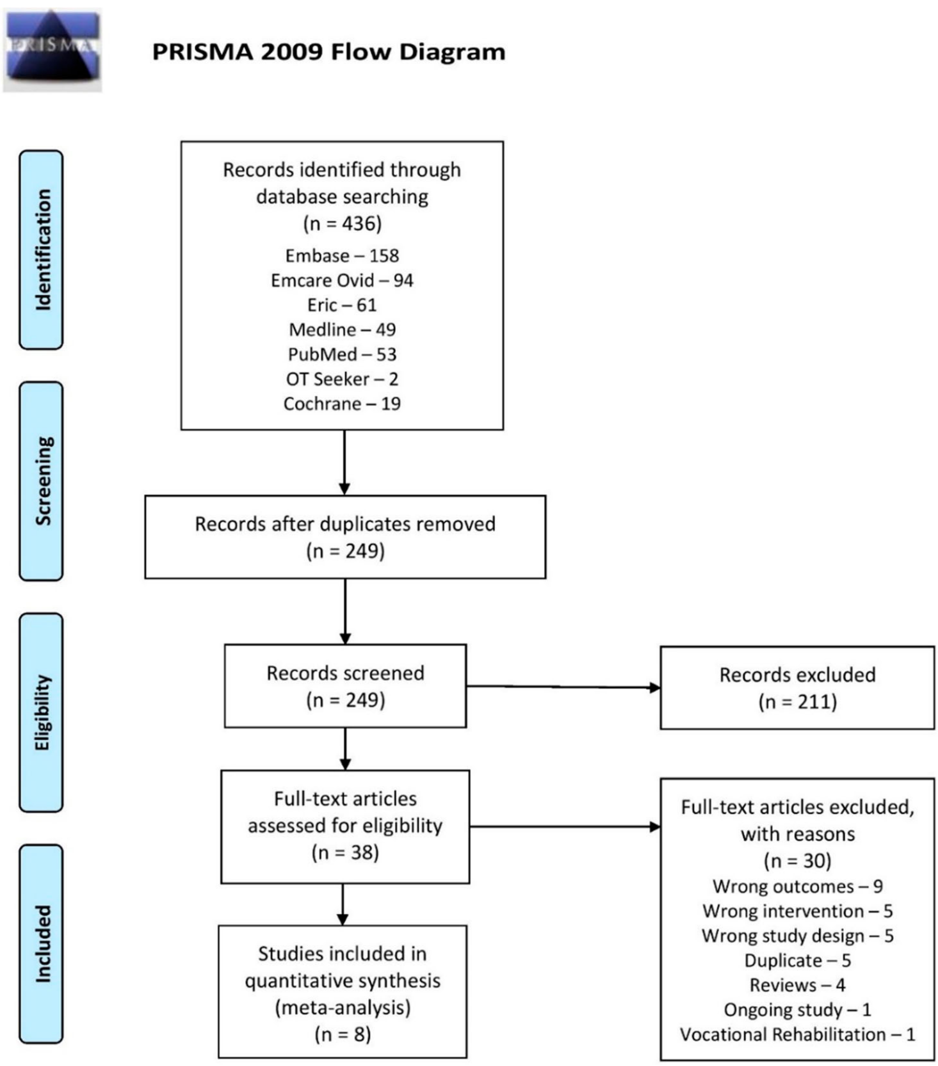



**Figure 1.** PRISMA flow diagram.

three RCTs (Cox et al., 2017; Ross et al., 2018; Smith et al., 2014). The method of randomisation was not demonstrated in two RCTs studies (Ross et al., 2018; Smith et al., 2014). Both widely utilised critical appraisal tools demonstrated consistent analysis based on the correlation of the raw score shown in Tables 1 and 2. For instance, the study by Smith and his team (2014) achieved the highest score in both tables.

## *Study characteristics*

Table 3 demonstrates the study characteristics of the included articles. It shows research designs, study age range, sample size, participants characteristics, interventions, and comparator/ control. Specific VR Intervention types of each study were explained in Table 4. All studies were published between January 2012 and February 2019 and conducted in various countries: India (Muneer et al., 2015), Singapore (Lu et al., 2018), Portugal (Simões et al., 2018), Italy (Saiano et al., 2015), and the United States (Burke et al., 2018; Cox et al., 2017; Ross et al., 2018; Smith et al., 2014).

## *Participants characteristics*

Total study sample sizes ranged from 5 to 60 participants with age ranging from 4 to 44. A total of 213 participants were included across the 8 studies. The number of participants with ASD ranged from 4 to 66. Five studies included children with typical development (Burke et al., 2018; Lu et al., 2018; Muneer et al., 2015; Simões et al., 2018; Smith et al., 2014), and two studies included participants with intellectual disability ($n = 4$) (Lu et al., 2018; Saiano et al., 2015). The majority of the participants were male. Participants were excluded if they had neurological, muscular or systemic diseases that may affect lower limbs and if there was a history of trauma or surgery. DSM-V and IV were the common diagnostic parameters of ASD. However, studies included participants with formal diagnosis by DSM-V and informal diagnosis from their family doctors. Thus, there was no consistency in terms of the diagnostic parameters for autism amongst the included studies.

## *Types of VR inventions*

VR interventions varied from targeting different ADL skills. Two studies used Virtual Reality Driving Simulation (VRDS) software as an intervention for driving skills (Cox et al., 2017; Ross et al., 2018). Two studies measured job interview skills using Virtual Interactive Training Agents (ViTA) (Burke et al., 2018) and Virtual Reality – Job Interview Training (VR-JIT) software (Smith et al., 2014). One study used Neuro VR software that was specifically developed to measure safety skills, including pedestrian skills and following road signs (Saiano et al., 2015). The other three studies used VR games as their interventions (Lu et al., 2018; Muneer et al., 2015; Simões et al., 2018). Details and descriptions of each intervention are listed in Table 4.

**Table 1.** Levels of evidence and modified McMaster results of methodological quality.

| Study | NHMRC level and study design | Modified McMaster items | | | | | | | | | | | | | | Raw score (total 14) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | |
| Burke et al. (2018) | IV-case series | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 10 |
| Smith et al. (2014) | II-RCT | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 11 |
| Ross et al. (2018) | II-RCT | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 8 |
| Cox et al. (2017) | II-RCT | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 10 |
| Muneer et al. (2015) | IV-case series | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 10 |
| Lu et al. (2018) | IV-case series | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 7 |
| Simões et al. (2018) | IV-case series | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 10 |
| Saiano et al. (2015) | IV-case series | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 10 |

Note: 1 = Yes, 0 = No, Not addressed.

**Table 2.** Levels of evidence and PEDro results of methodological quality.

| Study | NHMRC level and study design | PEDro scale items | | | | | | | | | | | Raw score (Total 10) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1* | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | |
| Smith et al. (2014) | II-RCT | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 7 |
| Ross et al. (2018) | II-RCT | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 5 |
| Cox et al. (2017) | II-RCT | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 6 |

Note: 1 = Yes, 0 = No, *item 1 does not contribute to the total raw score.

**Table 3.** Study characteristics.

| Study | Country | *N* participants with ASD | Age and gender | Ethnicity |
|---|---|---|---|---|
| Burke et al. (2018) | United States | *N* = 32; ASD=22 | Age 19–31; Male = 25 | African American, Biracial, Caucasian, Hispanic |
| Smith et al. (2014) | United States | *N* = 26; ASD = 16 | Age 18–31; Male = 12 | Caucasian, African American, other ethnicities |
| Ross et al. (2018) | United States | *N* = 60; All ASD | Age 16–25; Not addressed | Not addressed |
| Cox et al. (2017) | United States | *N* = 51; All ASD | Age 15.5–25; Male=40 | Not addressed |
| Muneer et al. (2015) | India | *N* = 5; ASD = 4 | Age 4–8; All male | Not addressed |
| Lu et al. (2018) | Singapore | *N* = 12; ASD = 10 | Age 8–16; Gender not addressed | Not addressed |
| Simões et al. (2018) | Portugal | *N* = 20; ASD = 10 | Mean age 18.8; Male=13 | Not addressed |
| Saiano et al. (2015) | Italy | *N* = 7; All ASD | Age 19–44; All males | Not addressed |

## Outcome measures

A range of outcomes and outcome measures (OMs) were employed to assess the application of VR. Outcomes have been clustered into three categories – job interview, driving, and other activity of daily living (ADL). There was a mixture of subjective and objective measures. Subjective methods included questionnaires that were self-reported (Smith et al., 2014), surveys for experience (Lu et al., 2018; Smith et al., 2014), driving attitude scales reported by parents (Ross et al., 2018), and feedback from teachers (Lu et al., 2018). Objective methods included skill specific parameter assessments such as knowledge tests (Saiano et al., 2015), electrodermal activity (EDA) (Simões et al., 2018), driving-specific executive function abilities and tactical driving skills (Cox et al., 2017), standardised job interview role play assessments (Smith et al., 2014), and the Marino Interview Assessment Scale (Burke et al., 2018). The duration of OMs also varied and ranged from two weeks to approximately six months**.** Two studies did not report the pre–post duration or dates (Lu et al., 2018; Simões et al., 2018). Table 5 provides a summary of various outcome domains and corresponding measures.

## Outcome domain – job interview

Two studies measured the application of VR intervention on job interview skills (Burke et al., 2018; Smith et al., 2014). One reported statistically significant improvement in job interview skills and increased self-confidence (Smith et al., 2014). While another one failed to determine the true effect due to long measurement points (Burke et al., 2018).

## Outcome domain – driving

Two studies measured the application of VR training on driving skills (Cox et al., 2017; Ross et al., 2018). Both used the same intervention, VR driving simulation (VRDS), and compared to the usual

**Table 4.** Study intervention and comparison.

| Study | Intervention: Type | Intervention: Frequency | Intervention: Intensity | Comparison | Measurement Duration |
|---|---|---|---|---|---|
| Burke et al. (2018) | Virtual Interactive Training Agent (ViTA) | 2 sessions a week | Practice with training agent 10–12 questions | Baseline: ViTa first session; Post-test: final face to face interview | 26 weeks |
| Smith et al. (2014) | Job interview training (VR-JIT) | 5 sessions | 120 min | Treatment As Usual | 2 weeks |
| Ross et al. (2018) | VR driving stimulator (VRDS) | 8–12 sessions | 60 min | Routine Training | 12 weeks |
| Cox et al. (2017) | VR driving stimulator (VRDS) | 8–12 sessions | 60 min | Routine Training | 12 weeks |
| Muneer et al. (2015) | VR games | 4–6 sessions | 30 min | Baseline: first VR session; Post-test: last VR session | 4 weeks |
| Lu et al. (2018) | VR Pink Dolphins Game | Not addressed | Not addressed | Not addressed | Not addressed |
| Simões et al. (2018) | VR Bus Taking Serious Game | 3 sessions | 20–40 min | Control group with typical development subjects | Not addressed |
| Saiano et al. (2015) | Neuro VR | 10 sessions | 45 min | Familiarisation: Session 1–5; Baseline: Session 6; Post-test: Session 10 | 10 weeks |

Note: Abbreviation: approx. = approximately.

routine training as an intervention. Both reported significantly improved driving skills. One study also reported improved executive functioning performance (Cox et al., 2017).

## *Outcome domain – ADLs*

Four studies measured the application of VR intervention on different ADL skills (Lu et al., 2018; Saiano et al., 2015; Simões et al., 2018). The study that measured four areas: motor, communication, cognitive and social/emotional skills reported statistically significant improvement in three domains, except communication (Muneer et al., 2015). The study that evaluated direction following, psychomotor skills and hand-eye coordination reported improvement in all outcomes and received informal positive feedback of enjoyment of the VR game (Lu et al., 2018). The study that calculated bus taking transportation skills reported statistically significant improvement in the theoretical knowledge of the process and decrease anxiety level during the process (Simões et al., 2018). The study that assessed safety skills (which included pedestrian skills and following road signs) reported no improvement in pedestrian knowledge according to the subjects' surveys, but positive changes in pedestrian behaviours in the parents' survey (Saiano et al., 2015).

## *Summary of results*

Table 5 provides an overview of results from the included studies which demonstrates the diversity of outcome measures within three domains. Despite this heterogeneity, the overall findings indicate that VR may have a positive impact across a range of outcomes. There was consistent evidence showing that VR successfully improved skill-specific performance in job interviewing, driving skills and other ADLs. Results also showed multiple significant changes in increased self-confidence during job interviewing (Smith et al., 2014), reduced negative attitude during driving (Ross et al., 2018), enhanced transportation knowledge (Simões et al., 2018), reduced anxiety level during bus taking (Simões et al., 2018), improved navigation skills (Saiano et al., 2015), and decreased errors following road signs (Saiano et al., 2015). Outcome results also demonstrated that the safety skills learnt were able to transfer into real life settings (Saiano et al., 2015). Skills of motor, cognitive, social/emotional, direction following, and communication by giving directions also significantly improved

**Table 5.** Summary of outcome domains, measures and results.

| Study | Outcome domain | | Outcome measure | Outcome Result |
|---|---|---|---|---|
| Burke et al. (2018) | Job interview | Job interviewing performance skills | Face to face interview Marino Interview Assessment Scale | ↑(+)* |
| Smith et al. (2014) | | Job interviewing performance skills | Standardised job interview role plays performance assessment | ↑(+) |
| | | Self-confidence level | Self-reported Likert's scale | ↑(+)* |
| | | Feedback of intervention | Training Experience Questionnaire (TEQ) | ↑(+) |
| Ross et al. (2018) | Driving | ASD novice driver attitude | Driving Attitude Scale Parent Report | ↓(+)* |
| Cox et al. (2017) | | General tactical driving performance | VRDS tactical testing | ↑(+) |
| | | Driving-specific executive function (EF) abilities | VRDS executive function test | ↔ |
| Muneer et al. (2015) | Other activities of daily living (ADLs) | Motor | 4-points rating scale | ↑(+)* |
| | | Communication | | ↑(+) |
| | | Cognitive | | ↑(+)* |
| | | Social/emotional | | ↑(+)* |
| Lu et al. (2018) | | Direction following | "Grid game" | ↑(+)* |
| | | Communicate by giving directions | "Grid game" | ↑(+)* |
| | | Psychomotor skills | "Do as you see" game | ↑(+)? |
| | | Hand-eye coordination | "Do as you see" game | ↑(+)? |
| | | Feedback of intervention | Participant survey | ↑(+) |
| | | | Teachers' questionnaire | ↑(+) |
| Simões et al. (2018) | | Bus taking transportation performance skills | Action accuracy calculated by intervention software | ↑(+) |
| | | | Debriefing accuracy | ↑(+)* |
| | | Anxiety level | Electrodermal Activity (EDA) | ↓(+)* |
| Saiano et al. (2015) | | Safety skills – navigation performance | Neuro VR application automatically record results | ↑(+)* |
| | | Safety Skills – following road signs number of errors | Neuro VR application automatically record results | ↓(+)* |
| | | Subjects' acquisition of street safety skills | McNemar's test | ↔ |
| | | Transfer of learned behaviour into real life | 6-point Likert-type questionnaire for parents/caregivers | ↑(+)* |

Note: ↑ = increase, ↓ = decrease, ↔ = no change, (+) = positive change/improvement, (–) = negative change, * = statistical significance ($p < .05$), (?) = significance not reported. Abbreviation: VRDS = Virtual Reality Driving Stimulator.

during participating in VR games (Lu et al., 2018; Muneer et al., 2015). Overall, participants' and caregivers' feedback of the interventions were positive (Lu et al., 2018; Smith et al., 2014).

## *NHMRC body of evidence matrix*

Table 6 demonstrates the analysis of results using NHMRC framework. Despite consistent positive outcomes reported across the studies, most studies ranked as low evidence base and performed minimal clinical impact and generality. For example, none of the studies calculated sample size estimates and reported clinical impact. Although there were consistent reports of outcomes, diverse intervention types and multiple study designs, implementation of recommendations should be applied with caution.

## *Discussion*

As VR is an emerging technology, its effectiveness on people with ASD remains uncertain. The goals of this systematic review are to determine (1) if VR is effective as a treatment for people with ASD in skill-specific performance and (2) if Occupational Therapists can employ VR in their practice. While the previous systematic review by Thai and Nathan-Roberts focused on social skills in ASD, this review

**Table 6.** NHMRC body of evidence matrix.

| Component | Grade | Comments |
|---|---|---|
| 1. Evidence Base | *D-Poor* =<br>Level IV studies, or level I to III studies with high risk of bias | • Studies ($n = 8$)<br>• 3 Level II-RCT studies<br>• 5 Level IV-Case-series studies<br>• No study calculated sample size estimates<br>• Fail to calculate statistical significance due to small sample size<br>• Low use of standardised assessments |
| 2. Consistency | *B-Good* = Most studies consistent and inconsistency may be explained | • Consistent reporting results of statistical significance<br>• Multiple study designs<br>• Diverse intervention types<br>• Varied outcomes and measurements |
| 3. Clinical impact | *D-Poor* =<br>Slight or restricted | • None of the studies report clinical significance<br>• Different types of interventions used with inadequate description for replicability in clinical practice<br>• Short intervention duration |
| 4. Generalisability | *D-Poor* =<br>Population/s studied in body of evidence differ to the target population and hard to judge whether it is sensible to generalise to the target population | • Small sample size<br>• Some participants have ASD and other conditions, such as an intellectual disability<br>• Age range: 4–44<br>• Majority of participants were male participants<br>• Included studies were conducted in multiple countries |
| Grade of recommendations | *D-Poor* =<br>Body of evidence is weak, and recommendation must be applied with caution | • In conclusion, most studies had a low level of evidence and lacked clinical significance. Although there was consistency in reporting the statistical significance in findings, there was a lack of uniform outcomes, intervention delivery protocols and adequate intervention time |

focuses on specific tasks in daily life which are more related to the Occupational Therapy practice (2018). Unlike the previous systematic review, this review employed NHMRC Framework to evaluate the strength of the evidence and give recommendations to professional practice (NHMRC, 2019).

A modest body of evidence base consisting of eight studies with different research designs was included. The summarised findings from this review indicated that VR may have a positive impact across a range of outcomes including job interview skills, driving skills, safety skills, bus taking skills, body function, and performance skills. Multiple studies demonstrated equivalent positive results despite different outcome measures. Two studies measured job interviewing skills and both demonstrated significant improvement through using standardised role play and face-to-face interview assessments (Burke et al., 2018; Smith et al., 2014). Two studies that assessed physiological feedback indicated that VR successfully decreased anxiety and increased confidence during task performance (Cox et al., 2017; Ross et al., 2018). Although some studies did not directly measure motor and cognitive skills as outcomes, researchers did observe that there was a significant increase in the movement of speed and attention level during VR interventions (Muneer et al., 2015; Saiano et al., 2015). VR is a technology that provides visual and auditory stimuli that immerses users into a realistic setting, where performance and body functions are required to interact with the environment to perform a task. Through improving different aspects of skill development, it would enhance the skill-specific performance level and the independence of people with ASD.

It is important to measure the acceptance of the intervention because people with ASD may have sensory processing dysfunction, as mentioned in the introduction. They may experience hypersensitivity to the demands of the environment in real life settings. However, none of the subject dropped out due to the discomfort of the intervention. It suggested that VR could be applicable to people with ASD. The enjoyment and effectiveness of VR were also reflected on the dropout rates and the feedback of the intervention. Low dropout rates were consistent among all included studies. Dropout reasons were solely due to scheduling (Saiano et al., 2015). Two studies, which evaluated participants' experiences of VR, explained VR is helpful in preparing participants for future interviews (Smith et al., 2014); fun and useful in motivating and reinforcing children with ASD (Lu et al., 2018). Other studies also mentioned that VR could provide a safe simulated environment for repetitive practice (Burke et al., 2018; Cox et al., 2017) which aligned with the positive feedback of the intervention. In addition, VR offers an adjustable platform that helps people with ASD relate to reality (Lu et al., 2018). Difficulty levels could be altered based on the users' abilities. Modification of VR systems also promotes sustainability of the intervention. Therefore, VR could be potentially adapted for clinical practice and applied to people with ASD.

Although all outcomes promoted a positive impact on skill performance, some results failed to show the statistical significance of improvement, such as communication skills. It could be due to the intervention time and the relationship with researchers. Social interaction is one of the difficulties that people with ASD struggle with the most. A longer period of time would allow rapport building and communication skill enhancement. In terms of motor skill development, one study showed statistically positive improvement (Muneer et al., 2015). However, the other study failed to show significant improvement (Lu et al., 2018). Most of the intervention periods were short (2 weeks to 6 months). As most people with ASD are visual learners, they may require more time to transfer visual information to motor behaviours. Therefore, the included studies may not allow sufficient time to show significant improvement in motor and communication skills.

Despite the consistency of the positive outcomes, the current body of evidence is weak. Recommendations must be applied with caution when interpreting the findings. After the process of critical appraisal, most of the studies lacked reporting the reliability and validity of the results. None of the current evidence reported the clinical importance and sample size justification. Given the heterogeneity of the included studies, a widely established framework (NHMRC Framework) was utilised to synthesise the descriptive data and constructed a poor grade of recommendation.

### *Limitation*

This systematic review has many limitations that need to be recognised. Firstly, there were concerns with the methodological quality including small sample size, gender ratio of the sample, allocation bias, unreliable outcome measure tools, and short intervention time. Secondly, due to the diversity of outcome measures, it is impossible to compare the outcomes directly among the selected studies. Thirdly, the use of questionnaires for parents and participants failed to demonstrate that the participants had successfully mastered the skills from the VR training and applied them in a real life setting. For example, pedestrian behaviours and following safety rules were hard to measure because of the safety concerns in real life situations (Saiano et al., 2015; Simões et al., 2018). Most of the studies reported increased skill in a specific task after VR training but failed to measure actual real-life performance such as an on-road driving performance (Burke et al., 2018; Smith et al., 2014). Despite the low dropout rates and positive feedback from participants, the contraindication to VR includes temporary discomforts, motion sickness and any long term effects that were not reported. Furthermore, publication and language bias should be recognised as the search relied on a relatively limited number of databases and the year was limited from 2012 to 2019. Studies that were only published in other databases, before 2012, and non-English languages may have been missed. As VR is a relatively new technology, there is no standard protocol and consistency in software development. A previous systematic review also mentioned there was no universal design that aimed at measuring the most necessary skills needed to be taught individuals with ASD (Thai

& Nathan-Roberts, 2018). Furthermore, most of the included studies developed their own VR training systems and conducted intervention in a short period of time. Thus, the results failed to reflect the true effectiveness of VR. It is important to address the limitations and to consider the application of the new technology in the treatment of ASD.

## Conclusion

### Implication for OT practice

In terms of the treatment for people with ASD, VR-based treatment may have some advantages over traditional treatments. One of the main advantages is that VR allows clients to immerse themselves into real life situations and practice skills in a safe and protective environment. VR is both a therapeutic tool and an assessment tool. It can be employed to assist with some OT treatments such as education, skill demonstration and home therapy. For home therapy, clients can practice certain skills at home without OT supervision. With a trustworthy recording system, it also allows therapists to follow up and monitor the progress. For clients with a lack of interest in traditional treatments, VR can provide a "just right challenge" which can enhance motivation and engagement. Although most results from the included studies are positive, OT clinicians should use VR cautiously in their client-centred practice since the cost, duration of application, and the weight of the VR headset may cause discomfort for specific individuals with ASD.

### Implication for future research

The current evidence for virtual reality as a treatment for people with ASD is profoundly limited. The development of a standardised protocol that researchers and clinicians can follow is required. Standardised assessments are also needed in order to have consistent and reliable measurements. More randomised controlled trials, with larger sample sizes and longer periods of intervention, are required in order to provide rigorous and precise recommendations.

## Acknowledgements

The systematic review topic was conceptualised, designed and implemented by the authorship team. Each member of the authorship team was involved in the searching, assessing, appraising and synthesising of the literature and equally contributed to the drafting and revising of manuscript. The authorship team is accountable for the work underpinning this systematic review and any issues related to the accuracy or integrity of any of this work will be appropriately investigated and resolved. The final manuscript has been read and approved by all authors. We would like to show our gratitude to Dr. Saravana Kumar for sharing his pearls of wisdom with us during the course of this research as a UniSA Associate Professor. We are also immensely grateful to Kiersten Fletcher for comments that greatly improved the manuscript.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

*Grace Mak* http://orcid.org/0000-0003-3624-3082
*Lea Zhao* http://orcid.org/0000-0003-2336-1557

## References

American Psychiatric Association. (1994). *Diagnostic and statistical Manual of Mental Disorders: DSM-IV*.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*.

Ausderau, K. K., Furlong, M., Sideris, J., Bulluck, J., Little, L. M., Watson, L. R., Boyd, B. A., Belger, A., Dickie, V. A., & Baranek, G. T. (2014). Sensory subtypes in children with autism spectrum disorder: Latent profile transition analysis using a national survey of sensory features. *Journal of Child Psychology and Psychiatry*, *55*(8), 935–944. https://doi.org/10.1111/jcpp.12219

Australian Bureau of Statistics. (2018, November 2). *4430.0 – Disability, ageing and carers, Australia: Summary of Findings, 2018*. https://www.abs.gov.au/ausstats/abs@.nsf/mf/4430.0

Australian Institute of Health and Welfare (AIHW). (2017). *Autism in Australia*. Australian Government. https://www.aihw.gov.au/reports/disability/autism-in-australia/contents/autism

Barnard, D. (2019). *History of VR – Timeline of events and tech development*. https://virtualspeech.com/blog/history-of-vr

Burke, S. L., Bresnahan, T., Li, T., Epnere, K., Rizzo, A., Partin, M., Ahlness, R. M., & Trimmer, M. (2018). Using Virtual interactive training agents (ViTA) with adults with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders*, *48*(3), 905–912. https://doi.org/10.1007/s10803-017-3374-z

Cox, D. J., Brown, T., Ross, V., Moncrief, M., Schmitt, R., Gaffney, G., & Reeve, R. (2017). Can youth with autism spectrum disorder use virtual reality driving simulation training to evaluate and improve driving performance? An exploratory study. *Journal of Autism and Developmental Disorders*, *47*(8), 2544–2555. https://doi.org/10.1007/s10803-017-3164-7

Didehbani, N., Allen, T., Kandalaft, M., Krawczyk, D., & Chapman, S. (2016). Virtual reality social cognition training for children with high functioning autism. *Computers in Human Behavior*, *62*, 703–711. https://doi.org/10.1016/j.chb.2016.04.033

Di Renzo, M., Di Castelbianco, F. B., Petrillo, M., Racinaro, L., & Rea, M. (2015). Assessment of a long-term developmental relationship-based approach in children with autism spectrum disorder. *Psychological Reports*, *117*(1), 26–49. https://doi.org/10.2466/15.10.PR0.117c15z8

Ho, B. P., Stephenson, J., & Carter, M. (2015). Cognitive–behavioural approach for children with autism spectrum disorder: A literature review. *Journal of Intellectual and Developmental Disability*, *40*(2), 213–229. https://doi.org/10.3109/13668250.2015.1023181

Law, M., Stewart, D., Pollock, N., Bosch, J., & Westmorland, M. (1998). *Critical review form – Quantitative studies*. McMaster University Occupational Therapy Evidence-Based Practice Research Group.

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lu, A., Chan, S., Cai, Y., Huang, L., Nay, Z. T., & Goei, S. L. (2018). Learning through VR gaming with virtual pink dolphins for children with ASD. *Interactive Learning Environments*, *26*(6), 718–729. https://doi.org/10.1080/10494820.2017.1399149

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264–269. https://doi.org/10.1371/journal.pmed.1000097

Muneer, R., Saxena, T., & Karanth, P. (2015). Virtual reality games as an intervention for children: A pilot study. *Disability, CBR & Inclusive Development*, *26*(3), 77–96. https://doi.org/10.5463/dcid.v26i3.456

National Health and Medical Research Council. (2009). *NHMRC evidence hierarchy*. Australian Government. https://www.mja.com.au/sites/default/files/NHMRC.levels.of.evidence.2008-09.pdf&gt

National Health and Medical Research Council. (2019). *NHMRC additional levels of evidence and grades for recommendations for developers of guidelines 2009*. Australian Government. https://www.mja.com.au/sites/default/files/NHMRC.levels.of.evidence.2008-09.pdf

Physiotherapy Evidence Database. (1999, June). *PEDro scale*. The Centre of Evidence-Based Physiotherapy (CEBP). https://www.pedro.org.au/

Quill, K. A. (1995). Visually cued instruction for children with autism and pervasive developmental disorders. *Focus on Autistic Behavior*, *10*(3), 10–20. https://doi.org/10.1177/108835769501000302

Ratajczak, H. V. (2011). Theoretical aspects of autism: Causes—A review. *Journal of Immunotoxicology*, *8*(1), 68–79. https://doi.org/10.3109/1547691X.2010.545086

Roley, S. S., Barrows, C. J., Brownrigg, S., Sava, D. I., Talley, V., Voelkerding, K., Lieberman, D., Honaker, D., DeLany, J. V., Smith, E., Toto, P., King, S., Lieberman, D., Baum, M. C., Cohen, E. S., Cleveland, P. A. M., & Youngstrom, M. J. (2008). Occupational therapy practice framework: Domain &. *American Journal of Occupational Therapy*, *62*(6), 625. https://doi.org/10.5014/ajot.62.6.625

Ross, V., Cox, D. J., Reeve, R., Brown, T., Moncrief, M., Schmitt, R., & Gaffney, G. (2018). Measuring the attitudes of novice drivers with autism spectrum disorder as an indication of apprehensive driving: Going beyond basic abilities. *Autism*, *22*(1), 62–69. https://doi.org/10.1177/1362361317735959

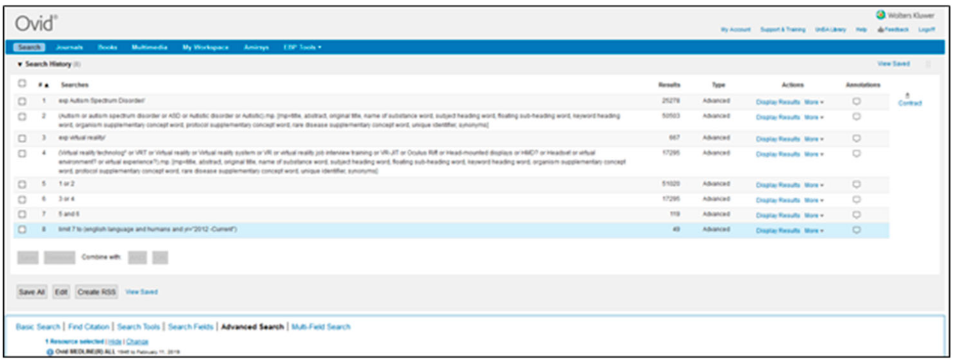
Saiano, M., Pellegrino, L., Casadio, M., Summa, S., Garbarino, E., Rossi, V., Dall'Agata, D., & Sanguineti, V. (2015). Natural interfaces and virtual environments for the acquisition of street crossing and path following skills in adults with autism spectrum disorders: A feasibility study. *Journal of Neuroengineering and Rehabilitation*, *12*(1), 17. https://doi.org/10.1186/s12984-015-0010-z

Shattuck, P. T., Durkin, M., Maenner, M., Newschaffer, C., Mandell, D. S., Wiggins, L., Lee, L.-C., Rice, C., Giarelli, E., Kirby, R., Baio, J., Pinto-Martin, J., & Cuniff, C. (2009). Timing of identification among children with an autism spectrum disorder:

Findings from a population-based surveillance study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *48*(5), 474–483. https://doi.org/10.1097/CHI.0b013e31819b3848
Simões, M., Bernardes, M., Barros, F., & Castelo-Branco, M. (2018). Virtual travel training for autism spectrum disorder: Proof-of-concept interventional study. *JMIR Serious Games*, *6*(1), e5. https://doi.org/10.2196/games.8428
Simpson, D. (2015). Coaching as a family-centred, occupational therapy intervention for autism: A literature review. *Journal of Occupational Therapy, Schools, & Early Intervention*, *8*(2), 109–125. https://doi.org/10.1080/19411243.2015.1040941
Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., Olsen, D. E., Bell, M. D., & Fleming, M. F. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *44*(10), 2450–2463. https://doi.org/10.1007/s10803-014-2113-y
Strickland, D. (1997). Virtual reality for the treatment of autism. *Studies in Health Technology and Informatics*, *44*, 81–86. https://pubmed.ncbi.nlm.nih.gov/10184809/
Thai, E., & Nathan-Roberts, D. (2018). Social skill focuses of virtual reality systems for individuals diagnosed with autism spectrum disorder; A systematic review. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *62*(1), 1469–1473. https://doi.org/10.1177/1541931218621333

## Appendices

### *Appendix 1. Database Medline search result*




### *Appendix 2. PEDro scale criteria*

| | |
|---|---|
| 1 | Eligibility criteria specified |
| 2 | Random allocation |
| 3 | Concealed allocation |
| 4 | Baseline similarity |
| 5, 6 & 7 | Blinding of subjects, therapists and assessors |
| 8 | Measures of key outcomes from more than 85% of subjects |
| 9 | Intention to treat analysis |
| 10 | Between-group statistical comparisions |
| 11 | Point measures of variability |