MASTERARBEIT | MASTER'S THESIS

Titel | Title

VR as an assessment tool in late childhood: A systematic review.

verfasst von | submitted by

Ahmad Sami Banai B.Sc.

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Psychologie
Betreut von | Supervisor:	Univ.-Prof. Giorgia Silani Privatdoz. PhD

Abstract

The increasing accessibility of immersive virtual reality has opened new avenues for psychological assessment, particularly in child populations. This systematic review synthesizes findings from 65 studies exploring how VR has been employed to evaluate various psychological constructs in children ranging from attention and executive functioning to emotion regulation, social information processing, and fear-related responses. The review identifies and categorizes key VR-based assessment frameworks, such as virtual classrooms, streets, and social simulations, and maps them against corresponding outcome measures, including behavioral, physiological, cognitive, and affective data. Findings highlight the strengths of VR in offering ecologically valid, interactive, and engaging environments that may overcome some of the limitations of traditional assessments. However, challenges remain regarding standardization, data interpretation, and individual variability in VR experience. This review underscores the potential of VR as a promising complementary assessment tool and emphasizes the need for further empirical validation and methodological refinement to ensure its robustness, particularly in clinical contexts.

Die zunehmende Zugänglichkeit immersiver virtueller Realität eröffnet neue Möglichkeiten für die psychologische Diagnostik, insbesondere bei Kindern. Diese systematische Übersichtsarbeit fasst die Ergebnisse von 65 Studien zusammen, die den Einsatz von VR zur Bewertung verschiedener psychologischer Konstrukte bei Kindern untersuchen – von Aufmerksamkeit und exekutiven Funktionen über Emotionsregulation und soziale Informationsverarbeitung bis hin zu Angstreaktionen. Die Übersichtsarbeit identifiziert und kategorisiert wichtige VR-basierte Bewertungsrahmen, wie virtuelle Klassenzimmer, Straßen und soziale Simulationen, und ordnet sie entsprechenden Ergebnismaßen zu, darunter Verhaltens-, physiologische, kognitive und affektive Daten. Die Ergebnisse unterstreichen die Stärken von VR: Sie bieten ökologisch valide, interaktive und ansprechende Umgebungen, die einige der Einschränkungen traditioneller Diagnostik überwinden können. Es bestehen jedoch weiterhin Herausforderungen hinsichtlich Standardisierung, Dateninterpretation und individueller Variabilität der VR-Erfahrung. Diese Übersichtsarbeit unterstreicht das Potenzial von VR als vielversprechendes ergänzendes Bewertungsinstrument und betont die Notwendigkeit weiterer empirischer Validierung und methodischer Verfeinerung, um ihre Robustheit, insbesondere im klinischen Kontext, sicherzustellen.

Table of Contents

1 State of the Art

1.1 Theoretical Background

A transformative revolution in mental health care is on the horizon, with virtual reality (VR) poised to take a leading role in this advancement (Freeman, et al., 2017). The definition of VR has evolved over time and remains a topic of discussion among researchers and clinicians (Rizzo & Shilling, 2017). Broadly, VR has been described as a technology that enables users to visualize, manipulate, and interact with complex digital data (Aukstakalnis & Blatner, 1992). Within this framework, VR represents a sophisticated form of human-computer interaction (Rizzo et al., 1997), offering users a more intuitive means of engaging with computer-generated three-dimensional graphical or photographic environments, which can be realistic or fictional and can reach levels beyond conventional input methods such as standard mouse and keyboard devices. Creating an engaging VR experience involves the integration of diverse interaction devices, sensory display technologies, and immersive content within the virtual environment. One of such devices is the head-mounted-display (HMD).

An HMD presents separate images to each eye, generating a stereoscopic 3D view. These images are rendered individually based on each eye's perspective within a mathematically defined virtual environment (VE). The HMD system typically incorporates head tracking, continuously capturing the user's head position and orientation to determine gaze direction. As users move their heads, the display updates in real time, usually at 60 frames per second, allowing for dynamic changes in the surrounding 3D scene and creating an immersive visual experience (Rizzo & Shilling, 2017).

A key mechanism of virtual reality is the immersion of the user into the VE. According to Jennett et al. (2008) the overall concept of immersion in games can be defined as the deep engagement in gameplay, resulting in a diminished awareness of time and the real world, alongside a feeling of being present within the game's environment. Clinical VR researchers have developed a wide range of VEs, ranging from airplanes, skyscrapers, fantasy worlds and battlefields to more mundane but important social settings like classrooms, homes, and supermarkets, designed to replicate real and imagined settings and enabling immersive simulations that align with the goals and procedure of specific clinical assessments or therapeutic interventions (Rizzo & Shilling, 2017). Recent technological advancements have made it possible to design highly realistic, interactive, and controllable VEs. Often referred to as the "Ultimate Skinner Box", VR has emerged as a powerful tool in both clinical and experimental psychology. VR enables immersive simulations for assessment, intervention, and performance measurement, offering capabilities that were previously unavailable (Rizzo & Koenig, 2017). Researchers consistently demonstrate that individuals react to these VEs in a manner that resembles real-life experiences (Gorini et al, 2010; Meehan et al., 2002).

An expanding body of scientific literature has begun to provide empirical support for the psychological applications of VR, documenting its outcomes and therapeutic effects across various domains (Rizzo et al., 2021). The systematic review of Freeman et al. (2017) identified 154 studies concerning the use of VR treatments for a range of mental health disorders. Other reviews focused more specific on the treatments of anxiety disorders or schizophrenia using VR (Carl et al., 2019; Maples-Keller et al., 2017; Rus-Calafell et al., 2018). Virtual reality has been utilized in the implementation of exposure-based therapies, which allows individuals to safely and systematically confront feared situations or environments within a controlled clinical setting (Bell et al., 2020). In a meta-analysis of nine randomized controlled trials that compared VR exposure with in vivo exposure for phobias, equivalent effects were observed for both interventions, which implies that similar underlying processes are at play (Wechsler et al., 2019).

The use of virtual reality has the potential for radical transformation not only in the landscape of mental health treatment but also mental health assessment (Bell et al., 2020). The measurement of psychological phenomena, encompassing symptoms, cognition, functioning and capacity, plays a pivotal role in both research and practice within the field of mental health (Baer & Blais, 2010). While researchers and clinicians can have a certain level of confidence in the current assessment instruments, it is important to acknowledge that there are numerous factors that pose threats to their reliability and validity (Myers & Winters, 2002). Results of traditional assessments like self-reports, observations from psychologists and paper-and-pencil cognitive assessments are often criticized regarding their generalizability to other life events (Parsons & Phillips, 2016). The disparities between the assessment context, such as laboratory or clinic, and the real world introduce several potential sources of bias, which can pose a threat to the accuracy of the results (Bell et al., 2020). While real-world assessments are feasible, they often come with significant costs and requirements and are limited in their access by factors such as geographical location and mobility constraints (Bell et al., 2020).

VR has the potential to overcome many of these limitations by providing highly controlled, immersive experiences that closely resemble real-world situations (Parsons, 2015). By exerting improved control over the perceptual surroundings, ensuring consistent presentation of stimuli, and enabling more precise and accurate scoring, the reliability of assessment can be enhanced within virtual reality environments (Parsons et al., 2019). By utilizing VR to immerse individuals in real-life scenarios, it becomes possible to conduct assessments that closely mimic the experiences encountered in everyday life (Bell et al., 2020). Virtual Environments can therefore also offer a potential answer to the issue of limited ecological validity (Parsons et al., 2019). Through the integration of immersive hardware such as HMDs, along with advanced input devices like eye trackers, gloves and body trackers, virtual reality has the capability to capture diverse performances in real-life tasks (Parsons et al., 2019). Virtual Reality has therefore also the potential to enhance the anticipation of a student's reaction to real-life challenges (Parsons et al., 2017). Different studies also imply the advantage of using VR-assessment as they are attractive to participants (Parsons et al., 2017).

1.2 Knowledge Gap

One key advantage of literature reviews is their ability to address broader theoretical questions that individual empirical studies cannot. By synthesizing patterns and relationships across multiple studies, literature reviews enable researchers to explore overarching themes and theoretical insights beyond the scope of a single investigation (Baumeister & Leary, 1997).

The first attempts have been made to give an overview of the vast literature about VR and its applications. Romero-Ayuso et al. (2021), who did a systematic review on the effectiveness of VR-based interventions for children and adolescents with ADHD, while Mesa-Gresa et al. (2018) focused in their systematic review on the effectiveness of VR for children and adolescents with autism spectrum disorder. Further, Freeman et al. (2017) did a systematic review of 285 studies on the potential of VR in mental health. These studies were divided into studies concerning assessment, theory development and treatment. In comparison to its other applications, the utilization of VR as an assessment tool has received relatively limited attention in the literature (Bell et al., 2020).

The beforementioned literature ranges from a review of virtual classroom environments for neuropsychological assessment from Parsons et al. (2019), which focuses on the assessment of cognitive functioning in neurological populations and puts a heavy emphasis on neurodevelopmental disorders and the use of virtual classrooms for assessing ADHD, to Geraets et al. (2022), who conducted a systematic review on the use of virtual reality in psychiatric diagnostic assessments to provide an overview in the field of different diagnoses. Other reviews focused on neurodevelopmental disorders (Satu et al., 2025), or on older populations (Tuena et al., 2020). As Geraets et al. (2022) suggests, future research may benefit from including a broader participant spectrum. Ranging from healthy and typically developing to clinical populations, and from applying more transdiagnostic perspectives.

The literature is missing a concise body of literature which focuses on a broader spectrum of younger children and extending the usual choice of children with neurodevelopmental disorders or psychiatric disorders. Further an overview is missing of VR as an assessment tool, which also includes outcome measures, which are not primarily relevant for clinical diagnosis. As Siddaway et al. (2019, p. 748) highlight "scientific literatures are now enormous or expanding exponentially, and knowledge is produced and shared rapidly across the world through the Internet ", it seems that systematic reviews can help to navigate through the complexity of the research topic by synthesizing the literature and providing an overview of certain topics. Therefore, this study aims to give a systematic review on VR as a tool for psychological assessment in late childhood.

1.3 Research Question and Objectives

Resulting from the existing knowledge gap and the missing of a systematic review about VR as a tool for psychological assessment in late childhood, this work aims to answer the following research questions.

1.How has virtual reality been used as a psychological assessment tool in late childhood?

2.Which psychological constructs have been successfully assessed using VR?

3.What are the advantages compared to traditional assessments?

The overarching goal of this systematic review is to synthesize current literature on VR as a tool for psychological assessment in late childhood and therefore provide a precise overview of the available articles in this specific research field. Ideally, this systematic review should answer the research questions by providing an illustration of the various ways of how VR as a tool is assessing various psychological constructs. Additionally, this review should give an overview of the advantages, disadvantages, and limitations in comparison to other assessment tools. This study wants to present its results also through a flow diagram to summarize the systematic literature search. Furthermore, this systematic review will not only summarize but also interpret the cumulative evidence and critically integrate the results to promote the understanding of the research topic as such critical thought and reflection are necessary for systematic reviews (Baumeister, 2013).

2 Methods

2.1 Systematic Literature Search

To answer the research questions a systematic review will be conducted to synthesize the vast literature. Siddaway et al. (2019) highlight six different key stages for conducting a systematic review, which are going to be explained in the following and adapted to this study. The last stage, Study Quality, will be skipped as this work aims to be a narrative review and not work with a meta-analysis, in which study quality plays a more important role.

2.1.1 Scoping

The first stage, Scoping, starts with the formulation of research questions. After formulating the research questions the breadth of the review should be considered and evaluated depending on the resources that the research group has. The researcher also must clarify whether a review has been done in this particular field of research. If the answer is yes, the researcher can continue their systematic review if the original systematic review is older than 10 years, was methodologically flawed and focuses on a certain aspect which is not the main focus of the review.
The research question of this study seems to fit the resources of the research team, as the research team consists of four members. No other review has been made in this specific area, therefore implying a need to investigate this research gap.

2.1.2 Planning

2.1.2.1 Search Terms

The second stage is named Planning and focuses on formulating adequate search terms, which operationalize the research questions. Creating a clear record-keeping system is also advised. Lastly, Siddaway et al. (2019) advise to adhere to reporting standards. The search terms that were used in this study are illustrated in Figure 1. These search terms were used in the combination with the words "OR" and "AND".

Table 1: Search Terms

Areas of Interest	Key Search Terms
Population	child*, adoles*, teenag*, young*, youth*
Application	Assess*, measur*, method*, evaluat*, test*, approach*, virtual reality

2.1.2.2 Inclusion and Exclusion Criteria

Inclusion and exclusion should be formulated to define clear boundaries of the review and ensure that only relevant work is included in the systematic review. These criteria should be justified by defensible grounds and revisited to reflect. The inclusion criteria for this systematic review are VR used as an assessment tool, children aged 8-12, psychological/psychophysiological measures, behavioral measure of psychological constructs. The exclusion criteria are conferences abstracts/

book chapter/ theses, which are not reporting on study aims, methods and results in a journal format, publications before 2013, language other than English, teenagers/adolescents specified as older than twelve or younger than eight, population of specific motor impairment (cerebral palsy, physiotherapy), targeting and measuring physiological problems (e.g. pain), not immersive VR environments, VR therapy without VR assessment and single case studies.

The idea to specify the age of 8-12 is based on the definition of the APA dictionary of psychology of what late childhood is. As this systematic review focuses on VR as an assessment tool studies were only included if outcome variables were measured in the virtual setting. Regarding the outcomes, purely physical outcomes like movement itself are excluded, except in cases where it is stated to be an indicator of a psychological construct. Further conference abstracts, book chapters and theses are excluded if they don't adhere to the standard of a journal article regarding its structure. Therefore, studies in conference abstracts were excluded if methods, objectives, and results were not reported. Due to the fast change in technological developments the systematic review only includes articles published after 2013 to avoid outdated VR applications. Populations with specific motor impairments were excluded because in these cases the VR application is usually used to only measure movement itself. Further, studies need to use a HMD to guarantee immersion and comparability. An online logbook has been constructed to create a clear record keeping system, which can be accessed by every member of the research team. To adhere to reporting standards the researchers of this study use PROSPERO.

2.1.3 Identification

The Third stage contains the Identification/Searching, whereby the researcher should at least search through two databases. Ideally the literature search should be conducted by two separate researchers to ensure inter-rater reliability. The following databases were used in this systematic search; Web of Science, PsycINFO on EBSCO, PubMed, and Scopus. To ensure inter-rater reliability two researchers identified the articles on the mentioned databases. Using the beforementioned search terms the first identification of studies led up to 6075 articles. After deduplication of the articles, 3150 articles were removed. Another three were removed from publication. In the end 2.922 articles were identified.

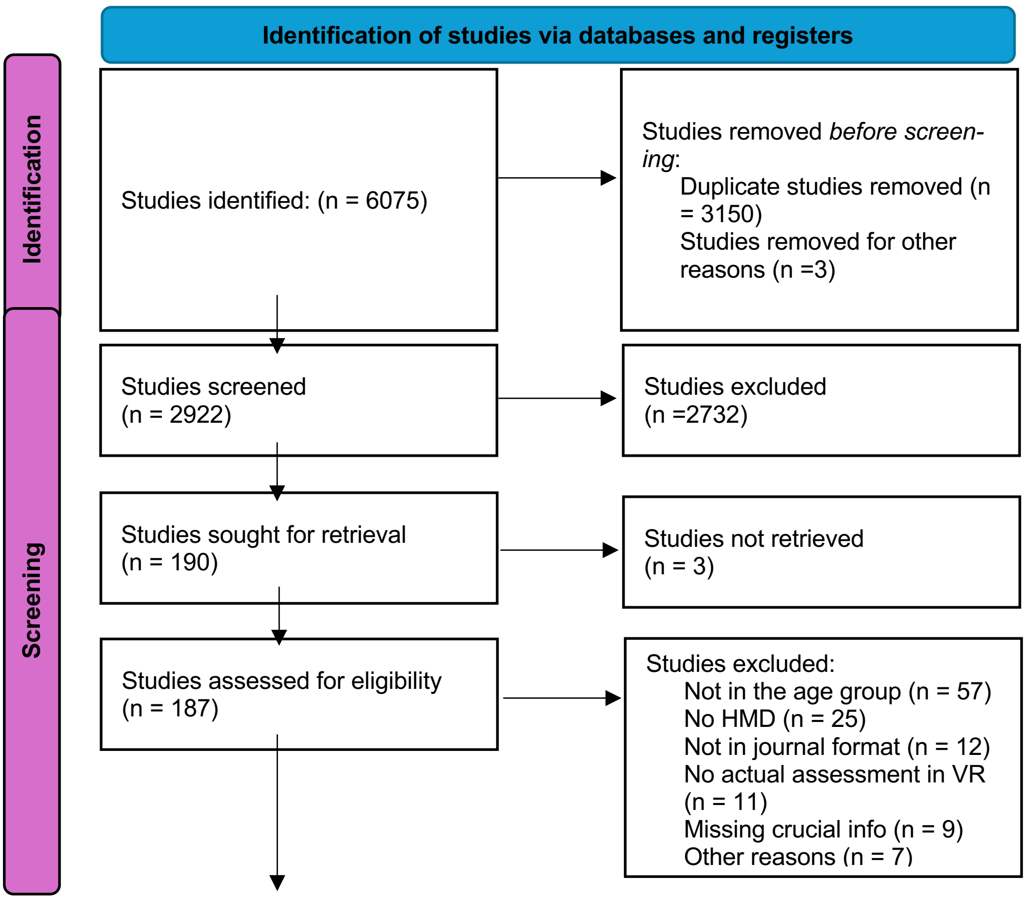
2.1.4 Screening

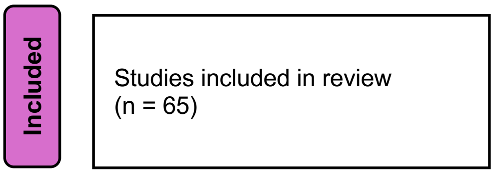
The screening process is mainly about reading the titles and abstracts to decide if the article should be included. Siddaway et al. (2019) advise to tend more on the side of sensitivity and include more to avoid missing important articles. The online platform, "Abstractr", was used to screen through the 2.922 abstracts with the established inclusion and exclusion criteria. 2732 articles were excluded based on their abstracts. Therefore 190 studies were checked for retrieval and eligibility.

2.1.5 Eligibility

Almost all 190 articles could be retrieved, leading to the extraction of all potentially relevant information by reading the whole article. Three studies could not be retrieved due to limited access. The following relevant information was extracted: Population, Number of Participants, Age Group, Study design, Duration, Purpose, Virtual Environment, Virtual Task, Scenario, Procedure, Main Findings, Outcome Measures, Limitations. After reading the articles 65 studies fit the inclusion and exclusion criteria. 57 studies were excluded because participants were either too young or too old, 25 studies didn't use HMDs, twelve studies did not report properly on methods, aims or results, ten studies had no actual assessment in VR and nine studies missed crucial information for the extraction and decision for inclusion or exclusion. Seven other studies were excluded because they measured constructs, which only refer to movement with no indication for psychological constructs (4), they had children with cerebral palsy (2), or the paper was not in English (1). Lastly, the Study Quality can be considered. Tools to assess study quality can help to gauge the appropriateness of the study design, although a study quality tool "should not be used to determine inclusion and exclusion from your review." (Siddaway et al., 2019, p. 765). As this study tries a narrative review approach, the study quality was not a primary focus.

Figure 1: PRISMA flow chart.





2.2 Narrative Review

Systematic reviews are often associated with meta-analyses. However, meta-analyses can also be applied to qualitative research using the same principles of transparency, rigor, and replicability. A narrative review is particularly appropriate when synthesizing quantitative studies that employ diverse methodologies or explore varying theoretical frameworks, constructs, or relationships (Baumeister, 2013). Unlike meta-analyses, narrative reviews focus on summarizing and interpreting findings without relying on statistical significance (Siddaway et al., 2019).

As Baumeister & Leary (1997) highlight this comparison between meta-analyses and narrative reviews, they underline the notion that a meta-analysis is typically preferred when numerous studies test the same hypothesis, as it serves a hypothesis-testing function. In contrast, narrative literature reviews are particularly valuable for synthesizing studies across varied topics, making them useful for theory development and hypothesis generation. One possible goal of literature review focuses on summarizing the current state of knowledge on a specific topic. These reviews aim to provide a comprehensive overview and integration of existing research without generating new theories. Further, these reviews are particularly useful for compiling established findings, such as in the context of grand applications or educational resources. As the goal of this work is to create a concise body of literature through a systematic review to help further researchers, a narrative review seems fitting.

2.3 Data Analysis

Baumeister & Leary (1997) in groups rather than commenting on each study individually as they can have similarities. The reader doesn’t need to know everything in which ways these are similar. In narrative syntheses, the number of studies included is often shaped by the scope and quality of the available literature. Unlike statistical meta-analyses, which can accommodate large datasets, narrative reviews require more hands-on analysis, making large numbers of studies more difficult to manage. To address this, studies are typically organized into smaller, more manageable groups early in the review process. These groupings may be refined as the synthesis progresses and are crucial for identifying patterns and facilitating meaningful comparisons. Importantly, grouping decisions should always align with the central research questions guiding the review. One way to group studies can be regarding the setting or context for the application (Popay et al. 2006). According to this approach of Popay et al. (2006) this review tries to answer the first research questions by grouping the 65 identified studies into shared VEs, as these VEs play a significant role in the assessment in VR.

To answer the second research question, it would be helpful to categorize the outcomes for better comprehension. For categorization deductive and inductive approaches can be used. According to Bell et al. (2020) clinically relevant outcomes can be divided into social functioning, cognition, and symptomatology, whereby these categories still need to be looked at more precisely, as these are not commonly used and just offer a first frame to categorize the different clinically relevant outcomes. Another possible categorization is the division of psychological constructs into psychological attributes such as processes or mechanisms (e.g., attention, memory), latent variables such as factors and traits (e.g. empathy) and operational definitions of behavior (e.g. walking distance) or physiological reaction (e.g. heart rate) (Slaney, 2017). Because of the relatively high number of studies included the outcomes are going to be categorized inductively. Further both virtual environments and outcomes are going to be tabulated as tabulation is a widely used method in systematic reviews for visually organizing both quantitative and qualitative data (Popay et al., 2006). This technique is especially helpful for providing an initial overview of the included studies and for identifying emerging patterns and relationships across the literature.

3 Results

3.1 Virtual Environments

The 65 identified studies can be divided into one of the seven categories, which can be seen in Table 1. A detailed tabulation of all of the studies can be found in the appendix at the end of the thesis. This following subchapter describes all of the VEs, their studies and the VR tasks used.

Table 2: Virtual Environments

VR Environment	VR Tasks or Topics	Population
Virtual Streets	Street Crossing Tasks, Hazard Anticipation Tasks on Bikes	TD, ADHD

Virtual Classrooms	VR-CPTs, InViRS, Aggression in VRC, Stuttering in VRC	ADHD, high IQ, TD, ASD
Virtual Apartments	EPELI, spEYEders, Fear of Spiders in Virtual Apartments	ADHD, TD, Children with anxiety disorder,
Gamified Virtual Environments	Why Did Baba Yaga Take my Brother, DEEP, Underwater Maze, BB and TTM	Children with Neurofibromatosis 1, TD
Neutral Virtual Environments	FBI Paradigm, Virtual Point to Origin Task, vrSocial, VR Lexical Decision Task, 3D Trajectory-Matching Task	ASD, children who stutter, TD
Multiple Virtual Environments	Memory, Learning, Stuttering, Fear Extinction Paradigm	ASD, children who stutter TD
Other Virtual Environments	Interpersonal distance, FBI Paradigm	Overweight children, TD

3.1.1 Virtual Streets

Pedestrian injuries represent a major global public health concern (WHO, 2017). Annually, traffic incidents claim approximately 1.2 million lives and cause injuries to 50 million people worldwide (World Health Organization, 2017). Children are particularly vulnerable, as traffic-related injuries are the leading cause of death for those over five years old (World Health Organization, 2018), with pedestrian injuries being the most common type. In 2017, Global Burden of Disease data reported that 34,000 children aged 5-14 died in road traffic crashes, with 25,000 of these fatalities occurring as pedestrian incidents (World Health Organization, 2018b). Another way to travel through traffic is transportation by bike. Despite its health benefits, cycling poses significant risks for young riders. Traffic accident data indicate that children up to 15 years old are the most frequently hospitalized cycling-related traffic victims (Nieuwkamp & Schoeters, 2018). Additionally, cycling is a particularly hazardous mode of transportation, with estimates suggesting that cyclists face a 23-times higher risk of serious injury per kilometer traveled compared to car drivers (Nieuwkamp & Schoeters, 2018). Maintaining the connection between perception and action is crucial for understanding how cyclists detect, anticipate, and respond to potential hazards (Zeuwts et al., 2023). Understanding the factors that heighten young children's risk of being struck by a vehicle while crossing streets is crucial for designing effective prevention strategies (Morrongiello

et al., 2019b). An effective way to identify these factors in safe settings could be the use of virtual street setups.
16 studies have been identified that use virtual street environments to assess outcomes related to pedestrian walking behavior in children. Further, two studies have been identified, which examine hazard anticipation on bicycles in virtual streets. Participants are immersed in environments consisting of roadways, which usually need to be crossed by the participants with slightly different objectives. In the case of the studies examining the hazard anticipation on bicycles the participants need to navigate through virtual streets, anticipating and avoiding threats. On one hand these virtual streets differ in size, traffic flow, vehicles and typical features of street environments such as buildings and trees which are placed on the sides of the streets. On the other hand, these studies differ regarding their instructions, number of trials and varying conditions such as gap sizes between the cars, vehicle speed and range of increments. Furthermore, the participants move differently through these virtual streets.

3.1.1.1 Virtual Street by Morrongiello and colleagues

The study of Morrongiello et al. (2015b) aimed to assess the ability of children to take evasive action, such as adjusting crossing speed, to prevent accidents (Morrongiello et al., 2015b). Further one study explored the factors contributing to injury risk for children (Morrongiello et al., 2016), while another compared street-crossing behavior between boys and girls to identify factors that may differently impact their risk of pedestrian injury (Morrongiello et al., 2020). Corbett and Morrongiello (2017) investigated how various measurement approaches influence the safety outcomes in child pedestrian research and their implication for both research and injury prevention.
Several studies explored the influence of different factors affecting street-crossing behaviors such as wearing a heavy backpack (Morrongiello & Corbett, 2021), the absence of auditory traffic cues (Morrongiello et al., 2021), time pressure (Morrongiello et al., 2015) and the presence of peers (Morrongiello et al., 2019b). Morrongiello et al. (2019) extended the aspect of peers in their study by examining whether children’s perception of their peers’ street-crossing behaviors influence their own behavioral norms and to assess the relationship between children’s self-reported crossing behaviors and their actual behavior in a virtual traffic environment. Furthermore, the study of Morrongiello & Corbett (2015) aimed to examine the alignment between parents’ expectations and their children’s actual street-crossing behaviors, assessing how accurately parents perceive their child’s pedestrian skills.
The most prominent virtual street setup that has been used consists of a two-lane road, 2.75 meters per lane, marked by a double yellow line in the center, flanked by sidewalks and houses set back

from the street (Morrongiello & Corbett, 2015; Morrongiello et al. 2015; Morrongiello et al. 2015 (b), Morrongiello et al. 2016; Corbett & Morrongiello, 2017; Morrongiello et al. 2019; Morrongiello et al. 2019 (b), Morrongiello et al. 2020; Morrongiello & Corbett, 2021; Morrongiello et al. 2021). Immersion is enhanced with features such as trees, shadows, textures, and traffic sounds, including the Doppler effect. Participants control their movement's direction and speed, allowing them to adapt their behavior while crossing. All traffic flows in a single direction, from the participant`s left, limited to the nearest lane (Morrongiello et al., 2020). If participants make a poor crossing decision, they have the option to step back onto the curb or accelerate to avoid the approaching vehicle (Morrongiello & Corbett, 2015).

Most of these studies name either two, three, four or five phases regarding data collection. Studies, in which participants each completed two phases, began with a familiarization phase, and ended with a testing phase (Morrongiello & Corbett, 2015; Morrongiello et al. 2015; Morrongiello et al. 2016; Corbett & Morrongiello, 2017; Morrongiello et al. 2019; Morrongiello et al. 2020). By implementing familiarization phases, the participants could watch the research assistants first, who demonstrate how to cross the street in the virtual environments via screen and then cross the virtual streets without cars by themselves to get used to the virtual environment (Morrongiello et al, 2020). Previous research has demonstrated that by the end of a 10-trial familiarization stage, children are fully accustomed to the VR equipment (Kennedy et al., 2000). Therefore, all these studies had ten familiarization trials, even when more phases are mentioned instead of the two. (Morrongiello & Corbett, 2015; Morrongiello et al. 2015; Morrongiello et al. 2015 (b), Morrongiello et al. 2016; Corbett & Morrongiello, 2017; Morrongiello et al. 2019; Morrongiello et al. 2019 (b), Morrongiello et al. 2020; Morrongiello & Corbett, 2021; Morrongiello et al. 2021). In phase 2 children were introduced to traffic and had to monitor the flow and cross the street when they judged it to be safe. An example for the testing phase can be seen in the study of Morrongiello et al. (2015). In phase 2 two types of trials were presented in a random sequence: (1) no time-pressure trials, and (2) time-pressure trials, in which participants were instructed the following:

"On trials when you hear a ticking clock, imagine that your close friend is already on the other side of the street, and she/he is about to leave. You need to speak to him/her right away about something important and you are worried that if you don't get across really soon, she/he will leave without you. So, you want to get across the street as quickly as you can but remember to go only when you feel it is safe." (Morrongiello et al., 2015, p.699).

Morrongiello et al. (2015b) named three phases, adding another in the beginning to estimate average walking speed and had participants walk between two green markers on the floor. Morrongiello et al. (2019b) used four phases for their study. The first phase included familiarization with children getting adjusted to the HMD and can practice walking without traffic. The second phase is meant to practice crossing with approaching cars. A baseline assessment is the main goal of the third phase leading to the last phase, which let the children repeat the phase three or watch a video of a peer crossing. Morrongiello & Corbett (2021) examined the effect of heavy backpacks on street crossing. Therefore, another phase was added after the familiarization measuring the street crossing in the presence of either traffic or a backpack or both. Another study also used five phases and examined the effect of street noise instead of backpacks (Morrongiello et al. 2021).
In some studies, the vehicles traveled at 50 km/h (Morrongiello & Corbett, 2015; Morrongiello et al. 2015(b); Morrongiello et al. 2019). 50 km/h is considered a moderate speed linked to an increased risk of pedestrian injury for children (Mueller et al. 1990; Roberts et al. 1995). Morrongiello et al. (2020) had vehicles traveling at either 30 or 70 km/h, whereas most of the studies included all three speeds (Morrongiello et al. 2015; Morrongiello et al. 2016; Corbett & Morrongiello, 2017; Morrongiello & Corbett, 2021; Morrongiello et al. 2021). Beside the vehicle speed, inter-vehicle gaps or gap thresholds, which are the average gap sizes they crossed into, differ between the studies (Morrongiello & Corbett, 2015). Some studies used three inter-vehicle gap sizes ranging from 3s and 5s to 7s (Morrongiello et al. 2015 (b); Morrongiello et al. 2016). Other studies used inter-vehicle gaps ranging from two seconds to six seconds with increments of one or 0.5 in between (Morrongiello & Corbett, 2015; Morrongiello et al. 2015; Morrongiello et al. 2020). The inter-vehicle gap size also varied from two to five seconds or from two to seven seconds with increments of 0.25 in one study (Corbett & Morrongiello, 2017; Morrongiello & Corbett, 2021, Morrongiello et al., 2021). Morrongiello et al. (2019) and Morrongiello et al. (2019b) expanded the inter-vehicle gap size even more, ranging from 1.5 till seven with increments of 0.25 or 1.5 till eight with increments of 0.75. The number of trials also vary in between these studies, ranging from 15 to 45 trials, depending on the amount of different vehicle speeds and inter-vehicle gap sizes used in the study Morrongiello & Corbett, 2015; Morrongiello et al. 2015; Morrongiello et al. 2015 (b), Morrongiello et al. 2016; Corbett & Morrongiello, 2017; Morrongiello et al. 2019; Morrongiello et al. 2019 (b), Morrongiello et al. 2020; Morrongiello & Corbett, 2021; Morrongiello et al. 2021).

3.1.1.2 Other virtual streets

Apart from the virtual street environment used by Morrongiello and colleagues six other virtual street setups could be identified. Luo et al. (2020) assessed the effectiveness of VR as a tool for identifying and reducing risky pedestrian behaviors among Chinese children. Wang et al. (2020) empathized on individual factors (sex and sensation seeking) and traffic conditions (vehicle speed and intervehicle distance) in relation to children' pedestrian safety. Stafford et al. (2022) focused on a specific variable, tau (a time-based factor), and how this variable influences decision-making regarding when and how to act in multi-lane traffic environments. Chung et al. (2020) compared children and young adults and how they adjust their walking speed when crossing roads under different traffic conditions. Tabibi et al. (2023) also examined two groups, by comparing TD children and children diagnosed with ADHD regarding their pedestrian skills and analyzing the relationship between pedestrian skills and cognitive functions such as attention, inhibition, and executive function in both. Parr et al. (2021) also explored the relationship between symptoms of ADHD and street crossing by investigating the relationship between inattention, hyperactivity, and motor timing difficulties, as well as how children's performance on basic timing task correlates with their behavior in a virtual road-crossing.

In contrast to Morrongiello and colleagues the virtual street environment of Wang et al. (2020) consists of a 5.6m wide road presenting bidirectional traffic with buildings, streetlights, and vehicles. Similar to beforementioned virtual setups, the sound of traffic in this virtual environment corresponds to vehicle distances and speeds, therefore also offering Doppler sound effects, which aims to increase immersion. In the case of cars hitting the participants the screen turns grey, and the sound of an ambulance sire is being played. Wang et al. (2020) present different conditions through varying vehicle velocities and inter-vehicle distances. Instructions were given verbally, and participants had practice trials, twelve experimental trials in total (Wang, et al., 2020)

Luo et al. (2020) confronted the participants with three challenges. In the first challenge the participants must safely cross a street to get soda at a convenience store. The second challenge requires the participants to decide whether to cross a wide street while the pedestrian crossing light flashes green, indicating insufficient time to complete the crossing. Challenge 3 introduces a hazardous situation where a school bus reverses toward the zebra crossing while the pedestrian crossing remains green (Luo, et al., 2020). This experiment consists of four trials.

Parr et al. (2021) used a scenario which consists of a 10-ft wide road with continuous traffic flow of 25mph coming from the left side. The temporal gaps between vehicles were presented in a random order, ranging from 2.5 to 5 seconds in increments of 0.5 seconds. Participants were instructed to cross the street without getting hit by a car in the process. Additionally, the children were instructed to wait for the purple car, the first vehicle in the sequence, to pass in front of them before initiating their crossing. They did two practice trials and twelve test trials (Parr et al., 2021). The Mobile Virtual Reality (MVR) replicates a real crosswalk near a local school and depicts a midblock crossing on a two-lane, bidirectional road marked with a crosswalk. In this simulation, children observe traffic flowing from both directions. Children press a button when they judge it safe to cross. Pressing the button initiates the crossing at a walking speed representative of typical child pedestrians. From a first-person perspective, the child observes the virtual traffic environment, where vehicles continue moving in both directions. This allows them to experience the outcome of their decision, whether successful or not. Positive feedback is given for safe crossings, while cautionary feedback is provided for collisions or near misses within one second of a collision. (Tabibi et al., 2023). Participants completed 21 street crossings in the MVR task, with seven crossings at each of three difficulty levels based on traffic speed and volume/density. Cars in the MVR appear at random intervals, predetermined by the system, and move at constant speeds (Tabibi et al., 2023).

Stafford et al. (2022) created a scenario which had a 6m wide road with a flow of oncoming traffic including 18 vehicles of bidirectional traffic according to UK traffic rules. This traffic flow included a specific gap where participants needed to decide whether to cross the road. Vehicle speed (32, 48, or 64 km/h) and distance (30, 40, 50, 60, 70 or 80m) were adjusted to create 18 distinct gaps. The time gaps between vehicles were designed to align with the average walking speed of an adult, ensuring that 50% of the gaps allowed for comfortable crossing and 50% did not. After the familiarization phase participants completed 13 practice trials where the sequence of gap closure rates was randomized. Following this practice session, participants undertook 54 pseudo-randomly presented trials, with each time gap presented three times. In each trial, participants were asked to assess whether a gap in the traffic flow provided enough time to safely walk across the virtual street. They were instructed to cross only when they perceived the gap as safe and suitable for crossing at comfortable walking speed. Participants were allowed to adjust their speed during the trial if necessary. If they determined that the gap was too small to cross safely, they were advised to wait for the trial to end and evaluate the gap in the next trial for crossing opportunities (Stafford et al., 2022). Participants were first instructed to cross the road five times without any traffic to familiarize themselves with navigating the virtual environment. At the beginning of each trial, they returned to the same starting point on the crosswalk, facing directly toward the opposite curb.

The virtual street used in the study of Chung et al. (2020) depicted a two-lane road, with each lane measuring 3.5 meters in width. The scene included trees, a skyline bordered by buildings, and a general street view of the roadway. Chung et al. (2020) used a treadmill for moving through the virtual environment. Chung et al. (2020) systematically altered three experimental variables: the participants' starting distance, the size of the inter-vehicle gap, and the vehicle size. Walking speed was assumed to range from 1.0 to 1.67 m/s for adults, approximately 1.17 m/s for children aged 6 to 12 years, and around 1.22 m/s from teenagers (Waters & Mulroy, 1999). For this experiment, an average walking speed of 1.1 m/s was assumed to calculate the initial distance. Participants began at distances of 3.5 m (near), 4.5 m (intermediate), or 5.5 m (far) from the curb to the point of interception with vehicles The gap between vehicles, treated as a dynamic entity (Chihak et al., 2010; Louveton et al. 2012), was defined as the space between the rear bumper of the leading vehicle and the front bumper of the trailing vehicle. The midpoint of the gap was programmed to arrive at the interception point four seconds after vehicles' appearance, equivalent to 33.2 m. Vehicle size was manipulated to reflect findings from Mathieu et al. (2017) on the influence of vehicle dimensions on crossing behavior. Simulations used either two white sedans (1.5 m wide and 3.5 m long) or two orange buses (2.4m wide and 11m long). Participants were tasked with crossing the gap between two vehicles traveling at a constant speed of 8.33 m/s (approximately 30 km/h) and walking until they reached the opposite side of the virtual road. At the start of each trial, participants viewed a black-and-white cartoon depiction of the virtual crosswalk to calibrate their perspective. Upon hearing a verbal "ready" signal, they prepared to cross, and at the "go" signal, the experimenter initiated the motion of the vehicles. Participants were required to immediately look left, assess the oncoming traffic, and cross if the gap appeared safe. To acclimate participants to the task and virtual environment, six practice trials were conducted: two free-walking trials without the HMD, two trials without vehicles, and two trials with vehicles traveling at a steady 25 km/h and separated by a 5-second gap. After the practice sessions, participants completed the task twice under each experimental condition. At the conclusion of each trial, feedback was provided through one of three outcomes displayed: success, collision, or failure. The term success was shown if the participant successfully crossed the gap and reached the opposite side of the road. Conversely, collision appeared if the participant collided with a vehicle, while failure was displayed if the participant missed the gap entirely. The trial presentation order was counterbalanced among participants, and each trial was repeated twice, based on findings from Plumert et al. (2011), which indicated short-term behavioral adjustments following specific road-crossing experiences.

3.1.1.3 Hazard Anticipation Task in Virtual Streets

Different to the prior virtual street setups, the following environments focus on hazard anticipation while cycling instead of the prior street crossing setups. The purpose of the research of Zeuwts et al. (2021) was to investigate the effects of induced mental fatigue on hazard perception and anticipation in young cyclists using a virtual reality bicycle simulator. Zeuwts et al. (2023) extended the research by aiming to develop an interactive and immersive VR-based hazard anticipation test for cycling simulators, assess its convergent validity and reliability, and evaluate its realism and potential for educational use.

The VR Hazard Anticipation Test presents children with 14 traffic-related scenarios designed to assess their ability to anticipate potential dangers while cycling (Zeuwts et al., 2021; Zeuwts et al., 2023). The hazards included in the test were selected in collaboration with three cycling safety education experts from the Flemish Traffic Engineering Foundation. Their selection was guided by Lachapelle et al. (2013) report on common mistakes made by child cyclists and previous research on hazard perception in young cyclists (Zeuwts et al., 2017). The test is divided into two sessions, each containing seven hazard scenarios and lasting approximately seven minutes (Zeuwts et al., 2021).

Hazardous traffic situations can be categorized based on their visibility, urgency, and level of threat. These distinctions help in understanding how cyclists and drivers assess and respond to different traffic hazards. (Lehtonen et al. 2017; Meir et al. 2015; Wetton et al. 2011).

The two studies used the following categorization:

1) Overt hazards involve visible road users engaging in dangerous behavior, such as a pedestrian suddenly stepping onto the road.
2) Covert hazards refer to potential dangers obscured by obstacles like parked cars or roadside vegetation, where a road user on a collision course remains hidden until the last moment.

Risks can also be classified by their time frame:

3) Abrupt or acute hazards such as a car door opening unexpectedly demand an immediate reaction.
4) Latent hazards develop gradually such as a pedestrian glancing over their shoulder before stepping onto the bike path, allowing for anticipation rather than instant action (Sahlberg et al. 2015; Vlakveld 2011; Wetton et al. 2011)

Lastly, traffic dangers can be distinguished by their level of threat:

5) Actual threats involve an immediate risk, such as a vehicle actively approaching on a collision path.
6) Potential threats arise in situations where visibility is obstructed, creating uncertainty- such as an obscured intersection where another road user may or may not appear.

Also, the cues which signalize can differ:

7) Visual cue signaling is an impeding hazard.

8) Auditory cue signaling an impeding hazard (Vlakveld, 2011; Zeuwts et al., 2017).

3.1.2 Virtual Classrooms

Attention-deficit/hyperactivity disorder (ADHD) is among the most prevalent psychiatric disorders in children and adolescents, affecting approximately 5-7% of this population (Thomas et al., 2015; Wang, et al., 2017). Beyond its core symptoms of inattention, hyperactivity and impulsivity, ADHD is also associated with secondary impairments in cognition, emotional regulation, and other related areas (Garbi, 2021). These challenges significantly impact academic performance, executive functioning, and social interactions (Powell et al., 2018). The diagnosis of ADHD is traditionally based on clinical interviews, self and observer reports, and rating scales. Further, the diagnostic criteria outlined in the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) and the International Statistical Classification of Diseases, 10th Revision (ICD-10) rely on assessing the presence of inattention and/or hyperactivity-impulsivity symptoms as reported by patients or their caregivers (American Psychiatric Association, 2013; World Health Organization, 1992). However, these subjective assessments have limitations in validity and reliability. To improve diagnostic accuracy, clinicians increasingly incorporate objective neuropsychological assessments alongside self-reported symptoms (Doyle, et al., 2005). Numerous studies have highlighted inconsistencies between objective measurements and subjective evaluations in ADHD diagnosis. (Manor, et al., 2008; Sumner et al., 2010; Choi et al., 2010; Corkum et al., 2001). Research suggests that incorporating objective measures could improve assay sensitivity in ADHD clinical trials (Sumner et al., 2010). Konrad et al. (2005) indicate with their study that teacher-rated ADHD symptoms were primarily influenced by motor activity changes and did not align with results from computerized attention assessments. Recently, there has been growing interest in developing objective evaluation tools for ADHD assessment. The Continuous Performance Test (CPT) is one such tool, requiring participants to respond to specific auditory or visual stimuli presented at regular intervals. Various CPT versions have been designed for objective ADHD evaluation. Research has shown that integrated visual and auditory CPTs (IVA) can effectively distinguish children with ADHD from their typically developing peers (Kim, et al., 2015). The study of Quinn (2003) indicates that individuals attempting to feign ADHD could manipulate subjective rating scales but were unable to do so on the IVA-CPT. This finding highlights the advantage of objective measures in accurately assessing ADHD. Fang et al. (2019, p. 1518) highlight this aspect as a demonstration of “the superiority of objective measurements in the evaluation of AHDH.”. Traditional CPTs are conducted in artificial settings, limiting their ability to assess real-world behaviors, such as those in a classroom. Studies have shown that most computer-based neuropsychological test have only low-to-moderate ecological validity (Parsons et al., 2007). Fang et al. (2019) further state, that CPT tasks are repetitive and require sustained focus, making them challenging and boring for children, especially those with ADHD. As a result, test completion rates, reliability, and validity may be affected, highlighting the need for more engaging assessment methods. To address these limitations, several VR applications have been designed and tested to assess cognitive functions like attention and executive function (Freeman, et al., 2017; Díaz-Orueta, et al., 2014; Iriarte, et al., 2016; Lalonde et al., 2013; Pollak et al., 2010; Rizzo, et al., 2009). Rizzo et al. (2000) are stated to be the "first to propose a VR classroom environment in order to improve the effectiveness of cognitive function assessment." (Yeh, et al., 2020, p. 1900). The Virtual Reality Classroom (VRC) environment (Rizzo, et al., 2009; Rizzo, et al., 2004; Rizzo, et al., 2000) was designed to replicate a typical classroom setting, which simulate an environment featuring desks, a female teacher positioned in front of a blackboard, two doors along the right-hand wall, and windows on the left-hand wall offering a view of the street. The virtual classroom provided a naturalistic simulation where virtual test challenges and distraction stimuli could be precisely controlled to facilitate performance evaluation across various experimental conditions (Rizzo et al., 2004). Other than the use of CPTs, virtual classrooms has been used to study joint attentional skills (Amat, et al., 2021), social information processing profiles (Verhoef et al., 2021; Verhoef, et al., 2022; Verhoef et al. 2022b) and stuttering in children (Boo et al. 2021).

3.1.2.1 VR- Classroom CPTs

The studies that have been included in this systematic review regarding the use of VR-CPTs have varying purposes. Many studies aimed to explore the feasibility and effectiveness of using VR as a tool for assessing ADHD symptoms (Fang et al., 2019) or develop an assessment tool to help clinicians in diagnosing ADHD (Yeh et al., 2020) by integrating multiple sensor technologies such as eye movement sensors and EEG sensors (Tan et al., 2019; Stokes et al., 2022). Hong et al. (2022) investigated the impact of VR distractors on attention and hyperactivity levels in children and adolescents diagnosed with ADHD. Bioulac et al. (2020) evaluated the effectiveness of a novel virtual classroom-based cognitive remediation program designed to alleviate symptoms in children with ADHD, while Coleman et al. (2019) aimed to introduce an ecologically valid approach to assessing the impact of working memory training on real-world attention performance. Several studies highlighted the influence of different factors on task performance in these virtual classrooms such as the introduction of environmental distractions (Stokes et al., 2022), such as the presence of a virtual teacher and social cues (Eom et al., 2019). Mühlberger et al. (2020) focused on performance differences by comparing medicated and unmedicated children with ADHD and a control group of typically developing children. Other studies focused on the link between ADHD symptoms and shifts in a child's field of view (Mangalmurti et al., 2020) and their head movements (Cho et al., 2022). The study of Negut et al. (2017) aimed to (a) assess the diagnostic validity of "ClinicalVR: Classroom-CPT" compared to a traditional CPT, (b) evaluate its task difficulty, (c) examine the impact of distractors on ADHD participants and healthy controls, and (d) compare both measures in terms of cognitive absorption.

The following virtual classrooms often differ in the number of stimuli, the number of distractors, the display duration of the stimuli and distractors and the intervals in between. The display duration differs from 100ms (Mühlberger et al., 2020) to 200ms (Negut et al., 2017), while the interstimulus interval can last between 1000ms (Negut et al., 2017) and 1900 ms (Mühlberger et al., 2020). The target stimuli, which were either numbers (Yeh et al., 2020), letters (Eom et al., 2019) or images (Stokes et al., 2022) and vary in amount from 320 (Mühlberger et al., 2020) presented stimuli to 500 (Stokes et al., 2022) presented stimuli. The tasks also differ regarding which buttons need to be pressed and what happens if you press the button at the wrong time.

To get a better understanding of the VRC, one study is going to be elaborated more on. The study of Mühlberger et al. (2020), which can be divided into two tasks, will be given as an example of a VR-CPT. The warm-up task involved the following steps: Upon the participant's arrival at the laboratory, the experimenter verified the completion of questionnaires and written informed consent forms. Next, the experimenter provided an overview of the VRC equipment and adjusted the HMD to fit the participant's head. The session began with a vision test, where the participant read letters displayed in the HMD. If no difficulties were observed in recognizing the letters, the experimenter initiated the VRC system, revealing the classroom environment. The participant was then instructed to spend one minute exploring the virtual classroom. Following this exploration, the virtual teacher assigned the first task, which consists of the warm-up activity. The participant viewed a sequence of numbers on the blackboard and was instructed to press the button on the response device whenever the number "5" appeared immediately after the number "9". Each number was displayed for 100 milliseconds, with an interstimulus interval of 1,900 milliseconds. A total of 20 number stimulus were presented, including five target sequences. If the participant struggled, the experimenter re-explained the task, and the warm-up trial was repeated. Successful completion of the warm-up trial, defined as no more than one error, allowed the participant to proceed to the main task after a brief break. In the main task, the virtual teacher instructed the participant to observe a sequence of letters displayed on the blackboard. The participant was required to press the response button as quickly as possible whenever the letter "K" followed the letter "A". A total of 320 letter stimuli presented, with each letter displayed for 100 milliseconds and an inter-stimulus interval of 1,900 milliseconds. The stimuli were divided evenly into four blocks, with each block containing 80 stimuli and lasting two minutes and 40 seconds. Each block featured ten target sequences presented at pseudorandomized intervals. Additionally, each block included a specific number of standardized distractors categorized as auditory (classroom sounds like whispering or pencil dropping), visual (a paper plane flying across the room), or a combination of auditory and visual (a noisy bus passing by outside the window). The design created varying levels of distraction: medium distraction in Blocks 1 and 4, high distraction in Block 2, and low distraction in Block 3.

1) Block 1 contained 26 distractors (twelve auditory, three visual, and eleven combined auditory/visual).
2) Block 2 included 51 distractors (21 auditory, eight visual, and 22 combined auditory/visual).
3) Block 3 had only five distractors (three auditory, one visual, and one combined auditory/visual)
4) Block 4 featured 28 distractors (13 auditory, four visual, and eleven combined auditory/visual).

Stokes et al. (2022) also used a virtual classroom, which was modeled after the original design developed by Rizzo et al. (2000). The updated classroom simulation from Strokes et al. (2022) incorporated features of modern class environments, emphasizing greater racial and ethnic diversity among the child avatars present in the virtual space. Stokes et al. (2022) extended the VR-CPT with a Stroop task and a Math task, which are going to be elaborated on more. The distraction that has been used in this study were consistent with those utilized in prior VR classroom studies (Rizzo, et al., 2009). The distractors fell into three primary categories, which differ from the beforementioned categories:

1. Window (Events occurring outside the left classroom window),
2. Student (Actions involving the student avatars),
3. Phone (Activities originating from two cell phones placed in the classroom).

The first distractor was triggered randomly between zero to ten seconds after the task began. After each distractor event, a 20-second cooldown period was observed before initiating the next distraction, which appeared within the subsequent zero to ten seconds. Each task block included a randomly selected set of distractors, most of which combined both auditory and visual elements. The duration of each distractor was determined by the longer of its auditory or visual components.

Participants engaged in three blocks each of the Stroop, Math, and AX-CPT tasks within the virtual classroom, during which subject responses and eye movement data were recorded.

1. The Stroop task displayed trial-by-trial on a virtual whiteboard (Lalonde et al., 2013), assesses executive functioning deficits and academic impairments commonly associated with ADHD, which contain greater risk for grade retention and lower academic achievement (Biedermann, et al., 2004). Stimuli included English color words shown in red, blue, or green ink in the center of the whiteboard, while the virtual teacher announced the ink color aloud. For congruent trials, the spoken color matched the ink color. During incongruent trials, the spoken color differed from the ink color. Participants responded via the Vive controller, pressing up for "Yes" (the teacher was correct) or down for "No". Instructions were presented using a PowerPoint slideshow before the task began in the VR headset. Each trial started with a fixation cross (1000 ms), followed by the word and auditory stimuli (2000 ms). Responses were required within this 2000 ms window. Each block consisted of 200 self-paced trials, with an equal distribution of congruent and incongruent trials.
2. The Math task adapted a match challenge previously designed for 7-9-year-olds (Rosenberg-Lee et al., 2011), aiming to simulate real-world classroom activities. Math performance, highly dependent on attention, was hypothesized to be influenced by distractibility. ADHD has been linked to academic impairments, particularly in math performance (Loe & Feldman, 2007; Efron, et al., 2014). Additionally, math tasks are often used in ADHD clinical trials to evaluate treatments effects (Wigal & Wigal, 2006; Faraone et al., 2007). Problems included simple and complex addition but excluded advanced problem-solving tasks. Participants also had to decide whether equations were correct by using a touchpad, while half of the trials featured correct equations.
3. The AX-CPT task replaced typical alphabet letters with child friendly images. Participants viewed sequences of images and were instructed to press the correct response button when the "X" image (dog) followed the "A" image (bear). Each block used a different set of developmentally appropriate images, such as animals, fruits, and colors. (Stokes et al., 2022).

Fang et al (2019) used the Virtual Reality Medical Center (VRMC) system to assess attention also through three distinct assessment testing modules. The VRMC designed for ADHD assessment consists of: Position Tracking (30 items), Stroop Test (41 items), and Object Recognition (60 items). Eom et al. (2019) designed two conditions based on the presence or absence of social cues from an avatar teacher. In the equipment control condition, participants completed the AX-CPT with letters displayed on a plain white screen, similar to a standard computerized CPT but delivered through the HMD. In the social cues condition, a male avatar appeared at the center of a whiteboard with CPT tasks displayed to his left or right. Participants were instructed to initially fixate on the teacher, positioned directly in front of them. The avatar provided visual cues by pointing and verbal cues, directing participants' attention to the indicated section of the whiteboard.

Cho et al. (2022) extended the social cue condition from Eom et al. (2019) by introducing three conditions, which varied based on how participants were guided to the screens on the right or left of the teacher avatar, whose upper body and face were visible. These cues were intended to guide participant's attention toward the stimuli in advance. Participants were required to complete CPT tasks presented on screens positioned to the left or right of the avatar. In the No-Cue Condition, participants completed the task without any directional prompt from the virtual teacher. This condition was designed to assess the influence of the teacher avatar's presence alone as a social cue. Trials were equally divided between the left and right screens, requiring participants to turn their heads to switch between them. The initial presentation of stimuli on either the left or right screen was counterbalanced across participants to ensure consistency. In the Visual Cue Condition, the teacher avatar used pointing gestures to indicate whether participants should focus on the left or right screen. In the Visual/Audio Cue Condition the teacher avatar combined pointing gestures with verbal prompts. The avatar said. "Look at me" to capture attention between stimuli and "Look at this" while pointing to the active screen. In both the visual and visual/audio cue conditions, the target screens alternated eleven times. The number of trials per screen was evenly distributed, and the sequence was counterbalanced across participants, adhering to the program's protocol. Other task parameters were consistently managed across all four conditions to ensure uniformity.

Yeh et al. (2020) introduces a novel VR classroom designed to assess sustained and selective attention, incorporating distraction events that include visual, auditory, and combined visual-auditory stimuli. Participants were tasked with completing two cognitive exercises, which include the beforementioned CPT and additionally an audio test, each lasting 6-8 minutes. The audio test aimed to assess sustained attention. In this test, participants listened to a series of digits and pressed the "confirm" button on the controller when two target digits appeared in a specific order. The order and digits were provided before each round, and participants had to identify the target digits correctly. If correct, an output of "true" was shown; otherwise, it displayed "false". Reaction times for each correct response were also recorded. The CPT was a visual task designed to assess selective attention. During the VR tasks, participants were exposed to various distractions (visual, auditory, or combined visual-auditory) to test their ability to maintain focus. Seven types of distractions events were introduced, including:

1) the teacher standing up, looking around, and sitting back down
2) a student yawning in front of the participant
3) a man walking down the aisle, opening the front door, looking around, and closing the door
4) a thunderstorm with accompanying thunder and lightning flashes outside
5) a car driving by and honking its horn
6) a fire truck with sirens passing by
7) a paper airplane flying across the user's field of view.

While the order of distractions events was fixed within each task, the order of events differed between the two task types. For each test, the first half was performed without any distraction events, while the second half included distraction events.

Tan et al., 2019 also used a combination of these two with the addition of the WCST. The audio test and the CPT are similar to the tasks in Yeh et al. (2020). The Wisconsin Card Sorting Test (WCST) assesses cognitive flexibility, problem-solving, and the ability to adapt to changing rules. Participants are given four stimulus cards and 128 response cards, each varying in number (one till four), color (red, green, yellow, blue), and shape (triangle, circle, cross, star). The sorting rule, based on number, color and shape, changes cyclically but is not disclosed to participants. They must determine the correct rule through trial and error. The test concludes when a participant correctly identifies the rule six times in a row or when all response cards have been used. The system records reaction time and accuracy for each sorting attempt.

Hong et al. (2022) adapted the Rapid Visual Information Processing Task (RVP), originally developed by Wesnes and Warburton (1984), into a virtual reality format to evaluate participants' sustained attention. The RVP task is widely recognized for its ability to assess sustained attention, working memory, and response inhibition across various psychiatric conditions, including ADHD (Maoz et al., 2018), bipolar disorder (Clark et al., 2002), depression (Weiland-Fiedler, et al., 2004), and schizophrenia (Hilti, et al., 2010). Gau and Huang (2014) further highlighted its potential as a cognitive endophenotype for ADHD research. In the VR-based RVP task, digits ranging from two to nine appeared sequentially at a rate of 150 digits per minute in random order displayed at the center of the classroom blackboard. Participants were instructed to press a controller touchpad as quickly as possible upon identifying any of three target sequences (3-5-7, 4-6-8) within an 800 ms response window. Hong et al. (2022) designed two distinct testing conditions, consisting of a no-distractor condition and a distractor condition. In the no-distractor condition, the VR-RVP task was conducted in a calm virtual classroom setting. Avatar classmates appeared attentive, focusing on the blackboard, the teacher avatar, or study materials on their desks. In the distractor condition, the same task was performed, but the environment included distractors. These distractors consisted of purely auditory, purely visual, or combined auditory/visual stimuli. The intervals between distractors varied randomly, set at 5, 10, 15, or 20 seconds, with each distractor displayed for five seconds. Over a 4-minute trial, there were eight distractor intervals and 18 distractors, presented in a randomized sequence regarding both their order and interval duration.

3.1.2.2 AULA Virtual Reality Test

AULA (Spanish for "classroom") is a virtual reality-based neuropsychological test designed to comprehensively evaluate attention processes and assist in diagnosing ADHD in children aged six to 16 years (Climent & Banterla, 2011; Diaz-Orueta et al., 2014; Zulueta et al., 2013). By simulating a realistic classroom environment, AULA provides a comprehensive assessment of attention, impulsivity, and hyperactivity, offering ecologically valid insights into ADHD symptoms and attention-related performance (Zulueta et al., 2019).

The test has demonstrated high test-retest reliability (Fernández-Fernández & Morillo-Rojas, 2012; Iriarte, et al., 2016) along with strong sensitivity and specificity (Rufo-Campos et al., 2012). AULA has demonstrated effectiveness in assessing inattention and distinguishing between neurotypical children and children with ADHD (Areces et al., 2018). AULA employs a CPT, incorporating various tasks and distractor conditions within a virtual school classroom setting (Climent & Banterla, 2011). The test comprises two primary exercises:

1. Non-X paradigm Task: participants press a button when they do not see or hear "apple".
2. X paradigm Task: participants press a button when they do see or hear "seven".

The AULA test provides several detailed scores such as scores comparing the non-X and X tasks, visual and auditory performance, performance with and without distractors and a total score (Diaz-Orueta et al., 2014). Stimuli are delivered visually on a classroom blackboard and audibly through headphones. Simultaneously, real-life ecological distractors - such as the sound of an ambulance, a child coughing, or the teacher walking around – are presented visually, auditorily, or in combination to mimic a realistic classroom environment (Zulueta et al., 2019). The order of task presentation, starting with No-X tasks followed by X tasks, can be deliberately chosen (Díaz-Orueta, et al., 2014). No-X tasks tend to cause overstimulation problems. Conversely, X tasks often lead to hypoactivation, manifesting as slower, more, variable, and less efficient answers (Artigas-Pallarés, 2009). This sequence is designed to better simulate the challenges faced by children with self-regulation difficulties, such as those with attention deficits (Díaz-Orueta, et al., 2014). This reflects the difficulty these children encounter when adjusting to new environmental demands after engaging in an overstimulating activity, as described by the state regulation model proposed by Sergeant et al. (1999).

Two studies focused on comparing the AULA to traditional CPTs by evaluating the convergent validity of these two (Diaz-Orueta et al., 2014), and the effectiveness of attentional variables in distinguishing ADHD cases (Rodriguez et al., 2018). Zulueta et al. (2019) had two main objectives: (a) to assess the external validity of the AULA by identifying its core factors and determining whether it accurately measures inattention, impulsivity, and hyperactivity, and (b) to examine AULA's accuracy in supporting ADHD diagnosis alongside subjective third-party reports. Zulueta et al. (2019). Other studies used the WISC-IV in addition to the AULA to compare cognitive variables and profiles (working memory and processing speed) and attentional profiles in students with and without ADHD, while also examining the predictive power of these variables in distinguishing between the two groups (Areces et al., 2018; Renati et al., 2022). Kakoulidou et al. (2021) extended the research by using a repeated-measures design exploring how text choice as an intrinsic motivator influences reading comprehension and enjoyment, comparing those with severe inattention to those without attentional difficulties.

3.1.2.3 InViRS

The study of Amat et al. (2021) introduces the design, development, and initial evaluation of the Interactive Virtual Reality System (InViRS), an adaptive, game-based platform aimed at improving joint attentional skills, specifically gaze sharing and gaze following (Amat, et al., 2021).

InViRS is a game-based system designed to help children with ASD practice gaze sharing and gaze following skills. While the system supports multiple game modes, its current version includes two main activities with a virtual avatar: 1. Tangram puzzle Game, which is used to practice 2. Bubble popping Game. Both games are utilized for pre- and post-assessment. Each game systematically assesses children's eye gaze responses to scaffolded prompts across multiple difficulty levels. InViRS offers various options for individualized and adaptive interaction with the child:

1. Provisioned varying gaze prompts: Adjusting gaze cues based on the task.
2. Least-to-Most prompting: Gradual delivery of prompts and visual aids, moving from minimal to maximum support.
3. Adaptive Interaction Module: The avatar adjusts its interaction level based on the child's performance.
4. Speed Variation of Gaze prompts: Changing the speed of prompts to test and develop the child's ability to follow gaze.

5. Real-Time performance Computation: Immediate assessment and tracking of game performance.

In InViRS, gaze sharing is defined as when a participant focuses their gaze within a specific area near the avatar's eyes, rather than directly on the eyes themselves. This approach allows gaze sharing to occur without causing discomfort or stress that individuals with ASD may experience from sustained, direct eye contact. (Yaneva et al., 2020; Tanaka & Sung, 2016). Amat et al. (2021) selected a minimum fixation duration of 200 ms, based on Rayner (1998), which identifies this as a typical characteristic of human gaze fixation. If a gaze lasts longer than 200ms, the avatar will initiate the next prompt by directing its gaze towards a game object, such as puzzle piece in the Tangram puzzle game or a bubble in the Bubble popping game. Amat et al. (2021) configured InViRS to allow a 30-second window for participants to focus on the avatar's eye region before advancing to the next step. This 30-second duration was selected in consultation with clinical psychologists specializing in ASD intervention to provide enough time for children to process the cue and respond to the avatar's prompt. A longer waiting period could lead to loss of focus or interest in the game. If a participant did not look at the avatar's eye region within 30 seconds, the system would trigger both audio and visual cues. In the Tangram puzzle (practice), a three-second bell ringing audio cue and a visual cue highlighting the avatar's eye region were used. In the assessment of the Bubble popping game, only the three-second bell ringing audio cue was played. If no eye contact was made within two minutes, the game ended.

Once a participant successfully shares their gaze with the avatar, InViRS prompts the avatar to direct its gaze toward a game object. The participant must then follow the avatar's gaze and direct their own attention to the game object to trigger the next event in InViRS. In the Tangram puzzle Game, after the participant looked at the correct object, its color was revealed, and they could then use the mouse to place the puzzle piece in the correct spot. If the participant failed to look at the correct game object within 30 seconds, InViRS activates assistive features from the Assistive Module to guide the participant. For example, the avatar would repeat the gaze prompt at a slower pace, and the puzzle piece would be highlighted to help direct the participant's focus. In the Bubble popping game, when the participant looked at the correct bubble, it would pop, and new bubbles would appear. If no gaze was detected on the correct bubble within 30 seconds, no assistive actions were triggered, and the avatar would move on to provide the next gaze prompt.

3.1.2.4 Aggression in Virtual Classrooms

Verhoef et al. (2021) address the issue that children often face difficult social situations, such as being laughed at, losing a game, or experiencing exclusion. How they mentally process these events shapes their subsequent behavior (Anderson & Bushman, 2002; Crick & Dodge, 1994; Lemerise & Arsenio, 2000). The social information processing (SIP) model outlines different key cognitive steps children go through before responding to these social situations. Over the years the SIP model has proven to be a valuable theoretical framework for understanding, preventing, and treating aggressive behavior problems (De Castro & van Dijk, 2017). Traditional methods for assessing aggressive SIP have notable limitations. Most studies rely on hypothetical vignettes, asking children to imagine and reflect on social situations as if they were real. However, this approach lacks ecological validity, as children may only exhibit aggressive SIP when they are emotionally engaged in actual social interactions (Anderson & Bushman, 2002). Strong emotions, such as anger or embarrassment, can activate aggressive thought patterns that might not emerge in calmer, hypothetical scenarios (Lemerise & Arsenio, 2000). Findings from a meta-analysis from Verhoef et al. (2019) indicate that children's aggressive behavior is more accurately predicted when SIP is assessed in real social interactions rather than through hypothetical scenarios presented in vignettes. Interactive VR offers a promising solution to overcome the limitations of traditional SIP assessments by immersing children in realistic social scenarios, enhancing emotional engagement, and improving ecological validity (Verhoef et al., 2022).

The pilot study from Verhoef et al. (2021) presents a novel interactive VR environment, which was developed to assess children's aggressive SIP, addressing limitations of existing assessment methods by providing standardized, emotionally engaging social interactions. Verhoef et al. (2022) continued their research by aiming to determine whether SIP profiles could be identified among boys exhibiting aggressive behavior problems. The virtual reality (VR) environment, developed by CleVR, featured a virtual classroom designed to resemble real-life setting, complete with established behavioral rules such as showing respect and friendliness towards others. Virtual peers in the classroom were portrayed as boys of a similar age group and average height, each with unique haircuts, facial features, and clothing patterns to enhance realism (Verhoef et al. 2021). The verbal responses of these virtual peers were pre-recorded by twelve children from theater schools, enabling consistent and standardized interactions across participants. These responses included general phrases ("I am ten. What's your age?") to maintain natural conversational flow. During interactions, the virtual peers displayed neutral emotional expressions, but their demeanor changed to show upset emotions if the participant exhibited aggression. Two engaging games were incorporated into the VR environment: building the tallest possible tower of blocks and attempting to knock over cans using five balls. Both games included high scores and bonus features to boost emotional engagement and maintain experimental control over gains and losses. Assessment scenarios revolved around these games, allowing participants to direct aggression toward the virtual peer (verbal insults or physical actions) or the peer's property (knocking over their tower). Game instructions, rules, and score tracking were clearly presented on a digital whiteboard and were verbally explained through pre-recorded standard instructions delivered by a female experimenter. Each scenario followed a structured sequence centered around the game being played by the participant and a virtual peer. Specific social events were introduced in a fixed order, though their exact timing was adjusted based on the participant's actions and progress during the game. At the conclusion of each scenario, the experimenter highlighted a key contextual event: either a provocation (the participant's game being disrupted by virtual character) or an instrumental gain (the virtual character winning the game).

Six scenarios were created: a practice scenario, a neutral scenario, two instrumental gain scenarios, and two provocation scenarios. These were all centered around one of the two games. The practice scenario familiarized participants with the VR environment by allowing them to play the game without any virtual characters. The neutral scenario assessed the participant's social information processing (SIP) in a neutral context involving small talk with a virtual peer, with no significant events occurring. The two instrumental gain scenarios focused on object acquisition, where participants had the option to steal a block or ball from the virtual peer for extra points, and competition, where participants could sabotage the virtual peer's game to win. The two provocation scenarios assessed SIP in response to social provocation, where participants were excluded from a game by two virtual peers, and object provocation, where the participant's game was deliberately sabotaged by the virtual peer (Verhoef et al., 2021).

3.1.2.5 Stuttering in Virtual Classrooms

While language impairments are not core diagnostic features of ASD or ADHD, age-appropriate receptive and expressive language skills are crucial for children's social and academic success (McIntyre, et al., 2017). Research comparing language abilities in school-aged children with ASD or ADHD to TD peers has produced mixed findings. Standardized assessments and narrative tasks using wordless picture books often indicate that structural language skills such as semantics and syntax, are comparable in these groups to those of mental-age-matched TD peers (Kim & Kaiser, 2000; Kuijper et al., 2017; Norbury et al., 2014). However, studies using more naturalistic, conversational contexts suggest that children with ASD or ADHD may exhibit less complex language compared to their TD counterparts (Losh & Capps, 2003; Redmond, 2004). This study compared the structural language profiles of verbally fluent school-aged children with ASD, ADHD, and comorbid ASD/ADHD to TD peers during a personal narrative task (Boo et al., 2022).

Participants engaged in a VR classroom experience using a HMD, which integrated head positioning and virtual avatar technology to assess social attention (Jarrold et al., 2013). A research assistant, acting as a teacher, introduced the virtual students and guided participants through a three-phase public speaking task:

1. Non-Social Phase – Participants viewed abstract lollipop-shaped targets.
2. Social Phase – The targets became human avatars introduced as students.
3. Higher-Demand Phase – Avatars gradually faded to 70% transparency if not fixated on but became opaque again when looked at.

Each phase lasted three minutes, during which participants answered self-referential questions about their daily routines, preferences, and experiences. The research assistant provided follow-up prompts to maintain verbal engagement. Participants typically answered 25 out of 38 possible questions, but analyses focused on 18 consistently asked across all participants. Audio recordings of responses were transcribed and verified for accuracy, ensuring consistency in data interpretation (Boo et al., 2022).

3.1.2.6 Social Stories in Virtual Classrooms

Halabi et al. (2017) conducted a usability study as a proof-of-concept to evaluate the feasibility of the designed system across three different interactive display types.

Halabi et al. (2017) designed a classroom “greeting” scenario centered on interaction between a virtual character and the user to enhance communication skills in high-functioning children with autism. This communication task module consists of two components: (1) the “greeting” scenario and (2) the 3D environment and characters used to implement it.

The “greeting” scenario was developed in collaboration with experts and educators at a local autism education center, incorporating their daily experiences with autistic children. The scenario draws inspiration from the “Social Stories” approach commonly used in autism-related VR solutions, providing a well-justified framework for the task. This scenario addresses a routine social interaction, which is greeting others, an interaction that occurs multiple times daily. The scenario is designed to assess and improve the user’s level of social engagement through repeated VR system usage.

The scenario begins with the user navigating virtually from outside the school, entering the building, and arriving in a classroom via the corridor. The classroom serves as the primary setting where the conversation task unfolds. Inside, a virtual teacher avatar, along with two virtual student avatars, stands near the chalkboard. The teacher initiates the activity by instructing all students, including the user, to participate in a greeting exercise. The virtual teacher avatar closely resembles actual instructors or therapists who work with the autistic students, while the classroom setting mirrors familiar real-world environment, which aims to foster relatability for the participants. The scenario is designed to align with the standard practices of traditional institutions and to address the unique challenges high-functioning autistic children face when initiating or responding to conversations with unfamiliar people.

Two interactive features are in the center of the scenario: Role-playing and Turn-Taking. The scenario emphasizes role-play to encourage responses from the user. Virtual student avatars demonstrate appropriate responses to greetings, modeling the behavior for the user without enforcing it. Halabi et al. (2017) aim to support observational learning, which helps the user understand how to greet back when prompted. To ensure a structured activity, each virtual student is called upon in turn to perform the greeting task. After the teacher provides instructions, each student is greeted and responds appropriately before the user is asked to complete the task. This approach aims to help the user learn essential social skills like responding to others and waiting for their turn in group interactions. By combining familiar settings, guided role-play, and structured turn-taking, the "greeting" scenario offers a supportive and effective method for enhancing communication skills in high-functioning autistic children.

The virtual environment was designed to implement the scenario in the following sequence:

1. Auto Navigation: The subject is guided automatically from outside the school, through the corridor, and into the classroom.
2. Welcome and Introduction: The virtual teacher introduces itself and the virtual student avatars.
3. Student Interaction 1: The teacher greets the first virtual student, who responds with verbal and nonverbal cues.
4. Student Interaction 2: The teacher greets the next virtual student, receiving a similar response, including a waving animation.
5. Subject Interaction: The teacher greets the next virtual student, receiving a similar response, including a waving animation.
6. Response Monitoring: The system tracks the subjects' voice and physical actions to record their response and response time.

3.1.3 Virtual Apartments

This group summarizes VEs which are portrayed in Virtual Apartments usually filled with either rooms with different functions or furniture, which is involved in the VR scenario. The studies using the EPELI VR-task (Seesjärvi et al., 2022; Seesjärvi et al., 2023; Merzon et al., 2022) use different rooms, which involve specific everyday scenarios task that need to be done in these rooms. In another virtual apartment participants were presented with the virtual stimuli on a table in front of them (Wechsler et al., 2021). Silva et al. (2022) offered the participants the ability to move around the virtual apartment and then being presented with the virtual stimuli also on a table placed in front of them.

3.1.3.1 EPELI

Human goal-directed behavior relies on continuous interaction with the environment, which is inherently dynamic, unpredictable, and filled with competing stimuli. Navigating such complexity requires adaptive cognitive processes, making the identification of real-life attention deficits particularly challenging (Risko & Kingstone, 2015; Gualtieri & Johnson, 2005). ADHD is one of the most well-recognized conditions where difficulties in regulating goal-directed behavior are central to its diagnosis (Barkley & Murphy, 2011; Barkley, 1991). Seesjärvi et al. (2022) developed the novel VR task EPELI (Executive Performance in Everyday Living) to quantify attention and executive function deficits in realistic, open-ended scenarios that mimic the contexts in which ADHD symptoms commonly appear. To assess goal-directed behavior and ADHD symptoms in real-life settings, Seesjärvi et al. (2022) evaluated the virtual reality task EPELI and its predictive, discriminant, and concurrent validity. Seesjärvi et al. (2023) continued the research on this topic by assessing key psychometric properties of EPELI, including its internal consistency, validity, and ability to accurately simulate real-life executive functioning challenges. Merzon et al. (2022) extended the research on the EPELI by collecting eye movement data from participants playing EPELI to enhance ADHD detection and gain deeper insights into the mechanisms underlying attention deficits.

The EPELI VR environment is designed as a virtual apartment with various rooms, including a child's room, living room, kitchen, open adult bedroom, utility room, and toilet/bathroom. The game involves completing 13 everyday task scenarios. Prior to starting, participants undergo a practice session where they learn to navigate the environment, interact with objects, and monitor time using a virtual watch displayed on the controller. A cartoon dragon character serves as a guide, introducing the practice session and providing instructions for each task scenario. Each scenario begins with the dragon giving verbal instructions for a series of subtasks (putting on clothes, eating breakfast, brushing teeth). The 13 scenarios are counterbalanced, with half the participants performing them in reverse order. Each scenario contains four to six subtasks, with instructions consisting of 30 to 66 words. In total, there are 70 subtasks: 52 can be completed at any time, 13 are time-based, and five are triggered by external cues (doorbell or phone tone). While participants are instructed to follow the subtasks order, completion order does not impact scoring. Each scenario lasts up to 90 seconds but ends earlier if all subtasks are completed correctly. Distractors, such as auditory (dog barking, coughing, traffic noises) and audiovisual elements (buzzing fly, running tap, TV program), are embedded in six to seven scenarios. Distractors are counterbalanced across participants along with the task order. Participants can turn off some distractors (running taps, TV, music), but others remain active throughout the scenario. The entire EPELI session lasts approximately 35 minutes. Afterward, participants complete a Repetition Task, where they verbally repeat instructions similar to those given by the dragon. This task evaluates the memory paradigm requires retaining and recalling instructions.

3.1.3.2 SpEYEders

Specific phobias, including spider phobia, are common in childhood and adolescence with animal phobias being the most reported, affecting around 8% of children annually, which are often cooccurring with other mental health disorders (Soo-Jin, et al., 2010). Exposure therapy is the preferred treatment for anxiety disorders (Wolitzky-Taylor et al., 2008) but integrating it into a game format may enhance engagement. Given children's positive associations with digital games, a VR-based serious game offers an innovative and appealing way to apply exposure therapy for spider phobia. The spEYEders project seeks to create a VR-based serious game to help prevent and treat spider phobia in school-aged children. Wechsler et al. (2021) designed an interactive gaze-based element in VR that presents users with a virtual spider capable of dynamically altering its appearance, transforming into different objects, or even speaking directly to the user. This interaction is facilitated through eye-tracking technology, where the system detects and responds to prolonged user gaze. Once the user maintains eye contact with the spider for a predetermined minimum dwell time, the system triggers an interactive response, enhancing engagement and immersion.

The task required participants to observe a virtual spider that appeared on a table in front of them. If participants maintained continuous eye contact with the spider for three seconds (for adults) or six seconds (for children), a specific visual feedback mechanism was triggered. This feedback was displayed for three seconds before participants were teleported to a different virtual room. In this new setting children rated their emotions using visual scales. The next condition then began with the idle spider, following the same process for each feedback variant. For children, feedback variants were shown in random order to account for potential changes in responses throughout the experiment. At the end of the VR session, children also answered open-ended questions. (Wechsler et al., 2021)

3.1.3.3 Fear of Spiders

As previously mentioned, phobias of spiders are cooccurring with different mental disorders (9). This paper examines (a) differences between school-aged children and adults in their perceptions and expectations of a VR phobic stimulus (a spider) before immersion (Study 1), and (b) the extent to which anxiety and phobic fear predict apprehension toward phobogenic VR stimuli (Study 2). The VR experience began in a virtual hallway leading to two apartments doors. Here, pre-recorded audio message introduced participants to VR and its role in treating childhood anxiety disorders, particularly in exposure therapy for spider phobia. Before entering the virtual apartment, participants answered several questions displayed on the HMD. These questions specifically reference two VR stimuli: a virtual spider (the experimental stimulus) and a virtual rabbit (the control stimulus), both used for exposure-based therapy. The spider was chosen as the experimental stimulus because it was previously used in the study by St-Jacques et al. (2010). The rabbit served as a neutral control stimulus, as prior clinical and experimental research has demonstrated significant differences in emotional valence between these stimuli (Cavanagh & Davey, 2003; Côté & Bouchard, 2005). After collecting apprehension data, participants explored one of the virtual apartments for five minutes. They could approach a virtual spider crawling on the counter. Before exiting the virtual environment, a final question appeared on the HMD, asking them to rate their fear of the virtual spider.

After listening to the pre-recorded introduction, participants were asked to describe their expectations regarding the virtual stimuli in the apartments. To evaluate their apprehension toward the virtual stimuli, a custom measure was developed, targeting five key dimensions commonly associated with animal phobias. Participants rated their apprehension on a scale from zero to ten, using a slide bar controlled via joystick buttons. The scale included five anchor points (from not dangerous to extremely dangerous). Each response had to be confirmed by pressing a button before the next question was displayed. The final version of the Apprehension Scale was composed of four items, measuring participants' expectations regarding how large, scary, dangerous, and disgusting the virtual stimuli are associated. The total sum of these items represented their overall apprehension score. After exploring the virtual apartment, a final question was displayed on the HMD to evaluate participants' actual fear response to the virtual spider they encountered (Silva et al., 2022).

3.1.4 Gamified VEs

Researchers used gamified VEs to investigate outcomes in a more playful way. These gamified virtual environments range from virtual imitations of the setting of the fairytale "The Magic Swan Geese" (Muravevskaia & Gardner-McCune, 2023), over virtual underwater worlds (van Rooij et al., 2016; D’Archangel et al., 2022) and rollercoasters (Oh et al., 2022) to virtual bridges (Chua et al., 2019).

3.1.4.1 Why Did Baba Yaga Take My Brother?

Developing empathy in children is crucial for reducing aggression, enhancing critical thinking, improving emotional regulation, and fostering interpersonal problem-solving skills (Borke, 1973; Cohen, 2001; Cotton, 1992). In developed countries, the time children spend on role-playing has significantly declined in recent years, which can play a role in fostering empathy (Alghamdi et al., 2016; Russ & Dillon, 2011). Some developmental psychologists attribute this decline to the growing use of digital technology among children (Ito, 2012; Veiga et al., 2016), raising concerns about its impact on social development. Conversely, researchers in Game Design and VR argue that VR technology can serve as powerful tool for fostering empathy in children (Apostolellis & Daradoumis, 2010; Bailey et al., 2019; Cummings & Bailenson, 2016). To assess VR’s potential for festering empathy in children aged 6-9, Muravevskaia & Gardner-McCune (2023) developed a VR Empathy Game framework based on participatory design research and early childhood development literature, followed by a qualitative evaluation of the resulting game. (Muravevskaia & Gardner-McCune, 2023).

The VR game is based on the Russian fairy tale “The Magic Swan Geese”, chosen for its emotionally rich characters and prosocial storyline. The player takes on the role of a sibling whose younger brother has been kidnapped, embarking on a journey to rescue him. The main antagonist, Baba Yaga, who kidnapped the brother, embodies anger, and her backstory reveals reasons behind her behavior. One way of how the backstory is revealed is by having the opportunity to learn about the past of Baba Yaga by looking through her glasses and investigating her memories. The player explores an enchanted forest, meeting three additional characters representing different emotions, which are happiness (Older Sister), sadness (Younger Sister), and fear (Goose). These characters provide the opportunity to ask questions, get to know behaviors related to their emotion and explore the perspectives of the characters. Further, the characters provide clues about how to interact with Baba Yaga using empathy actions such as asking about feelings, offering gifts, and hugging. Through these interactions, the player gradually builds a relationship with Baba Yaga, who eventually releases the kidnapped brother. The game begins without explicitly introducing the player’s character, allowing for an immersive first-person experience. Instructions guide movement and exploration, but the player must figure out independently how to retrieve their brother. The game’s core challenge is to recognize emotional cues and apply empathy-driven actions to resolve conflicts and achieve the final goal.

3.1.4.2 DEEP

A key factor in the development of anxiety disorders in children is physiological reactivity, which is the body's automatic response to stress, and is often reflected in changes in heart rate and breathing patterns (Antonovsky, 1979; Paul, 1983). Anxious children typically exhibit heightened arousal when faced with stressors and are more likely to avoid these situations rather than confront them (Weems et al., 2005). In contrast, physiological regulation refers to the ability to manage and reduce arousal levels (Dworkin, 1993). Breathing plays a fundamental role in this process with diaphragmatic breathing being a well-established technique for alleviating stress and tension (Sovik, 2000). Many evidence-based psychological treatments for anxiety incorporate relaxation and breathing exercises that encourage slow, controlled diaphragmatic breathing as a core component (Silverman et al., 2008). Van Rooij et al. (2016) state that despite the availability of evidence-based interventions, their effectiveness often varies, and long-term benefits remain limited. A key challenge is the lack of personalization, as existing approaches do not account for individual differences in responsiveness. To improve prevention and treatment outcomes for childhood anxiety, a new approach is needed, one that is both adaptive and supported by scientific methods to identify personalized profiles of change (van Rooij et al., 2016). The purpose of this study was to test a biofeedback virtual reality game, which summarizes therapeutic principles with intuitive learning process to strengthen anxiety regulation skills, and to examine if the game could reduce anxiety levels and negative affect and increase positive affect.

DEEP is a VR game designed to immerse players in a serene, underwater fantasy world where they can freely explore at their own pace. Unlike traditional games, DEEP does not impose specific objectives or challenges but instead focuses on fostering relaxation and mindfulness. A core feature of the game is its integration of diaphragmatic breathing exercises through biofeedback. A stretch sensor detects the player's diaphragm expansion during inhalation, with resistance decreasing as the diaphragm expands. This data is processed by a microcontroller and transmitted to the game, influencing various elements of gameplay. One immediate form of feedback is a visual representation of breathing, which can be portrayed as a dynamic circle that expands and contracts in sync with the player's breath, promoting awareness of their breathing patterns. Additionally, the player's breathing directly influences movement within the game. When lung capacity falls below 50%, gravity takes effect, subtly grounding the player. The direction and intensity of movement are also controlled by the breathing pattern, with inhales generating an upward force, while exhales provide an additional push. If the player is above the ground, inhaling results in forward motion rather than an upward lift. This dynamic interaction between breath and movement reinforces the practice of slow, deep breathing, making it an intuitive and engaging part of the gameplay experience. By integrating controlled breathing movement mechanics, DEEP offers an embodied and compelling incentive for diaphragmatic breathing.

3.1.4.3 Virtual Rollercoaster Ride

Behavioral interventions and psychological therapies, such as cognitive training, behavioral parent training, and social skills programs, are commonly recommended for children with ADHD. Among these approaches, VR-based cognitive training has gained attention due to its ability to engage users and create dynamic, interactive learning and clinical environments (Oh et al., 2022). This study evaluated the effectiveness of VR cognitive training for children with ADHD by measuring their sense of presence and EEG responses.

Oh et al. (2022) based their cognitive training on the Advanced Test of Attention (ATA), a diagnostic tool that assesses perceptual tracking, psychomotor speed, sequential processing, memory, attention, and concentration. The VR scenario was designed to strengthen behavioral inhibition, requiring participants to complete tasks in a virtual environment. To enhance engagement for children, the intervention incorporated a rollercoaster-themed VR experience, where participants followed instruction and targeted objects while riding a virtual outdoor rollercoaster.

The VR environment was designed to assess behavioral inhibition using various distraction stimuli, including objects and virtual avatars. The scenario featured an outdoor train ride, modified with gamification elements, and presented in a roller-coaster format. Before starting, participants selected the duration for each stage (10-50 seconds). The training introduced distraction stimuli that met specific conditions, while measuring positive responses, omission errors, and false alarm errors.

3.1.4.4 Build a Bridge and the Time Travelling Mailman

Previous research on VR in classroom learning has primarily explored non-immersive desktop VR (Hew & Cheung, 2010; Merchant et al., 2014), given its technological maturity and ease of implementation in schools (Mikropoulos & Natsis, 2011). However, the increasing accessibility of IVR has expanded its potential for education (Olmos et al., 2018). Affordable mobile VR headsets, utilizing smartphones for high-resolution, real-time content display, have made IVR more viable for classroom use. Presence, or the feeling of “being there”, is a key advantage of IVR (Steuer, 1992). While researchers have suggested a link between presence and learning (Mikropoulos & Natsis, 2011; Winn et al., 2002), empirical studies have yet to confirm that a stronger sense of presence directly enhances learning outcomes (Makransky et al., 2019). In recent years, IVR has emerged as a promising approach to technology-enhanced education, offering new possibilities for interactive and engaging learning experiences (Checa & Bustillo, 2020; Snelson & Hsu, 2020).

The first Concept "Build the First Bridge (BB)" has the learning goal to develop an understanding of structural design principles, with a specific focus on the history and construction of the first bridge. Students engage in an interactive simulation where they must construct the first bridge across a river (a structure that no longer exists today).

1) Interactive Version: Students select materials and structural components to build the bridge. The sequence of construction steps is crucial – if a component is placed incorrectly (selecting the bridge deck before installing support piles), the bridge fails, demonstrating the importance of proper planning. If an incorrect step is chosen, an audio guide provides feedback and suggestions, such as: "Try adding supports for the deck first." Once materials are correctly placed, additional narration explains their historical use ("Wooden piles were driven five meters into the riverbed for stability.").
2) Passive Version: Students observe the bridge being built while an audio narration delivers the same instructional content.

In both the Magic Window (MW) and Immersive Virtual Reality (IVR) conditions, students have control over their perspective, similar to a 360-degree video. However, they cannot directly interact with the environment or alter the pace of the construction process. In the screen condition students don't have any control over their viewpoint (Chua et al., 2019).

The learning goal in the second Concept "The time Travelling Mailman (MM)" is to understand the purposes for which structures are designed and what factors contribute to a well-designed structure. Students are transported to the 1800s, where they take on the role of a mailman tasked with delivering a letter and parcel across the river at two distinct time points. In the first time period, the bridge has not yet been constructed, and students must use a punt to cross the river. At the second point, the bridge is in place, but they must pay a toll and wait for the swing span to open and close in order to allow a boat to pass before they can cross.

1) Interactive Version: Students are given the freedom to explore the environment at their own pace. They can select objects and teleport them to different locations using a button press or screen tap.
2) Passive Version: Students experience the events unfold from a first-person perspective, with audio narration providing details about the punt, the bridge, and the historical context of each situation.

3.1.4.5 Virtual Reality Water Maze

Neurofibromatosis Type 1 (NF1) is a genetic disorder affecting approximately 1 in 3000 individuals (Friedman, 1999). Disruptions in neurofibromin function may lead to increased GABA transmission, potentially impairing long-term potentiation and affecting learning and cognition (Anastasaki et al., 2022). Although cognitive and academic difficulties in NF1 vary, around 70% of individuals struggle with academic performance (Doser et al., 2022; Lehtonen et al., 2013). One of the most consistent challenges in NF1 is visuospatial impairment, often considered a hallmark characteristic (Hyman et al., 2005). However, research on visuospatial deficits in NF1 remains limited, particularly regarding navigation and learning in complex environments. A previous mouse model of NF1 showed visuospatial learning deficits in the Morris Water Maze task, a widely used spatial learning test in animals (Costa et al., 2002). Further exploration of visuospatial skills using more advanced navigation-based tasks could provide deeper insights into the cognitive profile of individuals with NF1. This exploratory study investigated visuospatial processing skills using a virtual 3D Morris Water Maze task, with a focus on individuals with NF1 and reading difficulties (D'Archangel et al., 2022).

The HVS Water Maze is a virtual reality adaptation of the Morris Water Maze task (Morris, 1981), designed to assess spatial learning and memory in humans. Similar to the original animal-based version, participants are virtually placed in one of four quadrants of a 180 cm diameter water tank and must navigate to a hidden static platform. Using either the handheld controllers of the HMD or an Xbox controller, participants "swim" through the tank, guided by landmarks outside the tank, resembling a laboratory environment. Throughout the trials, session data such as path length and latency are recorded to analyze performance.

3.1.5 Neutral VEs

In contrast to other groups this cluster involves VEs with which are not primarily constructed by social settings of functionality like the virtual streets, virtual classrooms, or virtual apartments. Further these VEs tend to be more realistic than the VEs in the gamified group.

3.1.5.1 Full Body Illusion Paradigm

To explore how children and adults are differently influenced by avatar appearance and visuomotor coherence, Weijs et al. (2021) employed a within-participants design. The avatar's visual appearance was manipulated to resemble either a gender-neutral human figure or a skeleton. Additionally, visuomotor synchrony was varied by having the avatar's movements either align in real-time with the participant's movements or follow with a delay of 0.5 to 1 second. Embodiment was evaluated through an explicit measure, which consists of an embodiment questionnaire assessing participants' sense of ownership and agency, and also an implicit measure, which were recorded Skin Conductance Responses in response to a virtual threat. Participants viewed the avatar from both a mirror perspective and a first-person perspective. In one condition, the same avatar was used across the following stages: the embodiment induction, the interactive game, and the virtual threat scenario.

1) Embodiment Induction: participants observed the avatar as it moved, switching between first-person and mirror perspectives.
2) Interactive Game: participants interacted with white cubes that turned yellow upon successful interaction and could freely toggle between the two perspectives.
3) Virtual Threat: A diamond-shaped object descended slowly from the ceiling, passing through the avatar's body. It was first visible in the mirror perspective and later in the first-person perspective, with participants able to switch perspectives freely.

At the end, participants answered a six-statement embodiment questionnaire displayed on the HDM, using head movements to select their responses.

The experiment began with a 40-second embodiment induction phase, during which participants followed verbal instructions delivered through the headphones of the HMD. Participants were directed to lift and lower their arms and legs one at a time while observing either their virtual body or a virtual mirror positioned above a reclining chair, which reflected the full avatar. This was followed by a 90-second interactive game phase designed to fully engage participants with their virtual body and encourage active use of it. In this phase, participants were tasked with reaching for virtual cubes using their arms and legs. Successful interactions, which refer to when a tracker collided with a cube, caused the cube to change color from white to yellow. Regardless of success, each cube disappeared after five seconds. Further, a virtual threat was introduced to assess participants emotional reactions through skin conductance responses (SCRs). This threat took the form of a diamond-shaped virtual object that descended from the ceiling and passed through the virtual body. A trigger was sent to the electrodermal activity (EDA) recording software as soon as the object began its descent. Approximately six seconds later, the threat became visible to the participants, depending on where they were looking. The object struck the virtual body eight seconds after the trigger and fully passed through it by the 10-second mark. Following this event, the virtual body remained visible for an additional ten seconds without any movement prompts to ensure that the SCR was not affected by movement artifacts or visual changes. This resulted in a total embodiment phase duration of 150 seconds. Following the conclusion of the virtual threat phase, participants were presented with a six-item embodiment questionnaire displayed on the HMD. They responded to the questionnaire by selecting points on a Visual Analog Scale (VAS) using head movements. The questionnaire, derived from research on using the full body illusion and the rubber hand illusion, which was mentioned before (Cowie et al., 2018; Leggenhager et al., 2007, Cowie et al., 2016), assessed “subjective feelings of embodiment, overall enjoyment, and strangeness” (Wejs et al., 2021, p.6).

3.1.5.2 Virtual Point to Origin Task

Spatial updating, the ability to track one’s location relative to the environment while moving, relies on both piloting (visual landmark cues) and path integration (self-motion cues). Navigation is complex as it requires integrating multiple sources of information. Research suggests that adult females perform well in landmark-rich virtual environments but struggle when self-motion cues are removed (teleportation) (Barhorst-Cates et al., 2020). However, it remains unclear whether this applies to landmark-free environments or children. Barhorst-Cates et al. (2021) use VR to investigate spatial updating in young adults and 10- to 12-year-olds, aiming to determine (1) which self-motion cues are necessary, (2) whether children rely on different strategies than adults, and (3) if these effects generalize across different environments. Through two experiments, this study assessed how adults and children use dynamic self-motion cues to complete a point-to-origin task (Barhorst-Cates et al., 2021).

The experiment consisted of three conditions: full-dynamic (walking), visual-dynamic (controller induced), and no-dynamic (teleporting). The full-dynamic condition involved real walking at a comfortable pace, while the visual-dynamic condition required movement via a controller trigger. The no-dynamic condition involved participants teleporting by pointing with the hand controller to the goal position, while pressing the thumbpad would show the “arc” of the trajectory and releasing the thumbpad would lead to teleportation. Each trial began with participants locating a starting position (marked by a colored circle and pole) and memorizing it. The participants then traveled to two sequential target poles, which disappeared upon arrival, with a beep signaling their completion. After reaching the second pole, the screen turned black, and participants oriented back toward the starting position, verbally confirming readiness. The experimenter recorded their estimated position, then prompted a small step forward to reinforce the need for accurate spatial estimation. While the screen remained black, participants were led on a random circuitous path before starting the next trial.

3.1.5.3 vrSocial

Social communication relies heavily on nonverbal cues such as eye contact gestures, facial expressions, body language, and tone of voice. While these cues are essential for face-to-face interactions, they can be particularly challenging for individuals with autism. Innovative technologies, including IVR, offer promising solutions by enhancing social skill interventions. VR can provide real-time, interactive environments that concretize nonverbal communication, help individuals better understand and generalize these skills in social interactions (Boyd et al., 2018). The primary objective of this work was to assess the feasibility of immersive VR for children with autism by designing and evaluating three nonverbal communication applications.

Boyd et al. (2018) developed a multi-user game called vrSocial, which offers real-time feedback on key conversational dynamics, including proximity to other users, the user's speaking volume, and the time each player spends talking. This intervention effectively reveals previously hidden information about a user's participation in the conversation, fostering greater awareness and engagement. The intervention consists of four different versions and conditions.

The first condition offers no visual feedback on social skills. Users interact in an open space where they see and converse with the other user's static, floating avatar. This setup served as an initial exploration of the virtual environment and acted as the experiment's baseline condition. The second version includes a visualization of proximity. Concentric circles appear on the floor beneath each user: red for personal space, blue for appropriate social space, and yellow as an intermediate warning. Socially acceptable distance is maintained when the blue circles of both users touch, which corresponds to arm's length in physical space. If users move out of each other's social space (beyond the blue circle), the screen becomes 60% opaque white with a text cue prompting them to "step closer". Conversely, if a user enters another's yellow or red circle, the screen turns yellow with 50% opacity and displays warnings like "warning" or "step away".

The third version of the game provides visual feedback on participant's speaking volume and the time they spend talking. Speaking volume is monitored through the headset's microphone, with a pre-calibrated target range established by the research team. Each user receives feedback on their own volume but not the other's. A small grey speaker icon appears when the user is silent, a larger green speaker when the volume is appropriate, and a large red speaker if the user becomes too loud. Additionally, a "time spent talking" bar tracks conversation contributions. As users speak, the bar fills with orange from the left, while the other user's contributions fill the bar with purple from the right over a 60-second trial. The final version integrates visual feedback on all three social skills: proximity, speaking volume, and time spent talking. This condition aims to assess whether participants respond better, similarly, or worse when managing multiple feedback elements simultaneously.

3.1.5.4 VR Lexical Decision Task

Modern HMDs can record head and eye movements, providing direct insights into reading fluency and the effects of linguistic complexity (Mirault et al., 2020; Rayner, 1998). Traditional eye-tracking methods require fixed laboratory setups, making them impractical for classroom use. Further, reading assessments in school settings are often affected by external distractions. VR's immersive nature minimizes these distractions, creating a more controlled environment. Studying visual, orthographic, and attentional factors in reading (Facoetti et al., 2010; Ziegler et al., 2010) requires strict control over stimulus presentation, including fixed screen distances that determine visual angles and text size, which is something VR can achieve effectively (Rizzo & Buckwalter, 1997). The affordability of VR technology in recent years (Ray & Deb, 2016) makes large-scale implementation in schools feasible. It is essential to determine whether eye movement and reading assessments conducted via VR are reliable and externally valid. Despite VR's potential little research has focused on its use with primary school children (Eleftheria et al., 2013). Additionally, it remains unclear whether VR-based reading assessments can replicate classic laboratory findings, such as lexicality, word frequency, and length effects (Perry et al., 2010). Mirault et al. (2021) aimed to establish proof-of-concept for using a VR setup to measure reading fluency and track eye movements during silent reading in primary school children.

In this study, children completed two tasks: a VR Lexical Decision Task (VR-LDT) and a One Minute Reading aloud test, with task order counterbalanced between participants. For the VR-LDT, children sat at a school desk allowing free head and torso movement. Testing took place in a quiet room adjacent to the classroom, with two children assessed simultaneously- one performing the VR-LDT while the other completed the OMR test. The VR-LDT was introduced as a word detection game, where children identified real and nonexistent words. Each trial began with a 1000 ms fixation dot, followed by the stimulus -black monospaced text displayed in a neural VR environment (brown floor, blue sky). Participants read the word and responded using a gamepad, with pushing the right trigger for real words and the left for nonexistent words. The stimulus list, which consists of 100 items, was randomized but identical for all participants. The experimenter provided guidance through oral examples and corrections, but there were no practice trials or real-time feedback.

3.1.5.5 VR 3D Trajectory-Matching Task

Motor exploration, which is the variability in strategy use following failure, is essential for motor learning (Schulz, 2012). When children practice a new skill and experience failure, they must explore alternative strategies to achieve success (Thelen, 1995). Despite being a growing research area, little is known about the factors influencing motor exploration in children or how parents can support this process. Therefore, Brummelman et al. (2022) analyzed the impact of parental praise on children's motor exploration.

Children participated in a 60-trial VR 3D trajectory matching task designed to measure motor exploration and movement variability following failure. The participants were placed in a virtual rectangular room with a central pole and a grey knob, where the children hat to replicate a straight 3D line using a handheld controller. The starting point (grey knob) was adjusted to 80% of the child's height to ensure equal difficulty across participants. Before the main task, children familiarized themselves with the VR setup through two practice phases: 1. Free drawing, where the children saw their movements as a pink floating line., 2. Copying a straight 3D line, followed by five practice trials with varied angels and visible drawn paths to illustrate how their movements were recorded. In the main task, children completed 60 trials, each requiring them to copy a standardized straight 3D line. No visual feedback was provided during drawing, forcing reliance on exploratory movement adjustments. After each trial, participants received success or failure feedback for 500ms. Success was indicated by a green target and a flowing "ka-ching" sound, while Failure was indicated by a red target and a buzzer sound. To ensure a comparable number of failures across children, and adaptive success criterion was applied based on each child's past performance. The first trial always resulted in failure, and for the next four trials, the success threshold was adjusted based on the child's median error from available trials.

3.1.6 Multiple VEs

This group contains studies, which focuses on the use of multiple VEs. Participants are moved from virtual cities to virtual islands (Cadet et al. 2022), virtual streets and virtual bus rides (Almazaydeh et al., 2022), virtual apartments and virtual classrooms (Moise Richard et al. 2021) and the change of virtual indoor hallways and virtual outdoor hallways (Marusak et al., 2017, Marusak et al., 2018; Marusak et al., 2021).

3.1.6.1 Virtual City and Virtual Island

While VR technology is increasingly used in clinical settings, research on its impact on emotional, motivational, and cognitive functions in TD children remains limited (Cadet et al., 2022). Specifically, the relationship between immersion and memory in children is not well understood. Studies in adults suggest that different aspects of immersion, such as field of view, device type, and 3D asset quality, can influence memory retention of virtual experiences (Baños et al., 2008; Bowman & McMahan, 2007; Cadet & Chainay, 2020; Smith, 2019). However, further research is needed to explore whether these findings apply to children. This study explored how immersion, sense of presence, and emotions influence memory performance in VR among both adults and children.

Participants explored two virtual environments, the city and the island, each containing six locations to visit. The participants remained seated throughout the experiment, with head and hand tracking enabled via the HMD and controllers. The order of their visits could be chosen by the participants, but all six locations had to be completed in one environment before moving to the next. Stimuli were presented randomly within a 180-degree field of view, preceded by a spatialized sound cue to ensure attention. Participants had to point at each stimulus before it disappeared. After each visit, they answered self-reported questions on valence, arousal, and presence using a five-question scale. Upon completing all six locations in the first environment, participants removed the HMD, completed a French version of the ITC-SOPI questionnaire (Lessiter et al., 2001), and were asked to recall all stimuli they had encountered. Their responses were recorded, and ambiguous answers were clarified at the end. The same procedure was then repeated for the second environment. (Cadet et al. 2022).

3.1.6.2 Virtual City and Virtual Bus Ride

This study evaluates the effectiveness of a new VR-based learning environment designed to help children with autism safely practice and rehearse daily activities related to school life (Almazaydeh et al., 2022). The virtual city environment was created using the Unity platform, a widely used technology primarily designed for game and VR development. In this study, the virtual city was designed to include a diverse range of elements such as buildings, streets, transportation, people, trees, flowers, and rocks. These additions were intended to enhance the variety of scenes and contribute to a more authentic, natural atmosphere within the environment. The study was divided into three phases, which are going to be elaborated on.

The VR-Based Learning Practice consists of a conceptualization video, which was developed to present a virtual representation of a street beyond the school environment. The video serves primarily as an educational resource, designed to instruct viewers on specific guidelines for safely crossing the street with the aid of traffic lights. As a pre-activity demonstration, it provides foundational knowledge before engaging in interactive learning experience. To enhance realism, 3D modeling was used to simulate the flow of both traffic and pedestrians, incorporating traffic lights at road intersections and pedestrian crossings, The video showcases various street-crossing scenarios, distinguishing between safe and unsafe approaches. Additionally, distraction audio was layered over the video to regulate arousal levels and introduce an element of attentional challenge, reflecting real-world distractions pedestrians may encounter.

The VR-Based Test Practice phase is built on the learning and training outcomes established in the previous phase. This stage focuses on applying acquired skills in a realistic, immersive VR environment that extends beyond the school setting. In this phase, children actively engage in a fully immersive VR simulation designed to replicate real-world street-crossing scenarios. Each child is represented by an avatar, allowing for the systematic observation of their street-crossing behavior. The recorded behavioral measures include:

1. Standing position on the sidewalk and proximity to the crosswalk.
2. Head movements, specifically whether the participant looks left and right before crossing.
3. Response to traffic lights and adherence to pedestrian signals.
4. Reaction time, measuring how long it takes for the child to initiate and complete the crossing.

After the initial street-crossing attempt, incremental difficulty was introduced in a second round to assess behavioral adaptations under more challenging conditions. These modifications included auditory distractions (such as the sounds of pets, lawnmowers, and car horns), increased traffic density (with a higher number of pedestrians and vehicles), and higher vehicle speeds (requiring more precise timing and situational awareness). These two rounds provided valuable observational data, offering insight into participants' reactions and decision-making in varied and increasingly complex street-crossing scenarios.

In the VR-Based Tour Practice phase, participants experience a virtual school bus ride, during which key behaviors are recorded: 1. Social interactions, 2. Seat selection, and 3. Proper use of the bus stop signal. After completing the initial training, participants log into the system to track their performance in the remaining phases. User data is stored for post-analysis, enabling children, parents, and professionals to review behavioral patterns. This allows children to reflect on their decision-making and attention skills, while long-term tracking helps identify progress and areas needing improvement. Parents and therapists can provide targeted feedback, reinforcing a constructive and engaging learning process (Almazaydeh et al., 2022).

3.1.6.3 Virtual Apartment and Virtual Classroom

Many individuals who suffer from stuttering experience anxiety in various speaking situations, particularly public speaking (Vanryckeghem et al., 2017). Research indicates that stuttering in adulthood is often linked to a higher prevalence of social anxiety disorder (Iverach et al., 2011). While VR exposure therapy has proven effective for treating specific phobias in children and adolescents (Bouchard, 2011) and shows potential for treating social anxiety disorders (Parrish et al., 2016), its potential as an exposure-based intervention for individuals who stutter, remains underexplored. Moise Richard et al. (2021) aimed to evaluate the potential of a virtual classroom as a clinical tool for school-age children and adolescents who stutter.

During the VR session participants session, participants were immersed in four consecutive virtual environments: an empty apartment, a neutral classroom, a challenging classroom, and finally, a return to the initial empty apartment. The virtual apartment (VR training room; Loranger & Bouchard, 2017) served as a control condition, containing no anxiety-inducing stimuli. Its purpose was to familiarize participants with the VR experience and control for any anxiety triggered solely by immersion in VR. The virtual classroom featured seven students, and a teacher sat at desks, while the participant stood behind the teacher's desk. Originally designed for social anxiety treatment, the classroom had two variations. In the neutral condition, students simply looked at the participant without being explicitly distracting or supportive. In the challenging condition, the researcher introduced social stressors, such as students yawning, laughing, or appearing disinterested, as well as a teacher displaying dissatisfaction. These reactions were triggered at specific time intervals and tailored to each participant's perceived difficulty in speaking in front of a class. This personalized approach allowed for a gradual increase in anxiety exposure, aligning with each participant's individual hierarchy of feared situations.

3.1.6.4 Virtual Hallways

One key factor in the development of PTSD and other stress-related disorders are disruptions in fear and extinction learning (Lissek, et al., 2005). Since these conditions often originate in childhood or adolescence, investigating the underlying neural and behavioral mechanisms in young individuals is essential (Kessler, 2005). A deeper insight into fear and extinction learning and the underlying neurobehavioral mechanisms in children can help (1) identify developmental vulnerabilities, (2) determine mechanisms that contribute to pathology, and (3) guide the development of refinement of targeted treatment for youth. Furthermore, early alterations in extinction processes may explain (4) why adults with a history of childhood exposure to fear-inducing events, such as violence or abuse, are more susceptible to stress-related disorders. Research on fear conditioning and extinction have utilized Pavlovian fear conditioning paradigms, which typically involve associating a neutral stimulus with an aversive event, allowing researchers to study how fear responses are learned and later diminished (Marusak et al., 2017). Marusak et al. (2017) introduce and validate a novel virtual reality paradigm designed to study threat-related learning and extinction in children, replicating real-world fear-inducing events and environments linked to the development of fear- and stress-related disorders. The novel paradigm offers a more comprehensive assessment of fear responses beyond physiological markers like SCRs or subjective measures such as far and expectancy ratings and allows participants to move freely within the VR environment, providing valuable behavioral data, such as avoidant responses during fear extinction. Marusak et al. (2018)

extended their research by examining the brain and behavioral correlates of extinction recall in pre-adolescent children, using the novel VR paradigm with robust context manipulations that model interpersonal threat-related experiences and assess conditioned fear responses at multiple levels. Further research has been made by Marusak et al. (2021) by applying the novel paradigm to investigate the effects of childhood trauma exposure on fear conditioning, extinction learning, and extinction recall in pre-adolescent children.

Children participated in a two-day fear-extinction paradigm. On the first day, participants completed fear conditioning and extinction learning in separate VR environment contexts, with a ten-minute break between the two phases. The sessions of the first day were carried out in a controlled laboratory setting. Day 2 focused on testing extinction recall and fear renewal. As the primary interest was extinction recall, this session was conducted during fMRI scanning using a 2D adaptation of the task. After the recall session, participants underwent a fear renewal test conducted on a laptop outside the scanning environment. To evaluate whether the fMRI scanning environment influenced extinction recall, a subgroup of participants completed all four sessions exclusively in VR. This study used virtual reality to examine how children learn and extinguish fear in response to interpersonal threats. Fear conditioning was conducted in a "danger" environment (CXT+) followed by extinction training in a separate "safety" environment (CXT-), with a 10-minute break between the two phases. A burst of white noise was used as the unconditioned stimulus (US), and trial numbers were based on prior pilot testing with children. The conditioned stimuli (CS) were virtual adult male avatars chosen to reflect real-world interpersonal threats, such as those involving violence or abuse, which children may encounter. The avatars represented different racial and ethnic groups to match the diversity of the study's participants. Both the assignment of context (danger or safety) and the selection of CS types were randomized to ensure balance across participants. Children navigated the virtual environment and used a joystick to rate their fear and expectations of the US. Using a joystick, participants actively navigated through the environment, enhancing spatial contextual representation and engaging hippocampal-dependent learning (Burgess et al., 2002). Electrodes and headphones remained in place throughout all sessions, being removed only during scheduled breaks (Marusak et al., 2018).

Physiological responses were captured by measuring skin conductance from their non-dominant hand, providing real-time data on their reactions during the task. Skin conductance responses (SCRs) and ratings of fear and US expectancy were utilized as physiological and subjective measures of fear, consistent with standard practices in fear-extinction research. Uniquely, our paradigm allowed children to actively interact with their environment during the fear conditioning and extinction process, an element not explored in previous pediatric fear-extinction studies. This approach enabled the assessment of conditioned fear across multiple dimensions: self-reported fear, physiological responses, and behavioral tendencies such as approach or avoidance actions.

3.1.7 Other Virtual Environments

This group categorizes all VEs, which didn't fit in any other of the categories.

3.1.7.1 Virtual Schoolyard

Other than the virtual classroom other virtual school related environments have been used. A virtual schoolyard has been used to study interpersonal distance (Martarelli et al. 2015).

Childhood obesity remains a major global health concern (Lobstein, et al., 2015; Lobstein et al., 2004), with recent estimates indicating that over 40 million children under five and over 330 million aged 5-19 were overweight or obese in 2016 (Fanzo, et al., 2018). Various biological, psychological, and social factors contribute to its development and persistence (Lehrke & Laessle, 2003; Puder & Munsch, 2010). The role of peer influence in obesity has gained interest, as peer selection processes may reinforce weight-related behaviors, impacting further weight development (Christakis & Fowler, 2007; Hammond, 2010). However, research on peer selection in childhood obesity remains limited. This study explored peer selection dynamics using an immersive VR environment. VR offers a valuable tool for studying these dynamics, providing high experimental control, external validity, and replicability (Blascovich et al., 2002). Additionally, VR effectively measures interpersonal distance, offering insight into approach and avoidance behaviors in social interactions (Martarelli et al., 2015).

The experiment included four one-minute conditions, each featuring different avatar groups in a virtual schoolyard: (1) normal-weight eating avatars, (2) normal-weight playing avatars, (3) overweight eating avatars, and (4) overweight playing avatars. All participants experienced all four conditions. To begin with, participants familiarized themselves with the VR environment for about a minute. Each scenario started with a school bell sound, after which the avatars appeared, approached the participant. And stopped at a predefined position. Participants were free to move and explore once the avatars reached their positions. To create eating and drinking avatars, images of unhealthy, palatable foods (sandwich, toast, donut) and soft drinks (cola) were sourced from Google SketchUp. Each of the four avatars held one of these items and performed a repetitive arm movement, raising and lowering it toward their mouth to simulate eating or drinking. In the playing condition, avatars participated in a traditional Swiss schoolyard game ("sigg-sagg-sugg"), where the children placed one foot in the center and made a decision to withdraw or kept it in place. The trial was lost for the child whose foot was the only one either remaining or removed from the center. After the VR task, participants were shown screenshots of the four avatar groups and asked which group they would prefer to play with in real life.

3.1.7.2 FBI paradigm based on Lenggenhager et al. (2007)

This paradigm extensively studied and replicated (Aspell et al., 2009; Aspell et al., 2013; Ionta et al., 2011; Lenggenhager et al., 2009; Palluel et al., 2011; Salomon et al., 2013), involves participants observing their own body from behind while experiencing visuo-tactile stimulation on both their actual and virtual bodies. During this process, participants watch a stick stroking the back of the virtual body, while simultaneous (synchronous) or mismatched (asynchronous) strokes are applied to their real back. In adults, synchronous stroking prompts stronger effects than asynchronous stroking, leading to increased identification with the virtual body and a shift in the perceived location of touch, as reported through questionnaires. Moreover, synchronous stroking causes a change in self-location, where individuals feel drawn closer to the virtual body. This displacement is measured using a walking task: participants, after being moved backward with their eyes closed, are asked to return to their original position. They consistently stop slightly ahead of their prior spot, moving toward the virtual body. This forward drift in self-location demonstrates how synchronous stroking influences bodily self-awareness. The FBI paradigm is thus utilized to explore how multisensory processes shape bodily self-consciousness in children, focusing on tactile perception, self-identification, and self-location Each participant was placed on a 1 x 4-meter strip of carpet and followed a sequence consisting of a baseline trial, the recording of a video for the asynchronous trial, one synchronous and one asynchronous test trial in randomized order, and a second baseline trial. After completing each test trial, participants underwent a self-location drift measurement and responded to questionnaire items. White noise was delivered through headphones connected to an iPod throughout the experiment (Cowie et al., 2018).

3.2 Outcomes

Several outcomes have been measured with the help of VR. These outcomes are divided into behavioral outcomes, task specific outcomes, cognitive and psychological outcomes, physiological measures and observational measures, emotions and affects and lastly a category named others for all the outcomes that didn’t fit into the other categories.

3.2.1 Behavioral Outcomes

This category summarizes outcomes that relate to specific behaviors. These outcomes include crossing behavior, cycling behavior and bus-related behavior.

3.2.1.1 Crossing Behavior

Crossing behavior, which is also often referred to as pedestrian behavior (Wang et al., 2020) or pedestrian skill (Tabibi et al. 2023) summarizes a plethora of different factors and behaviors, which are being examined to determine pedestrian safety outcomes.

Start Delay

Start Delay or Start Gap (Wang et al., 2020) is defined as the „time (in seconds) between the rear bumper of the first car in the chosen gap crossing the participant's path and the time when they step off the curb." (Corbett & Morrongiello, 2017, p. 299). Further, Morrongiello et al. (2015, p. 700) refer to "Start Delay" as an "index of cognitive processing time when watching traffic before crossing". (This delay before initiating a crossing into the traffic gap is viewed as a measure of children's cognitive processing efficiency in pedestrian scenarios (Wang et al., 2020).27/28))

Gap size choice

Gap size choice, which is also referred to as temporal gap (Morrongiello et al., 2016), average inter vehicle gap size (Morrongiello et al., 2019) or inter-vehicle gap choice (Morrongiello et al., 2019b) refers to the "average size of the inter-vehicle gap for which the participant chose to cross into", measured in seconds instead of distance (Morrongiello et al., 2015). It is calculated as the time interval (in seconds) between the rear bumper of the first vehicle in the gap passing the participant`s position and the front bumper of the second vehicle reaching the same point. (Morrongiello et al., 2021)

Average inter-vehicle gap threshold

The study of Morrongiello & Corbett (2015) extended the measure of "Gap size choice" to "Average inter-vehicle gap threshold" by including the judgment of the participants' parents. Morrongiello & Corbett (2015, p.3) based the average inter-vehicle gap threshold in seconds on "parent judgement of what gap size their child would cross into and children's actual street crossing". ("Safe" crossing was defined as either the child choosing not to cross or successfully crossing with more than 1 second of time left to spare. A "risky" crossing was defined as a scenario where the child was hit, or the vehicle passed within a 1-second gap from the child.)

Time Left to spare

Time left to spare refers to the remaining time (in seconds) before the approaching car intersects the child`s path (Morrongiello et al., 2015). It is calculated as the time (in seconds) between the child and the approaching vehicle at the moment the child steps off the curb (Morrongiello et al., 2020).

Time left to spare on gap entry

TLS on gap entry refers to the time (in seconds) between the participant stepping into the vehicle's path and the front bumper of the approaching car reaching their position. TLS on gap entry is calculated as Gap Size – Start delay.

High-risk time left to spare

High-risk time left to spare concludes trials with time left so spare in a range from 0 (hit) to .75 (close call) seconds (Morrongiello et al., 2020) or 0 to 0.25 (Corbett & Morrongiello, 2017; Morrongiello et al., 2019).

Dangerous crossings

Wang et al. (2020) assessed dangerous crossings, which are very similar to the beforementioned HRTLS. Dangerous crossings are the "number of crossings when the child was either struck by a virtual car or had a close call, which was defined as being within 1 s of being struck by an approaching vehicle." (Wang, et al., 2020, p. 419)26/27.

Hits

Another measure which comes very close to the HRTLS is the assessment of "Hits". A hit is defined as an instance where the participant is struck by a car (Morrongiello et al., 2016). Hits can also be measured in proportions of trials in which children were hit (Morrongiello et al., 2019). This can be determined by checking if the coordinates of the vehicle's front bumper intersect with the participant's position coordinates.

Time car in view

This construct, which also was referred to as "Checking traffic "(Morrongiello et al., 2016) or "Visual attention to the traffic" (Morrongiello & Corbett, 2021) or just "visual attention" (Morrongiello et al., 2020) describes the duration (in seconds) that the child spent observing an approaching vehicle while it was in their path. This is determined based on whether the vehicle was visible within the VR headset (Morrongiello et al., 2020). It is calculated as the percentage of time the child spends walking within the car's path while observing traffic, which is indicated by a head turn of at least 65 degrees toward the approaching vehicle (Morrongiello et al., 2016). Tabibi et al.

(2023) assessed a similar measure, which is attention to traffic, measured as the proportion of looks to the left and right relative to the total time spent waiting to cross the street.

Walking Speed

Walking speed is determined by dividing distance by time (Morrongiello et al., 2016). It is calculated every 1/20th of a second, allowing for precise measurement of speed at different points during the crossing, as the distance covered in 0. 05 seconds in minimal (Corbett & Morrongiello, 2017). Morrongiello & Corbett (2021) further assessed walking speed when crossing separately without traffic.

Increased walking speed in Car path

This metric represents the relative change in velocity when crossing a gap. It is calculated by determining the difference between the maximum velocity reached while participants are on the car's path and the velocity at the moment of entering the gap, then dividing this difference by the initial velocity upon entering the gap. This provides a ratio indicating how much the pedestrian adjusted their speed while crossing. (Morrongiello & Corbett, 2021).

Evasive actions

Evasive actions represent the change in velocity (measured in m/s) between the moment the participant entered the gap and their maximum velocity while in the path of the vehicle (Morrongiello et al., 2020).

Running

Close to walking, running was another outcome measured by Wang et al. (2020). Wang et al. (2020) define running by the number of instances in which a child ran while crossing the road, characterized by both feet being off the ground at the same time. Such behavior is significant for safety for two main reasons. First, running may result in the child failing to notice oncoming traffic. Second, running increases the likelihood of tripping or falling during the crossing. Additionally, children who run can confuse drivers, potentially leading to pedestrian-vehicle collisions due to inappropriate driver responses.

Missed opportunities.

Different authors assessed missed opportunities, which are the number of times the participants refrain from crossing the road, despite having the opportunity to do so safely, based on their crossing speed during the trial and the size of the inter-vehicle gap (Morrongiello et al. (2015). Often, missed opportunities conclude traffic gaps, which were 1.5 times longer than the time needed for the participants to cross safely (Wang et al., 2020).

Estimated Crossing Behavior measures

Different to other studies, Corbett & Morrongiello (2017) assessed estimated outcomes for the crossing behavior. Missed opportunities (estimated) are the number of inter-vehicle gaps that were rejected, even though these gaps were at least 1.5 times longer than the child's estimated crossing time, calculated using their average walking speed. Time Left to Spare (Estimated) is determined based on the child's estimated walking speed. An estimated TLS value of 0 indicates that the child would have been struck by the vehicle. The average walking speed was calculated by having the participant walk a distance equivalent to the street's width ten times, without wearing any VR equipment.

Pedestrian Safety Rules

Different to other studies, Luo et al. (2020) examined five pedestrian safety rules, which they use to measure crossing behavior. These consists of "B1, do not attempt to dash into the street; B2, do not cross the street when the light is red; B3, do not cross the street when the light is blinking; B4, check for traffic immediately before crossing; and B5, do not use the zebra crossing when the school bus is backing up.

Timing of Entry

Parr et al. (2021) primary measure of interest was the precision of children's timing relative to the lead vehicle in the gap when they entered the car's path, measured in milliseconds. Smaller times reflect tighter and more precise timing of entry. Additionally, a variability score for timing of entry was calculated by determining the standard deviation of entry timing across all trials.

Stafford et al. (2022) used the following constructs to assess street crossing behavior; 1. Actions selection consists of gap acceptance, % of collisions and information for gap selection, 2. Action Timing consists of timing of entry and variability of timing. Timing of entry and variability of timing, 3. Action control consists of crossing time, duration on collision course, magnitude of safety and time to spare. Crossing duration/time.

Position and Velocity Profiles

Chung et al. (2020) evaluated crossing behavior through (a) each participant's position and velocity profile as they approach the interception point, (b) the timing of entry (c) the participant's location within the gap at the point of interception. Apart from the timing of entry the other outcomes are much broader than the beforementioned walking speed or gap choice which is why these outcomes are mentioned separately.

Action Selection

One of the big three outcomes that Stafford et al. (2022) uses is "Action Selection" which consists of gap acceptance, % of collisions and information for gap selection. Gap acceptance is defined as Gaps between cars, which were classified as accepted crossings when the participants walked more than one meter across the six-meter road, % of collisions were noted if the inter-vehicle gap closed before the participant arrived at the opposite sidewalk. Information for gap selection compromises the time-based variable tau. To be more precise, the tau value here is determined by dividing the current size of a gap by its rate of closure.

Action Timing

This outcome consists of timing of entry and variability of timing. For this outcome, the time (in seconds) between the tail of the lead car in the gap passing the participant and the moment the pedestrian entered the traffic path was recorded. Additionally, a variability of timing score was calculated for each participant by determining the standard deviation of their entry timing across the 54 trials.

Action Control

This outcome consists of crossing time, duration on collision course, magnitude of safety and time to spare. Crossing time describes the time (in seconds) it took for participants to travel from entering the inter-vehicle gap to reaching the opposite sidewalk. The duration on collision course aimed to evaluate whether participants were attuning to the optical variable tau to adapt their walking behavior while crossing the road. The experiment modeled the scenario as the simultaneous closure of two gaps: one between the approaching vehicle and the pedestrian (tauX) and another between the pedestrian and the opposite sidewalk (tauZ). To evaluate the duration a participant remained on a safe course of passage, the percentage of the crossing time during which the tau value of gap Z was closer to zero than the tau value of gap X was calculated. Magnitude of safety is a measure aimed at evaluating the temporal margin of safety after a participant was no longer on a collision course. Therefore, the normalized area under the curve of the positive portion of the tau difference was calculated. At last, time left so spare is the time (in seconds) remaining until the vehicle reached the participant`s line of crossing at the moment the participant arrived at the opposite sidewalk.

3.2.1.2 Cycling behavior

During the critical time frame of each hazard scenario, several key performance metrics were recorded (Zeuwts et al. 2021):

1. Speed (km/h): The participant's average cycling speed throughout the event.
2. Response Rate (%): A binary measure indicating whether the participant applied the brakes within the designated time frame (1 = braked, 0 = did not brake)
3. Reaction Time (s): The time elapsed between the onset of the hazard and the participant's initial braking response.
4. Hazard Perception Score: A rating system evaluating how well the participant navigated the hazard; 0 = Collision, 1 = Near-miss or failure to visually fixate on the hazard, 2 = Successfully identifying and safely avoiding the hazard.

3.2.1.3 Bus-Related Behavior

Bus-Related Behavior in the study of Cadet et al. (2022) refers to the way participants interacted socially before getting into the bus, their selection of an appropriate seat and the proper use of the bus stop signal.

3.2.2 Task Specific Outcomes

These outcomes refer to measurements which are specific to the VR task. Overall game performance measures are also included in this category.

3.2.2.1 EPELI

In EPELI framework, several indices and scores were developed to assess task performance and behavior. For event-based subtasks, only those completed within 10 seconds after the triggering event were considered correct. Similarly, for time-based subtasks, only those completed within a 10-second window before or after the target time were deemed accurate.

To streamline the analysis, highly correlated variables (correlation bigger/same .85) in the control group were reduced to a single representative variable. The key metrics retained included:

1. Total Score: The number of correctly performed subtasks.
2. Task Efficacy: The percentage of relevant actions (necessary to complete subtasks) out of all recorded actions, excluding navigation clicks.

3. Navigation Efficacy: The total score divided by the distance covered, by encompassing both movement and object interactions.
4. Controller Motion: The angular movement of the controller during task execution.
5. Total Actions: The combined number of clicks and physical interactions, including clicks during instructional phases and non-relevant actions (hitting drums in the children's room)

Secondary variables related to prospective memory included time-based subtask scores, event-based subtask scores, and the frequency of clock checks. These secondary variables were not part of the primary analysis, as there were no pre-registered hypotheses for time or event-based tasks in EPELI. Thus, the study focused on the five primary performance metrics (Seesjärvi et al. 2022).

3.2.2.2 Virtual Maze Task Performance

The performance metrics in the virtual maze task include path length, latency, Gallagher cumulative search error and Gallagher proximity measure. Path Length refers to the total distance traveled from the starting point to the endpoint, measured in meters. During learning trials, it represents the entire distance covered until reaching the platform. Floating (remaining motionless) does not contribute to path length. In probe trials, path length is measured as the total distance traveled within 90 seconds. Latency is defined as the time taken to locate the target area (40 cm circle around the 20cm platform) after entering the virtual water tank. This measure applies to both learning (with a hidden platform) and probe (without a platform) trials. If a participant never reaches the platform, the maximum time is recorded (60 seconds for learning trials and 90 seconds for probe trials). Gallagher Cumulative Search Error is assessed by the sum of distances between the participant and the goal, averaged every second and corrected for the starting position. Lower values indicate better learning performance. This metric is commonly used to assess performance during learning trials. Gallagher Proximity Measure refers to the participant's average distance from the platform, adjusted for the starting position. This measure is typically used in probe trials, with lower values indicating a better understanding of the platform's location.

3.2.2.3 Overall Task performance

The overall game performance in the study of Amat et al. (2021) referred to the time to complete the game and the response time to each prompt. In the case of Halabi et al. (2017) the task performance was assessed by recording the start time and each participant's response time, which was displayed upon successful task completion. Response time was logged when the system detected speech via voice input. Therapists could monitor task progress in real time and track nonverbal responses. Additionally, the system could record response times for nonverbal inputs, such as waving, using an external monitor.

3.2.3 Cognitive and Psychological Outcomes

This category includes outcomes linked closely to cognition. These outcomes range from attention-related outcomes to social information processes, memory, and language related outcomes.

3.2.3.1 Attention-Related Outcomes

Omission

Omissions or omission errors occur when a participant fails to respond to a target stimulus when required and is often associated with inattention (Díaz-Orueta, et al., 2014). These errors are linked to selective and focused attention. In the AULA system, omission scores are categorized in several ways: total omission, omission by sensory modality (visual versus auditory), omissions influenced by distractors (with and without distractors), and omissions based on task type (No-X versus X tasks) (Díaz-Orueta, et al., 2014). To measure the effect of distractions on task performance, Stokes et al. (2022) computed a performance distraction index. This was calculated by subtracting the mean response count during distractor-on periods from the mean response count during distractor-off periods and dividing the difference by the combined mean response rates for both conditions. This index quantifies the proportion of task completed in the presence of distractors compared to when no distractors were present. The performance distraction index is conceptually similar to measuring omission errors, which are often indicative of attention lapses, a common challenge in individuals with ADHD (Buzy et al., 2009). However, this measure specifically accounts for the presence of known distractors in the testing environment.

Commission

Commission or commission errors happen when a participant incorrectly responds without the presence of a target stimulus and is often times associated with impulsivity (Díaz-Orueta, et al., 2014). These errors are associated with difficulties in motor control or response inhibition. Like omissions, AULA tracks commissions across total scores, sensory modalities, the presence or absence of distractors and the task type. (Díaz-Orueta, et al., 2014).

Reaction time

The average response time or reaction time measures how quickly a participant reacts to stimuli, recorded in milliseconds. This metric serves as a reliable indicator of information processing speed. AULA collects these averages across various conditions, such as tasks with and without distractors, No-X and X-tasks, and responses to auditory versus visual stimuli. Reaction times are also analyzed for both correct responses and commission errors to provide deeper insight into performance nuances (Díaz-Orueta, et al., 2014).

Reaction time variability

The variability in reaction times, measured by standard deviation, reflects the consistency – or lack thereof- in a participant's responses. This metric can also highlight lapses in vigilance or attention over time. (Díaz-Orueta, et al., 2014). Sustained Attention measures the ability to stay focused over an extended period, assessed by the duration of continuous attention (Tan et al., 2019). Kakoulidou et al. (2021) state that high Omission errors and reaction time variability can indicate difficulties with sustained attention. Mühlberger et al. (2020) analyzed cognitive tempo by assessing both reaction time and reaction time variability.

Accuracy

Fang et al (2019) operationalized the outcome accuracy as the percentage of correct items to total items y 100. In the study of Tan et al (2019) accuracy was assessed differently. The audio test and CPT included 390 test points, including the numbers of letters requiring a response. Accuracy is calculated by dividing the correct items by 390 items. A lower accuracy reflects poorer performance. The WCST ends when the participant correctly classifies six consecutive cards. A higher number of trials indicates greater difficulty in identifying the correct rule, signifying weaker cognitive flexibility. Longer response times and delayed task completion suggest slower cognitive processing and poorer performance overall.

Quality of Attention Focus (QF)

QF assesses the participant's ability to maintain attention on relevant stimuli, specifically when viewing the classroom blackboard. This measure combines visual omission and commission errors occurring while the blackboard is within the participant's field of view, providing insights into how internal or external distractions affect attention quality. (Díaz-Orueta, et al., 2014).

Response Rate

Stokes et al. (2022) used response rate as a metric to measure attentiveness during tasks. The response rate reflects the number of responses made during specific periods of interest and was calculated separately for each task. This metric was prioritized over response accuracy due to individual variability in the difficulty of math problems. Stokes et al. (2022) categorized periods of interest into Distractor-off periods and Distractor-on periods. Distractor-off periods are defined as moments when no distractors were active, corresponding to the default classroom setting, while Distractor-on periods are referred to as moments when distractors were active in the environment during a task.

Other Attention-Related Outcomes

Despite the other standardized outcomes which relate to attention, there were few outcomes which single studies used; number of correct responses to CPT tasks (Fang et al., 2019), number of incorrect responses (Fang et al., 2019), total time taken for the attention tasks (Tan et al., 2019), percent of completion (Tan et al., 2019) and probability of hits in attention tasks (Hong et al., 2022)

3.2.3.2 SIP

Participants' aggressive behavior was assessed by observing their actions within the VR scenarios. Their social information processing (SIP) was evaluated immediately after each scenario through a series of questions. During these one-minute assessments, participants remained immersed in the VR environment with the glasses on. The Sip measures included assessments of their anger, intent attributions, interaction goals, aggressive response generation, and evaluation of their actual aggressive responses in VR.

Participants' behavioral responses in the VR scenarios were evaluated through direct observation. A trained graduate student documented participants' actions in real-time. Subsequently, the behaviors were coded following standard procedures (De Castro et al., 2005) into three categories:

1) 0: Non-aggressive behavior (prosocial actions or avoidance).
2) 1: Mildly aggressive behavior (coercion or verbal aggression).
3) 2: Severely aggressive behavior (physical or destructive aggression).

To ensure reliability, a second rater independently coded 50% of the behavioral descriptions.

3.2.3.3 Memory

Cadet et al. (2022) assessed memory for the presented stimuli in their study by using a free recall task.

3.2.3.4 Language related outcomes

Structural Language Profiles

Boo et al (2022) assessed structural language profiles by the following outcomes: number of utterances, mean length of utterances, noun types and tokens, verb types and tokens, pronoun types and tokens, and discourse marker types and tokens, 1st person pronoun tokens, 3rd person pronoun tokens and ratio of children who produced ambiguous pronouns.

Reading Fluency

To measure reading fluency in the VR lexical decision task, reading speed and reading of correct words can be observed. Further one of two classic criteria were used to assess internal validity, which are effects of lexicality and word length on reading speed and reading of correct words (Mirault et al. 2021).

Stuttering severity

Stuttering severity was measured by assessing the percentage of stuttering instances while doing the VR task (Moïse-Richard et al., 2021).

Volume and Duration of time spent talking

Boyd et al. (2018) assessed speaking volume and time spent talking by using logs, which were generated by the VR system.

3.2.4 Physiological Measures and Observations

This group of outcomes includes physiological measures, which range from head movement, EEG signals, skin conductance responses, diaphragm expansion, heart rate and interpersonal distance and eye movement Further this category includes observational measures.

3.2.4.1 Head Movement

Head movements have been widely studied as a measure of hyperactivity in individuals with ADHD (Hong et al., 2020). Further, head rotation data was collected and analyzed as a neuro-behavioral indicator that could reflect the users' intentions while performing cognitive tasks (Yeh, et al., 2020). Prior research emphasizes that head movements aimed at target stimuli serve as a crucial indicator of attentional orientation, underscoring the importance of further exploration in this domain (Teicher et al., 1996; Lavine et al., 2002; Seyama, 2006; Frischen et al., 2007). Head movement can be quantified by tracking degrees of motion across three axes, with higher values indicating greater movement (Coleman et al., 2019). The three dimensions are horizontal (yaw), vertical movement (pitch), and tilting (roll) (Mangalmurti et al. 2020). Head movement can also be used to determine the participant's field of view at any moment (Tan et al. 2019). Another term which is often used synonymously in AULA or deduced from head movements is "Motor

Activity”. Using motion sensors integrated into the VR headset, the system records head movements to evaluate their frequency and relevance. This includes identifying whether movements are necessary or unnecessary. The system can also track how far and when a participant’s head deviates from the focus of attention or target stimuli (Díaz-Orueta, et al., 2014). Cadet et al. (2022) also measured minimum and maximum head position and head rotation (pitch, yaw, and roll), which serve to provide objective support for data.

3.2.4.2 EEG

EEG signals can be evaluated in both the time and frequency domains to extract relevant features, such as band power and fractal dimensions, which enables differentiation between focused attention and distraction (Tan et al. 2019). Oh et al. (2022) used EEG measurements to evaluate the physiological effects of VR content. EEG devices recorded brain activity, including alpha, beta, delta, gamma, and theta waves. In the study of Oh et al. (2022) a two-channel EEG device was attached to the forehead to measure frontal lobe activity. EEG captures electrical signals generated by neural transmission, with brain wave frequencies classified into distinct bands.

3.2.4.3 SCR

Chua et al (2019) identified several physiological markers, that possibly can reflect engagement and increased mental effort in participants. Specifically, for electrodermal activity, the authors focused on skin conductance responses (SCRs), which are characterized by peaks in skin conductance and are associated with increased arousal or excitement. Further, in the case of Marusak et al. (2018) SCRs were assessed through the non-dominant hand to have a physiological measure of conditioned fear. The assumption is that a decrease in physiological responses over time reflects habituation.

3.2.4.4 Diaphragm expansion

Throughout gameplay of the underwater VR game “DEEP”, diaphragm expansion data was continuously recorded at a sampling rate of 60 Hz (van Rooij et al. 2016).

3.2.4.5 Heart Rate

For heart rate measurements, the authors performed both time-domain analyses, with a particular focus on the low-frequency (LF) and high frequency (HF) components of heart rate variability (HRV). An elevated LF/HV ratio in HRV indicates increased mental effort (Chua et al. 2019).

3.2.4.6 Interpersonal Distance and Proximity

A LED marker mounted on the HMD tracked the participant's position within the VE using a Precision Position Tracking system, which included four optical cameras placed in each corner to the lab. The captured orientation and position data were transmitted to the rendering computer to update the visuals shown in the HMD in real time. Avatars were positioned approximately seven meters in front of the participants within a 3D schoolyard scene. For later analysis, the minimum distance between the participant and the nearest avatar was calculated (Martarelli et al. 2015). In the case of Boyd et al. (2018) proximity to other VR avatars was measured by using logs generated by the VR system.

3.2.4.7 Eye Movement

Eye movements and fixations serve as indicators of overt visual attention. A fixation is defined as the period during which the fovea remains relatively stationary on a specific point for at least 90 milliseconds. Fixations shorter than 90 milliseconds often occur as minor corrections between saccades, while those lasting between 90 and 140 milliseconds contribute to pre-attentive processing. Fixations exceeding 140-200 milliseconds indicate more detailed cognitive processing.

To evaluate gaze behavior for each event, the following parameters were analyzed:

1. Fixation Rate (Binary: 1 or 0): Indicates whether the participant fixated on the area of interest (AOI) during the event.
2. Entry Time of First Fixation (s): The time elapsed between the onset of the event and the participant's first fixation on the AOI.
3. Number of Fixations: The total count of fixations directed at the AOI.
4. Total Dwell Time (s): The cumulative duration of all fixations within the AOI, excluding saccadic movements.

These measures provide insight into how participants allocate their visual attention while navigating potentially hazardous cycling environments (Zeuwts et al. 2021).

Tan et al. (2019) assess the eye movement, which records X and Y coordinates of the eyeball position to map visual attention. To evaluate how much attention participants directed toward the task, Stokes et al. (2022) calculated the percentage of time their eye gaze (fixations) remained focused on the whiteboard during specific task periods. An invisible boundary was drawn around the whiteboard, simulating a stand-sized classroom setup. Fixations within this designated area were classified as on-task, while all other fixations were considered off-task. To assess how distractions influenced eye gaze, the authors calculated a gaze distraction index. This index was determined by subtracting the whiteboard fixation percentage during the post-distractor from the pre-distractor fixation percentage, then dividing the result by the pre-distractor fixation percentage. Cadet et al. (2022) also assessed pupillary dilation from the first to the fifth second and fixation time. These physiological measures aim to provide objective support for their data. Amat et al. (2021) measured gaze fixation, which refers to fixation points, ratio of gaze fixation on eye to gaze fixation on other facial features.

3.2.4.8 Observational Data:

In the study of Muravevskaia & Gardner-McCune (2023) the task performance was assessed by the strategies, which were used to find the virtual brother in the game, the way children perceive the game and the VR identity while playing the game. All of the above were identified by themes based on observation notes. In other studies video and screen recordings were captured to support the validation of physiological data through triangulation (Chua et al. 2019).

3.2.5 Emotions and Affects

Outcomes which relate closely to emotions and affects are gathered in this group. These include measurements of emotions, apprehension, valence, and arousal.

3.2.5.1 Emotions

Marusak et al. (2017, 2018, 2021) used three different ways to measure conditioned fear. Subjective measures of conditioned fear were measured through fear and US expectancy ratings. Fear was assessed through the question “How scary is this?”, which could be answered from “1=not scary” to “5= very scary”. The US expectancy ratings “Do you think you will hear a loud sound with this?” also ranged on a scale from 1 to 5 with 1 being “definitely not” and 5 “definitely yes”. These ratings were submitted at the start, in the middle and the end of each session and were projected into the VE. Further SCRs were assessed through the non-dominant hand to have a physiological measure of conditioned fear. The assumption is that a decrease in physiological responses over time reflects habituation. As a behavioral measure of conditioned fear approach/avoidance behavioral action tendencies were measured. To operationalize these action tendencies a heatmap of movement in the VEs was constructed. The number of times participants visited a quadrant is implied by warm and cool colors with warmer colors indicating higher numbers of visiting.

In the case of Wechsler et al. (2021) emotional responses were measured using visual rating scales, assessing positive emotions (joy, interest, fun) and negative emotions (scare, fear, disgust), with each scale ranging from zero to ten. Sum scores for positive and negative affect ranged from 0 to 30.

3.2.5.2 Apprehension of virtual stimuli

To evaluate apprehension toward virtual stimuli, a custom measure was developed targeting five dimensions commonly linked to animal phobias. Participants rated each dimension on a 0-10 scale using a joystick-controlled slider, with anchors ranging from “not dangerous” to “extremely dangerous”. After confirming each response with a button press, participants assessed: (A) their fear of the animal (spider or rabbit), (b) its perceived size, (c) how scary it appeared, (d) how dangerous it seemed, and € how disgusting the participants the stimuli would be (Silva et al., 2022).

3.2.5.3 Valence

Valence was assessed with a 9-point illustrated scale from “very unpleasant” to “very pleasant”, which was measured with a virtual scale, which can be used by head motion (Cadet et al., 2022).

3.2.5.4 Arousal

Arousal was assessed from “weakly arousing” to “very arousing” by using a virtual analogue scale. (Cadet et al., 2022). Another way to assess arousal levels is by measuring heart rates and respiratory rates. Arousal levels were determined by comparing participants’ average heart and respiratory rates during the VR experience to their baseline measurements (Huang et al., 2020).

3.2.6 Other Outcomes

This category includes all outcome measures which didn’t fit into the other categories. These range from sense of presence, motor exploration, drift in self-location and angular error.

3.2.6.1 Sense of Presence

Cadet et al. (2022) used a 7-point scale (ranging from not at all to absolutely) to assess various dimensions of presence.

- General Presence: Perceived sense of being in the virtual environment.
- Spatial Presence: Impression of being able to explore the environment.
- Tactile Presence: Perception of being able to physically interact with objects.
- Attentional Focus: Level of concentration on the assigned task
- Motivation: Willingness to complete the given task

3.2.6.2 **Motor Exploration**

Motor Exploration was assessed through movement variability following failure in the VR-based 3D trajectory matching task. Using a handheld controller, the participants attempted to replicate an identical straight 3D line across all 60 trials. Crucially, the drawn paths were not visible to the participants, requiring them to rely on internal feedback and exploratory adjustments rather than visual cues to improve their performance. To ensure that all children encountered a consistent and sufficient number of failures for assessing post-failure movement variability, an adaptive success criterion was implemented. The criterion was based on each child's own prior performance, specifically the deviation in length and elevation between the target and the drawn trajectory (Brummelman et al., 2022).

3.2.6.3 Drift in self-location

Drift was assessed by measuring the distance, in centimeters, between the participant's starting position and their final position, specifically in the direction of the virtual body (Cowie et al., 2018).

3.2.6.4 Angular Error

The absolute angular error was defined as the smallest difference in degrees between the participant's final heading direction at the last pole and their heading direction immediately after initiating the response turn, but before stepping forward (Barhorst-Cates et al., 2021).

4 Discussion

4.1 Recapitulation and Interpretation

In the research field of VR, it remained unclear how VR has been used as an assessment tool in late childhood and which constructs have been assessed with VR throughout the last decade. Further it was unclear which advantages and limitations the use of VR had in its role as an assessment tool in late childhood. To fill this research gap, a systematic review has been conducted. Baumeister & Leary (1997) suggested different ways of synthesizing the results of systematic reviews by "zooming out". The zoom out contains the process of connecting grouping findings and concepts. Following the process of the systematic literature search 65 studies have been identified which fit the inclusion and exclusion criteria. These 65 studies could be divided into seven groups, which differ in the VEs the participants were put into.

Regarding the first research question "How VR has been used as an assessment tool in late childhood?" all studies could be divided into six categories, which can be differentiated by the environment the participants have been put into. Research on VR assessment in late childhood from 2013-2023 has a heavy emphasis on virtual street environments with 16 studies being centered around street crossing tasks and two studies using hazard anticipation tasks on bikes.

Another big emphasis in the research field on VR assessment in late childhood is in virtual classrooms with a total of 24 studies, especially with virtual classrooms using a form of VR-CPT to assess attention-related outcomes. Other than the 18 studies which focus on VR-CPT, three other studies are centered on aggression in virtual classrooms and the remaining three consist of reading and speaking exercises in the virtual classrooms.

Five studies used gamified environments, which used different aspects of fantasy-worlds to add to the virtual experience. In contrast to these gamified environments five studies used neutral environments with a stronger emphasis on the virtual task that was needed to be done. Throughout the systematic search five studies have been found that use virtual apartments virtual environment, with the different rooms and furniture playing a significant role in the VR tasks. The use of varying virtual environments has been applied in six studies. These studies benefit from effects that the change of environments comes with. Lastly two studies remain that didn't fit into the other categories. One study used a virtual schoolyard to measure interpersonal distance. The category virtual classroom could have been broader by expanding the space virtual school. As all the other studies play just in the virtual classroom this thought was abandoned to keep the categories as homogenous as possible. The other study, which examined the full body illusion, didn't give any information about the virtual environment, which led to the inclusion in this group as crucial information is missing.

Regarding the second research question "Which psychological constructs have been successfully assessed using VR?" this systematic review identified a plethora of different outcomes. These outcomes range from behavioral outcomes, task-specific outcomes, cognitive and psychological outcomes and physiological measures and observations to emotion and affects and a few other outcomes that didn't fit into the other categories.

The third research question "What are the advantages compared to traditional assessments?" can be answered in many ways. One big advantage comes in the benefit of having high ecological validity. Because of the realistic simulations VR environments can closely replicate real-world scenarios, allowing researchers to assess outcomes in environments that can mirror different daily life activities and situations more accurately than traditional laboratory settings. Also, the tasks and environments can be modified precisely and embedded in meaningful and dynamic contexts, leading to more natural responses from participants. Environmental variables like distractors or stimuli in VR-CPTS or speed of cars in street-crossing tasks can be manipulated consistently across participants. These identical conditions can be reproduced across sessions and studies, supporting reliability and validity. Further, risky behaviors can be studied ethically without real-world consequences. Exposure to fear inducing situations can be done in a safe, controlled manner.

Another big advantage lies in multimodal data collection. VR assessment can have real-time tracking of gaze, movement, reaction times and task performance. EEG, heart rate monitors, or eye-tracking devices can be integrated to gather objective physiological responses. Further, observational data can be collected with a lot of details through logs of in-VR interactions. The immersive and interactive nature of VR can enhance children's engagement and task adherence. This is especially relevant in pediatric populations, where attention and motivation can be common limitations in traditional assessments. Gamified elements can further support sustained involvement and reduce nervousness. Furthermore, VR assessments can be tailored in real time to match participants' abilities or diagnostic profile. Adaptive features such as incrementally increasing task difficulty or incorporating personalized stimuli make VR suitable for diverse developmental and cognitive profiles, including children with ADHD or ASD.

4.2 Limitations

At worst, a literature reviewer may adopt the role of an "intuitive lawyer" rather than that of an objective scientist, selectively presenting evidence to advocate for a predetermined position, rather than impartially evaluating the full scope of available research (Baumeister & Newman, 1994). While this systematic review provides a comprehensive overview of how VR has been applied as a psychological assessment tool in children, several limitations must be acknowledged. Out of the 16 street crossing tasks ten studies include the same author. Further, these studies are methodologically similar, which on one hand facilitates the comparability between the studies, but on the other hand imply a certain bias, as the outcome variables and parameters in the virtual environment are quite similar. As Baumeister and Leary (1997) state literature reviewers must consider not only the volume but also the methodological diversity of the evidence. A hypothesis backed by 50 studies using the same method may be less robust than one supported by five studies employing varied methodologies. Narrative literature reviews are well-suited to recognize and integrate such methodological diversity, making it a critical factor in drawing meaningful and well-rounded conclusions. Therefore, the results in these studies have to be interpreted carefully and with consideration to the bias that comes from the same authors. With the help of different authors using the same VE and street crossing task the validity of the beforementioned can be strengthened. The heterogeneity of studies could also limit the interpretation of the results. The included studies vary considerably in terms of their methodological designs, sample characteristics, VR hardware, and outcome measures. This heterogeneity limits the ability to synthesize findings quantitatively and may reduce comparability. Confusion bias due to the absence of control groups and selection bias due to the absence of randomization can occur because of this heterogeneity of studies. These biases are all lowering internal validity and limiting the possibility of drawing clear conclusions (Segawa et al., 2020).

One of the beforementioned advantages, which is the possibility of integrating physiological measurements, also has a downside. When physiological measures are applied in VR-based assessments, they are typically used to capture-real-time responses to specific stimuli or events within the virtual scenario. However, unlike traditional laboratory environments, VR settings involve dynamic movement and complex visual input, which may introduce noise or artifacts in physiological recordings. As a result, baseline measurements, which are rather than reactive responses to rich, multisensory VR environments, might offer more reliable data for certain psychiatric assessments. (Geraets et al. 2022).

The conceptualization of psychological constructs (attention, behaviors, emotions) differed notably between studies. This diversity complicated the categorization of outcomes and may have led to interpretive ambiguity. A possible limitation in the interpretation of the results can occur through the grouping of the various psychological constructs. The outcomes that have been grouped up inductively can also be categorized differently. Physiological outcomes like "eye gaze" or "head movement" could have been put into the "attention-related outcomes" group, because these are assessed most of the time to measure attention. The plethora of possible variations of the way the outcomes have been categorized highlighted the possible bias in interpretation that this systematic review has. This ambiguity also applies to virtual environments as some virtual environments could fit in multiple groups. This systematic review struggled with conceptualizing groups for the outcomes and the virtual environments. The help of different researchers could help, and inter-rater variability could help facilitate the decision on how to categorize the findings.

Many studies were exploratory, lacked long-term follow-ups or had small sample sizes. Especially in clinical populations. Consequently, the diagnostic or predictive validity of VR-based assessments remains underexplored, and generalizability is limited. Further only studies in English were included, which may have excluded relevant findings published in other languages or formats, potentially introducing a publication bias. Further, several studies lacked detailed reporting on hardware specifications, calibration procedures, or sensory parameters, which are crucial for replication and interpretation of VR-based results. Additionally, because technology is advancing so rapidly, hardware becomes outdated quickly and proprietary issues limit the availability of VR applications across newer platforms (Bell et al., 2020).

4.3 Conclusion

By following these research questions, this work wants to provide a precise synthesis of different scientific articles on VR as a tool for psychological assessment in late childhood and therefore build a steppingstone for further researchers, to have an overview of the different assessment tools, measured psychological constructs and limitations of VR assessment. This systematic review highlights the growing utility of immersive VR as a versatile and innovative tool for psychological assessment in children. Across 65 included studies, VR was applied to evaluate a diverse array of constructs including attention, behavior, motor exploration, executive function, memory, emotion, social information processes and phobic response. Its ability to simulate ecologically valid environments while maintaining experimental control offers significant advantages over traditional assessment methods.

VR-based assessments were found to be particularly effective in increasing engagement, allowing for multi-data collection, and enabling the safe reproduction of real-world challenges that would otherwise be difficult or unethical to observe directly. Importantly, the integration of behavioral, cognitive, and physiological measures within these controlled environments enables a richer, more dynamic understanding of child functioning.

However, further research is needed to standardize protocols, establish normative data, and validate VR-based tools against established clinical benchmarks. As technology becomes more accessible and affordable, future work should also focus on the development of adaptive, scalable platforms suitable for diverse developmental profiles. In conclusion, while challenges remain, VR offers a promising and evolving frontier for the assessment of psychological functioning in children, with the potential to transform both research and clinical practice.

5 Appendix

Table 3: Complete List of the included studies

VR Environment	Authors	Area of exploration	Sample	Outcome Measures in VR
Virtual Street	Chung et al. 2022	Differences in velocity regulation	TD, 16 children and 16 adults	Timing of Entry, Position and Velocity Profiles
Virtual Street	Corbett & Morrongiello, 2017	Extent of measurement bias	TD, 130	Start Delay, Gap Size Choice, Time Left to Spare, High-risk Time Left to Spare, Walking Speed, Evasive Actions, Missed Opportunities, Estimated Crossing Behaviors
Virtual Street	Luo et al. 2020	Effectiveness of VR as a tool for identifying and reducing risky pedestrian behaviors	TD, 79	Pedestrian Safety Rules: Dashing into the street, Checking for traffic, Crossing when it is green, Not using the Zebra crossing when the bus is backing up
Virtual Street	Morrongiello & Corbett 2015	Comparison of parents' expectations and actual crossing behavior of children	TD, 139	Average inter-vehicle gap
Virtual Street	Morrongiello et al. 2015	Effect of time pressure on crossing behavior	TD, 143 children and 28 adults	Start Delay, Gap Size Choice, Time Left to Spare, Hits, Missed Opportunities
Virtual Street	Morrongiello et al. 2015b	Evasive action skills	TD, 95	Walking Speed, Evasive Actions
Virtual Street	Morrongiello et al. 2016	Factors contributing to injury risk	TD, 95	Start Delay, Gap Size Choice, Time Left to Spare, Time Car in View, Hits, Walking Speed

Virtual Street	Morrongiello et al. 2019	Relationship between children's perception of peers' behavioral norms for crossing and actual crossing behavior	TD, 86	Start Delay, Gap Size Choice, High-Risk Time Left to Spare, Hits
Virtual Street	Morrongiello et al. 2019b	Impact of peers on crossing behavior	TD, 137	Start Delay, Gap Size Choice, High-Risk Time Left to Spare
Virtual Street	Morrongiello et al. 2020	Differences between boys and girls in injury risk	TD, 133	Start Delay, Gap Size Choice, Time Left to Spare, High-Risk Time Left to Spare, Time Car in View, Evasive Actions
Virtual Street	Morrongiello & Corbett, 2021	Influence of heavy backpacks on crossing behavior	TD, 75	Start Delay, Gap Size Choice, Time Left to Spare, Time Left to Spare on Gap Entry, Time Car in View, Walking Speed, Increased Walking Speed in Car Path
Virtual Street	Morrongiello et al., 2021	Influence of absence of auditory information on crossing behavior	TD, 75	Start Delay, Gap Size Choice, High-Risk Time Left to Spare, Time Car in View, Increased Walking Speed in Car Path
Virtual Street	Parr et al. 2021	Relationship of inattention/ hyperactivity and motor timing difficulties	92 TD, 10 ADHD	Timing of Entry
Virtual Street	Stafford et al. 2022	Relationship of a time-based variable and decision-making in street crossing	TD, 15 children and 30 adults	Crossing behavior; Action Selection, Action Timing, Action Control
Virtual Street	Tabibi et al. 2023	Comparison regarding attention and crossing behavior	22 TD, 28 ADHD	Start Delay, Dangerous Crossings, Time Car in View
Virtual Street	Wang et al. 2020	Influence of individual differences and traffic environment factors on crossing behavior	TD, 101	Dangerous Crossings, Running

Virtual Street	Zeuwts et al. 2021	Investigation of the effects of induced mental fatigue on hazard perception	TD, 49	Cycling Behavior; Speed, Response Rate (%), Reaction Time, Hazard perception Score. Gaze Behavior; Fixation Rate, Entry of First Fixation, Number of Fixations, Total Dwell Time
Virtual Street	Zeuwts et al. 2023	Development and evaluation of an interactive and immersive VR-based hazard anticipation test	TD, 130	Cycling Behavior; Speed, Response Rate (%), Reaction Time, Hazard perception Score. Gaze Behavior; Fixation Rate, Entry of First Fixation, Number of Fixations, Total Dwell Time
Virtual Classroom	Amat et al. 2021	Design, development, and initial evaluation of the (InViRS), aimed at improving gaze sharing and gaze following	9 ASD, 9 TD	Game performance, Gaze prompt speed and Gaze fixation
Virtual Classroom	Areces et al. 2018	Comparison of cognitive variables and profiles and attentional profiles and examining predictive power of these variables	50 ADHD, 38 TD	Omission errors, commission errors, Head Movement, reaction time
Virtual Classroom	Bioulac et al. 2020	Evaluation of a cognitive remediation program embedded within VR-CPT	ADHD, 51	Omission errors, commission errors
Virtual Classroom	Boo et al. 2022	Comparison of structural language profiles of children during a personal narrative task	21 ASD, 24 ADHD, 31 comorbid ASD/ADHD, 22 TD	Structural language profiles
Virtual Classroom	Cho et al. 2022	Relationship between head movements and ADHD symptoms	20 ADHD, 17 TD	Head movement
Virtual Classroom	Coleman et al. 2019	Introduction of an ecologically valid approach to assess the impact of working memory training on attention performance	13 TD	Omission errors, commission errors, reaction time, reaction time variability, Head movement

Virtual Classroom	Díaz-Orueta et al. 2014	Analysis of convergent validity between AULA and CPT of Conners	57 ADHD	Omission errors, commission errors, reaction time, reaction time variability, Head Movement, quality of attention focus
Virtual Classroom	Eom et al. 2019	Influence of social cues on attention performance	20 ADHD, 18 TD	Omission errors, commission errors, reaction time, reaction time variability and accuracy
Virtual Classroom	Fang et al. 2019	Feasibility of VR-CPT in Chinese population	77 ADHD, 63 TD	Correct items, incorrect items, accuracy, and total time
Virtual Classroom	Halabi et al. 2017	Feasibility of a newly designed virtual system across three different interactive display types	3 ASD, 7 TD	Task performance (Social communication)
Virtual Classroom	Hong et al. 2022	Impact of distractors on sustained attention and head movement	21 ADHD, 19 TD	Probability of hit, hit reaction time, omission errors, commission errors, accuracy
Virtual Classroom	Kakoulidou et al., 2021	Influence of text choice reading comprehension and enjoyment	92, ADHD and TD	Omission errors and reaction time variability
Virtual Classroom	Mangalmurti et al. 2020	Influence of field of view on attentional performance and ADHD symptoms	45 ADHD, 40 TD	Head Movements/ Field of view, omission errors, commission errors reaction time, reaction time variability, accuracy,
Virtual Classroom	Mühlberger et al. 2020	Performance differences between (un)medicated children	94 ADHD, 34 TD	Omission errors, commission errors, reaction time, reaction time variability, Head Movement
Virtual Classroom	Negut et al. 2017	Assessment of diagnostic validity of VR-CPT, evaluation of its task difficulty and examination of the influence of distractors on participants in terms of cognitive absorption.	33 ADHD, 42 TD	Omission errors, commission errors, reaction time, total correct responses

Virtual Classroom	Renati et al. 2022	Relationship between cognitive profile and attentional profile	15 intellectually gifted children	Omission errors, commission errors, reaction time and reaction time variability
Virtual Classroom	Rodríguez et al. 2018	Comparison of traditional CPT and VR-CPT	237 ADHD, 101 TD	Omission errors, commission errors, reaction time and reaction time variability
Virtual Classroom	Stokes et al. 2022	Influence of distractors on attentional performance and evaluation of eye-tracking as assessment	20 ADHD	Eye; Whiteboard fixation percentage, response rate, gaze distraction index, performance distraction index, Omission Errors
Virtual Classroom	Tan et al. 2019	Development of an ADHD diagnosis tool	50 ADHD, 50 TD	Task Performance Data: time, percent completion, accuracy, and reaction time. Physiological Data: eye movement, Head movement, brain electricity activity
Virtual Classroom	Verhoef et al. 2021	Introduction of a novel interactive VR environment developed to assess aggressive social information processing (SIP) in children	18 TD and 14 children with aggressive behavior	Aggressive responses; anger, intent attributions, interaction goals
Virtual Classroom	Verhoef et al. 2022	Introduction of a novel interactive VR environment developed to assess aggressive social information processing (SIP) in children	18 TD and 14 children with aggressive behavior	Anger, intent attributions, interaction goals, outcome expectancies and response evaluation
Virtual Classroom	Verhoef et al. 2022 (b)	Investigation whether distinct SIP profiles could be identified among boys exhibiting aggressive behavior problems	118 boys from special education for disruptive behavior problems, 66 boys from regular education	Anger, intent attributions, interaction goals, behavioral responses, outcome expectancies, response evaluations
Virtual Classroom	Yeh et al. 2020	Development an assessment tool for ADHD	37 ADHD, 31 TD	Omission errors, commission errors, reaction time, accuracy, Head Movement

Virtual Classroom	Zulueta et al. 2019	Testing the external validity and diagnostic accuracy of AULA	213 ADHD, 194 TD	Omission errors, commission errors, reaction time, reaction time variability, Head Movement
Gamified Virtual Environment	Chua et al. 2019	Role of interactivity and immersion in learning	36 TD	Skin conductance responses, heart rate, observational data
Gamified Virtual Environment	D'Archangel et al. 2022	Investigation of visuospatial processing skills using a virtual 3D Morris Water Maze task	17 NF1, 14 TD	Path length, latency, Gallagher cumulative search error, Gallagher proximity measure
Gamified Virtual Environment	Muravevskaia & Gardner McCune 2023	Evaluation of an empathy-based VR Game	14 TD	Empathy Scale, VR Identity while playing the game, Ways Children perceive the game, Strategies to find the VR brother
Gamified Virtual Environment	Oh et al. 2022	Evaluation of effectiveness of VR cognitive training for children with ADHD by measuring their sense of presence and EEG responses	8 ADHD, 8 TD	Sense of Presence and EEG
Gamified Virtual Environment	Van Rooij et al. 2016	Testing a biofeedback VR game, and examining if the game could reduce anxiety levels and negative affect and increase positive affect	86 TD	Diaphragm expansions
Virtual Apartment	Merzon et al. 2022	Enhancement ADHD detection and gain deeper insights into the mechanisms underlying attention deficits	37 ADHD, 36 TD	Total Score, Task Efficacy, Navigation Efficacy, Controller motion, Total Actions and Total Score, Head Rotation Angle, Head Rotation Speed
Virtual Apartment	Seesjärvi et al. 2022	Evaluation of the virtual reality task EPELI and its predictive, discriminant, and concurrent validity	38 ADHD, 38 TD	Total Score, Task Efficacy, Navigation efficacy, Controller Motion and Total Actions
Virtual Apartment	Seesjärvi et al. 2023	Assessment of key psychometric properties of EPELI, including its internal consistency, validity, and ability to accurately simulate real-	77 TD	Total Score, Task Efficacy, Navigation efficacy, Controller Motion, Total Actions, Time-based prospective memory score, Clocked

		life executive functioning challenges		checks, and Event-based prospective memory score
Virtual Apartment	Silva et al. 2022	Examination of differences in the perceptions and expectations of a VR phobic stimulus before immersion (Study 1), and the extent to which anxiety and phobic fear predict apprehension toward phobogenic VR stimuli (Study 2)	Study 1: 241 TD, Study 2: 37 TD	Apprehension of virtual stimuli
Virtual Apartment	Wechsler et al. 2021	Development of VR Game for the prevention and treatment of fear of spiders	14 TD	Emotions (joy, interest, fun) negative affect (scare, fear, disgust)
Neutral VE	Barhorst-Cates et al.	Investigation of spatial updating, aiming to determine (1) which self-motion cues are necessary, (2) whether children rely on different strategies than adults, and (3) if these effects generalize across different environments	Study 1: 30 TD, Study 2: 25 TD	Angular Error
Neutral VE	Boyd et al. 2018	Assessment of the feasibility of immersive VR	11 ASD	Proximity, volume, duration of time spent talking, combination of these
Neutral VE	Brummelman et al.	Impact of parental praise on children's motor exploration	202 TD	Motor Exploration
Neutral VE	Mirault et al. 2021	Evaluating proof-of-concept for using a VR setup to measure reading fluency and track eye movements during silent reading in children.	90 TD	Reading Fluency; Correct Words, Reading Speed Eye Movement
Neutral VE	Weijs et al. 2021	Investigation of the difference between children's and adults' subjective and objective responses to a	32 TD children, 33 Adults	Skin conductance response

		virtual-reality-based, visuomotor, first-person perspective full body illusion.		
Multiple Virtual Environments	Almazaydeh et al. 2022	Evaluating the effectiveness VR-based learning environment to help children with autism practice and rehearse daily activities	9 ASD	Daily activity skills; street crossing behavior (Head Movement) and bus-related behavior,
Multiple Virtual Environments	Cadet et al. 2022	Influence of immersion, sense of presence, and emotions on memory performance in VR	TD, 48 adults and 40 children	Valence, arousal, sense of presence and memory, head position, head rotation, pupillary dilation, fixation time
Multiple Virtual Environments	Marusak et al. 2017	Introduce and validate a novel VR paradigm designed to study threat-related learning and extinction in children	43 TD	Measures of conditioned fear; Fear and US expectancy ratings, SCR, Approach/Avoidance behavioral action tendencies
Multiple Virtual Environments	Marusak et al.	Examining the brain and behavioral correlates of extinction recall in pre-adolescent children, using the novel VR paradigm	43 TD	Measures of conditioned fear; Fear and US expectancy ratings, SCR, Approach/Avoidance behavioral action tendencies
Multiple Virtual Environments	Marusak et al.	Investigate the effects of childhood trauma exposure on fear conditioning, extinction learning, and extinction recall in pre-adolescent children	20 Trauma Exposed, 24 TD	Measures of conditioned fear; Fear and US expectancy ratings, SCR, Approach/Avoidance behavioral action tendencies
Multiple Virtual Environments	Moise Richard et al. 2021	Potential clinical utility of VC for assessing school-aged children and adolescents who stutter.	10 children with Stuttering Diagnosis	Stuttering severity; the %SS
Other Virtual Environments	Martarelli et al. 2015	Exploration of peer selection dynamics using an immersive VR environment	16 children with overweight, 41 children with typical weight	Interpersonal Distance
Others Virtual Environments	Cowie et al. 2018	Exploration of the development of bodily self-consciousness by examining how children respond to the full-body illusion and its	31 TD adults, 28 6-7 years old children, 21 8-9 years old, 19 10-11 years old	Drift in self-location

	relationship with multisensory bodily information		

6 List of Tables

7 List of Figures

8 Bibliography

Alghamdi, M., Regenbrecht, H., Hoermann, S., Langlotz, T., & Aldridge, C. (2016). Social presence and mode of video communication in a Collaborative Virtual Environment. *acific Asia Conference on Information Systems, PACIS 2016 - Proceedings Pacific Asia Conference on Information Systems*, (p. 126).

Almazaydeh, L., Al-Mohtadi, R., Abuhelaleh, M., & Al Tawil, A. (2022). Virtual reality technology to support the independent living of children with autism. *International Journal of Electrical & Computer Engineering, 12*(4), pp. 4111-4117.

Amat, A. Z., Zhao, H., Swanson, A., Weitlauf, A. S., Warren, Z., & Sarkar, N. (2021). Design of an interactive virtual reality system, InViRS, for joint attention practice in autistic children. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 29*, pp. 1866-1876.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5.* Arlington: American psychiatric association.

Anastasaki, C., Orozco, P., & Gutmann, D. H. (2022). RAS and beyond: the many faces of the neurofibromatosis type 1 protein. *Disease Models & Mechanism, 15*(2), p. dmm049362.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual review of psychology, 53*(1), pp. 27-51.

Antonovsky, A. (1979). *Health, Stress, and Coping.* San Fransisco: Jossey-Bass Inc.

Apostolellis, P., & Daradoumis, T. (2010). "Exploring the value of audience collaboration and game design in immersive virtual learning environments.". *Proceedings of the 9th international conference on interaction design and children.*, (pp. 326-330).

Areces, D., Dockrell, J., García, T., González-Castro, P., & Rodríguez, C. (2018). Analysis of cognitive and attentional profiles in children with and without ADHD using an innovative virtual reality tool. *PloS one, 13*(8).

Artigas-Pallarés, J. (2009). Modelos cognitivos en el trastorno por déficit de atención/hiperactividad. *Revista de neurología, 49*(11), pp. 587-593.

Aspell, J. E., Heydrich, L., Marillier, G., Lavanchy, T., Herbelin, B., & Blanke, O. (2013). Turning body and self inside out: visualized heartbeats alter bodily self-consciousness and tactile perception. *Psychological science, 24*(12), pp. 2445-2453.

Aspell, J. E., Lenggenhager, B., & Blanke, O. (2009). Keeping in touch with one's self: multisensory mechanisms of self-consciousness. *PloS one, 4*(8), p. e6488.

Aukstakalnis, S., & Blatner, D. (1992). *Silicon mirage; the art and science of virtual reality.* Peachpit Press.

Baer, L., & Blais, M. A. (2010). *Handbook of clinical rating scales and assessment in psychiatry and mental health.* New York: Humana Press.

Bailey, J. O., Bailenson, J. N., Obradovic, J., & Aguiar, N. R. (2019). Virtual reality's effect on children's inhibitory control, social compliance, and sharing. *Journal of Applied Developmental Psychology, 64*, p. 101052.

Baños, R. M., Botella, C., Rubió, I., Quero, S., García-Palacios, A., & Alcañiz, M. (2008). Presence and emotions in virtual environments: The influence of stereoscopy. *CyberPsychology & Behavior, 11*(1), pp. 1-8.

Barhorst-Cates, E., Stefanucci, J. K., & Creem-Regehr, S. H. (2020). A comparison of virtual locomotion methods in movement experts and non-experts: testing the contributions of body-based and visual translation for spatial updating. *Experimental Brain Research, 238*(9), pp. 1911-1923.

Barhorst-Cates, E., Stoker, J., Stefanucci, J., & Creem-Regehr, S. H. (2021). Using virtual reality to assess dynamic self-motion and landmark cues for spatial updating in children and adults. *Memory & Cognition, 49*(3), pp. 572-585.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of abnormal child psychology, 19*, pp. 149-178.

Barkley, R. A., & Murphy, K. R. (2011). The nature of executive function (EF) deficits in daily life activities in adults with ADHD and their relationship to performance on EF tests. *Journal of Psychopathology and Behavioral Assessment, 33*, pp. 137-158.

Baumeister, R. F. (2013). *Writing a literature review. The portable mentor: Expert guide to a successful career in psychology.* Springer.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of general psychology, 1*(3), pp. 311-320.

Baumeister, R. F., & Newman, L. S. (n.d.). How stories make sense of personal experiences: Motives that shape autobiographical narratives. *Personality and Social Psychology Bulletin, 20*(6), pp. 679-690.

Bell, I. H., Nicholas, J., Alvarez-Jimenez, M., Thompson, A., & Valmaggia, L. (2020). Virtual reality as a clinical tool in mental health research and practice. *Dialogues in clinical neuroscience, 22*(2), pp. 169-177.

Biedermann, J., Monuteaux, M. C., Doyle, A. E., Seidman, L. J., Wilens, T. E., Ferrero, F., . . . Faraone, S. V. (2004). Impact of executive function deficits and attention-deficit/hyperactivity disorder (ADHD) on academic outcomes in children. *Journal of consulting and clinical psychology, 72*(5), p. 757.

Bioulac, S., Lallemand, S., Rizzo, A. A., Philip, P., Fabrigoule, C., & Bouvard, M. P. (2012). Impact of time on task on ADHD patient's performances in a virtual classroom. *European Journal of Paediatric Neurology, 16*(5), pp. 514-521.

Bioulac, S., Micoulaud-Franchi, J.-A., Maire, J., Bouvard, M. P., Rizzo, A. A., Sagaspe, P., & Philip, P. (2020). Virtual remediation versus methylphenidate to improve distractibility in children with ADHD: A controlled randomized clinical trial study. *Journal of Attention Disorders, 24*(2), pp. 362-335.

Blanke, O. (2012). Multisensory brain mechanisms of bodily self-consciousness. *Nature Reviews Neuroscience, 13*(8), pp. 556-571.

Blanke, O., & Metzinger, T. (2009). Full-body illusions and minimal phenomenal selfhood. *Trends in cognitive sciences, 13*(1), pp. 7-13.

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K. R., Hoyt, C. L., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological inquiry, 13*(2), pp. 103-124.

Boo, C., Alpers-Leon, N., McIntyre, N., Mundy, P., & Naigles, L. (2022). Conversation during a virtual reality task reveals new structural language profiles of children with ASD, ADHD, and comorbid symptoms of both. *Journal of Autism and Developmental Disorders, 52*, pp. 2970–2983.

Borke, H. (1973). The development of empathy in Chinese and American children between three and six years of age: A cross-cultural study. *Developmental psychology, 9*(1), p. 102.

Bouchard, S. (2011). Could virtual reality be effective in treating children with phobias? *Expert review of neurotherapeutics, 11*(2), pp. 207-213.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer, 40*(7), pp. 36-43.

Boyd, L. E., Gupta, S., Vikmani, S. B., Gutierrez, C. M., Yang, J., Linstead, E., & Hayes, G. R. (2018). VrSocial: Toward immersive therapeutic VR systems for children with autism. *In Proceedings of the 2018 CHI conference on human factors in computing systems* , pp. 1-12.

Brummelman, E., Grapsas, S., & van der Kooij, K. (2022). Parental praise and children's exploration: a virtual reality experiment. *Scientific reports, 12*(1), p. 4967.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron, 35*(4), pp. 625-641.

Buzy, W. M., Medoff, D. R., & Schweitzer, J. B. (2009). Intra-individual variability among children with ADHD on a working memory task: an ex-Gaussian approach. *Child Neuropsychology, 15*(5), pp. 441-459.

Cadet, L. B., & Chainay, H. (2020). Memory of virtual experiences: Role of immersion, emotion and sense of presence. *International Journal of Human-Computer Studies, 144*, p. 102506.

Cadet, L., Reynaud, E., & Chainay, H. (2022). Memory for a virtual reality experience in children and adults according to image quality, emotion, and sense of presence. *Virtual Reality, 26*(1), pp. 55-75.

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., . . . Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of anxiety disorders, 61*, pp. 27-36.

Cavanagh, K., & Davey, G. (2003). Access to information about harm and safety in spider fearful and nonfearful individuals: when they were good they were very very good but when they were bad they were horrid. *Journal of Behavior Therapy and Experimental Psychiatry, 34*(3-4), pp. 269-281.

Checa, D., & Bustillo, A. (2020). A review of immersive virtual reality serious games to enhance learning and training. *Multimedia Tools and Applications, 79*(9), pp. 5501-5527.

Chihak, B. J., Plumert, J. M., Ziemer, C. J., Babu, S., Grechkin, T., Cremer, J. F., & Kearney, J. K. (2010). Synchronizing self and object movement: how child and adult cyclists intercept moving gaps in a virtual environment. *Journal of experimental psychology: human perception and performance, 36*(6), p. 1535.

Cho, Y. J., Yum, J. Y., Kim, K., Shin, B., Eom, H., Hong, Y.-j., . . . Kim, E. (2022). Evaluating attention deficit hyperactivity disorder symptoms in children and adolescents through tracked head movements in a virtual reality classroom: The effect of social cues with different sensory modalities. *Frontiers in Human Neuroscience, 16*(943478), pp. 1-12.

Choi, J., Yoon, I.-Y., Kim, H.-W., Chung, S., & Yoo, H. J. (2010). Differences between objective and subjective sleep measures in children with attention deficit hyperactivity disorder. *Journal of Clinical Sleep Medicine, 6*(6), pp. 589-595.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine, 357*(4), p. 370379.

Chung, H. C., Choi, G., & Azam, M. (108). Changes in children's perception-action tuning over short time scales: Bicycling across traffic-filled intersections in a virtual environment. *Journal of experimental child psychology, 108*(2), pp. 322-337.

Chung, H. C., Choi, G., & Azam, M. (2020). Effects of initial starting distance and gap characteristics on children's and young adults' velocity regulation when intercepting moving gaps. *Human factors, 62*(2), pp. 1002-1018.
Clark, L., Iversen, S. D., & Goodwin, G. M. (2002). Sustained attention deficit in bipolar disorder. *The British Journal of Psychiatry, 180*(4), pp. 313-319.
Climent, G., & Banterla, F. (2011). *AULA: evaluación ecológica de los procesos atencionales.* San Sebastian, Spain: Nesplora.
Cohen, J. (2001). Social and emotional education: Core concepts and practices. In *Caring classrooms/intelligent schools: The social emotional education of young children* (pp. 3-29).
Coleman, B., Marion, S., Rizzo, A., Turnbull, J., & Nolty, A. (2019). Virtual reality assessment of classroom–related attention: an ecologically relevant approach to evaluating the effectiveness of working memory training. *Frontiers in psychology, 10*, p. 1851.
Corbett, M. R., & Morrongiello, B. A. (2017). Examining how different measurement approaches impact safety outcomes in child pedestrian research: Implications for research and prevention. *Accident Analysis & Prevention, 106*, pp. 297-304.
Corkum, P., Tannock, R., Modolfsky, H., Hogg-Johnson, S., & Humphries, T. (2001). Actigraphy and parental ratings of sleep in children with attention-deficit/hyperactivity disorder (ADHD). *Sleep, 24*(3), pp. 303-312.
Costa, R. M., Federov, N. B., Kogan, J. H., Murphy, G. G., Stern, J., Ohno, M., . . . Silva, A. J. (2002). Mechanism for the learning deficits in a mouse model of neurofibromatosis type 1. *nature, 415*(6871), pp. 526-530.
Côté, S., & Bouchard, S. (2005). Documenting the efficacy of virtual realityexposure with psychophysiological andinformation processing measures. *Applied psychophysiology and biofeedback, 30*, pp. 217-232.
Cotton, K. (1992). *Developing empathy in children and youth.* Northwest Regional Educational Laboratory.
Cowie, D., McKenna, A., Bremner, A. J., & Aspell, J. E. (2018). The development of bodily self-consciousness: changing responses to the Full Body Illusion in childhood. *Developmental science, 21*(3).
Cowie, D., Sterling, S., & Bremner, A. J. (2016). The development of multisensory body representation and awareness continues to 10 years of age: Evidence from the rubber hand illusion. *Journal of experimental child psychology, 142*, pp. 230-238.
Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological bulletin, 115*(1), p. 74.
Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media psychology, 19*(2), pp. 272-309.
D'Archangel, M. A., Del Tufo, S. N., Cutting, L. E., Harrison, F. E., Ess, K. C., & Barquero, L. A. (2022). Virtual reality water maze navigation in children with neurofibromatosis type 1 and reading disability: An exploratory study. *Journal of Pediatric Neuropsychology, 8*(4), pp. 155-167.
De Castro, B. O., & van Dijk, A. (2017). "It's Gonna End Up with a Fight Anyway" Social Cognitive Processes in Children with Disruptive Behavior Disorders. In *The Wiley handbook of disruptive and impulse-control disorders* (pp. 237-253).
Díaz-Orueta, U., Garcia-López, C., Crespo-Eguílaz, N., Sánchez-Carpintero, R., Climent, G., & Narbona, J. (2014). AULA virtual reality test as an attention measure: Convergent validity with Conners' Continuous Performance Test. *Child Neuropsychology, 20*(3), pp. 328-342.
Dobricki, M., & De La Rosa, S. (2013). The structure of conscious bodily self-perception during full-body illusions. *PloS one, 8*(12). doi:e83840
Doser, K., Belmonte, F., Andersen, K. K., Østergaard, J. R., Hove, H., Handrup, M. M., . . . Kenborg, L. (2022). School performance of children with neurofibromatosis 1: a nationwide population-based study. *European Journal of Human Genetics, 30*(12), pp. 1405-1412.
Doyle, A. E., Willcut, E. G., Seidman, L. J., Biederman, J., Chouinard, V.-A., Silva, J., & Faraone, S. V. (2005). Attention-deficit/hyperactivity disorder endophenotypes. *Biological psychiatry, 57*(11), pp. 1324-1335.
Dworkin, B. R. (1993). *Learning and physiological regulation.* Chicago, London: University of Chicago Press.
Efron, D., Sciberras, E., Anderson, V., Hazell, P., Ukoumunne, O. C., Jongeling, B., . . . Nicholson, J. M. (2014). Functional status in children with ADHD at age 6–8: a controlled community study. *Pediatrics, 134*(4), pp. e992-e1000.
Ehrsson, H. H. (2007). The experimental induction of out-of-body experiences. *Science, 317*(5841), p. 1048.
Eleftheria, C. A., Charikleia, P., Iason, C. G., Athanasios, T., & Dimitrios, T. (2013). An innovative augmented reality educational platform using Gamification to enhance lifelong learning and cultural education. *IISA 2013* (pp. 1-5). IEEE.
Eom, H., Kim, K., Lee, S., Hong, Y.-J., Heo, J., Kim, J.-J., & Kim, E. (2019). Development of virtual reality continuous performance test utilizing social cues for children and adolescents with attention-deficit/hyperactivity disorder. *Cyberpsychology, Behavior, and Social Networking, 22*(3), pp. 198-204.

Facoetti, A., Corradi, N., Ruffino, M., Gori, S., & Zorzi, M. (2010). Visual spatial attention and speech segmentation are both impaired in preschoolers at familial risk for developmental dyslexia. *Dyslexia, 16*(3), pp. 226-239.

Fang, Y., Han, D., & Luo, H. (2019). A virtual reality application for assessment for attention deficit hyperactivity disorder in school-aged children. *Neuropsychiatric disease and treatment, 15*, pp. 1517-1523.

Fanzo, J., Hawkes, C., Udomkesmalee, E., Afshin, A., Allemandi, L., Assery, O., . . . Schofield, D. (2018). *2018 Global Nutrition Report: Shining a light to spur action on nutrition.* Bristol, UK: Development Initiatives. Retrieved March 22, 2025

Faraone, S. V., Wigal, S. B., & Hodgkins, P. (2007). Forecasting three-month outcomes in a laboratory school comparison of mixed amphetamine salts extended release (Adderall XR) and atomoxetine (Strattera) in school-aged children with ADHD. *Journal of attention disorders, 11*(1), pp. 74-82.

Fernández-Fernández, M., & Morillo-Rojas, M. (2012). Test—Retest validation of AULA Nesplora (virtual reality continuous performance test) for ADHD. *Poster presented at the 2nd International ADHD Conference*, (pp. 23-25). Barcekibam Spain.

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D. M., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological medicine, 47*(14), pp. 2393-2400.

Friedman, J. M. (1999). Epidemiology of neurofibromatosis type 1. *American journal of medical genetics, 89*(1), pp. 1-6.

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological bulletin, 133*(4), p. 694.

Gallagher, S. (2005). *How the body shapes the mind.* Oxford: Clarendon press.

Garbi, M. (2021). National Institute for Health and Care Excellence clinical guidelines development principles and processes. *Heart, 107*(12), pp. 949-953.

Gau, S. S.-F., & Huang, W.-L. (2014). Rapid visual information processing as a cognitive endophenotype of attention deficit hyperactivity disorder. *Psychological Medicine, 44*(2), pp. 435-446.

Geraets, C. N., Wallinius, M., & Sygel, K. (2022). Use of virtual reality in psychiatric diagnostic assessments: a systematic review. *Frontiers in Psychiatry, 13*, p. 828410.

Gorini, A., Pallavicini, F., Algeri, D., Repetto, C., Gaggioli, A., & Riva, G. (2010). Virtual reality in the treatment of generalized anxiety disorders. *Annual Review of Cybertherapy and Telemedicine 2010*, pp. 39-43.

Gualtieri, C. T., & Johnson, L. G. (2005). ADHD: Is objective diagnosis possible? *Psychiatry (Edgmont), 2*(11), p. 44.

Halabi, O., El-Seoud, S. A., Alja'am, J., Alpona, H., Al-Hemadi, M., & Al-Hassan, D. (2017). Design of Immersive Virtual Reality System to Improve Communication Skills in Individuals with Autism. *International Journal of Emerging Technologies in Learning, 12*(5), pp. 50-64.

Hammond, R. A. (2010). Social influence and obesity. *Current Opinion in Endocrinology, Diabetes and Obesity, 17*(5), pp. 467-471.

Hew, K. F., & Cheung, W. S. (2010). se of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British journal of educational technology, 41*(1), pp. 33-55.

Hilti, C. C., Hilti, L. M., Heinemann, D., Robbins, T., Seifritz, E., & Cattapan-Ludewig, K. (2010). Impaired performance on the Rapid Visual Information Processing task (RVIP) could be an endophenotype of schizophrenia. *Psychiatry research, 177*(1-2), pp. 60-64.

Hong, N., Kim, J.-j., Kwon, J.-H., Eom, H., & Kim, E. (2022). Effect of distractors on sustained attention and hyperactivity in youth with attention deficit hyperactivity disorder using a Mobile virtual reality school program. *Journal of attention disorders, 26*(3), pp. 358-369.

Hyman, S. L., Shores, A., & North, K. N. (2005). The nature and frequency of cognitive deficits in children with neurofibromatosis type 1. *Neurology, 65*(7), pp. 1037-1044.

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., . . . Blanke, O. (2011). Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. *Neuron, 70*(2), pp. 363-374.

Iriarte, Y., Diaz-Orueta, U., Cueto, E., Irazustabarrena, P., Banterla, F., & Climent, G. (2016). AULA—Advanced virtual reality tool for the assessment of attention: Normative study in Spain. *Journal of Attention Disorders, 20*(6), pp. 542-568.

Ito, M. (2012). *Engineering play: A cultural history of children's software.* MIT Press.

Iverach, L., Jones, M., Menzies, R. G., O'Brian, S., Packman, A., & Onslow, M. (2011). Response to walter manning and J. Gayle Beck: Comments concerning iverach, jones et al.(2009a). *Journal of Fluency Disorders, 36*(1), pp. 66-71.

Jarrold, W., Mundy, P., Gwaltney, M., Bailenson, J., Hatt, N., McIntyre, N., . . . Swain, L. (2013). Social attention in a virtual public speaking task in higher functioning children with autism. *Autism Research, 6*(5), pp. 393-410.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies, 66*(9), pp. 641-661.

Kakoulidou, M., Le Cornu Knight, F., Filippi, R., & Hurry, J. (2021). The effects of choice on the reading comprehension and enjoyment of children with severe inattention and no attentional difficulties. *Research on Child and Adolescent Psychopathology, 49*(11), pp. 1403-1417.

Kennedy, R. S., Stanney, K. M., & Dunlap, W. P. (2000). Duration and exposure to virtual environments: sickness curves during and across sessions. *Presence: Teleoperators & Virtual Environments, 9*(5), pp. 463-472.

Kessler, R. (2005). Lifetime prevalence and age-of-onset distributeions of DSM-IV disorders in the National Comorbidity. *Arch Gen Psychiatry, 62*, pp. 593-602.

Kim, J., Lee, Y., Han, D., Min, K., Kim, D., & Lee, C. (2015). The utility of quantitative electroencephalography and integrated visual and auditory continuous performance test as auxiliary tools for the attention deficit hyperactivity disorder diagnosis. *Clinical Neurophysiology, 126*(3), pp. 532-540.

Kim, O. H., & Kaiser, A. P. (2000). Language characteristics of children with ADHD. *Communication disorders quarterly, 21*(3), pp. 154-165.

Konrad, K., Günther, T., Heinzel-Gutenbrunner, M., & Herpertz-Dahlmann, B. (2005). Clinical evaluation of subjective and objective changes in motor activity and attention in children with attention-deficit/hyperactivity disorder in a double-blind methylphenidate trial. *Journal of Child & Adolescent Psychopharmacology, 15*(2), pp. 180-190.

Kuijper, S. J., Hartman, C. A., Bogaerds-Hazenberg, S. A., & Hendriks, P. (2017). Narrative production in children with autism spectrum disorder (ASD) and children with attention-deficit/hyperactivity disorder (ADHD): Similarities and differences. *Journal of abnormal psychology, 126*(1), p. 63.

Lachapelle, U., Noland, R. B., & Von Hagen, L. A. (2013). Teaching children about bicycle safety: an evaluation of the New Jersey Bike School program. *Accident Analysis & Prevention, 52*, pp. 237-249.

Lalonde, G., Henry, M., Drouin-Germain, A., Nolin, P., & Beauchamp, M. H. (2013). Assessment of executive function in adolescence: a comparison of traditional and virtual reality tools. *Journal of neuroscience methods, 219*(1), pp. 76-82.

Lavine, R. A., John, S. L., Gokturk, M., & Dickens, B. (2002). Eye-tracking measures and human performance in a vigilance task. *Aviat. Space Environ. Med., 73*, pp. 367-372.

Lee, M.-H., Patel, P., & Ranganathan, R. (2022). Children are suboptimal in adapting motor exploration to task dimensionality during motor learning. *Neuroscience Letters, 770*, p. 136355.

Legrand, D. (2006). The bodily self: The sensori-motor roots of pre-reflective self-consciousness. *Phenomenology and the cognitive sciences, 5*(1), pp. 89-118.

Lehrke, S., & Laessle, R. (2003). Adipositas. In U. Ehlert, *Verhaltensmedizin. Springer-Lehrbuch* (pp. 497-529). Berlin, Heidelberg: Springer.

Lehtonen, A., Howie, E., Trump, D., & Huson, S. M. (2013). Behaviour in children with neurofibromatosis type 1: cognition, executive function, attention, emotion, and social competence. *Developmental Medicine & Child Neurology, 55*(2), pp. 111-125.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child development, 71*(1), pp. 107-118.

Lenggenhager, B., Mouthon, M., & Blanke, O. (2009). Spatial aspects of bodily self-consciousness. *Consciousness and cognition, 18*(1), pp. 110-117.

Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. *Science, 317*(5841), pp. 1096-1099.

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments, 10*(3), pp. 282-297.

Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour research and therapy, 43*(11), pp. 1391-1424.

Lobstein, T., Baur, L., & Uauy, R. (2004). Obesity in children and young people: a crisis in public health. *Obesity reviews, 5*(1), pp. 4-85.

Lobstein, T., Jackson-Leach, R., Moodie, M. L., Hall, K. D., Gortmaker, S. L., Swinburn, B. A., . . . Wang, Y. (2015). Child and adolescent obesity: part of a bigger picture. *The Lancet, 385*(9986), pp. 2510-2520.

Loe, I. M., & Feldman, H. M. (2007). Academic and educational outcomes of children with ADHD. *Journal of pediatric psychology, 32*(6), pp. 643-654.

Loranger, C., & Bouchard, S. (2017). Validating a virtual environment for sexual assault victims. *Journal of traumatic stress, 30*(2), pp. 157-165.

Losh, M., & Capps, L. (2003). Narrative ability in high-functioning children with autism or Asperger's syndrome. *Journal of autism and developmental disorders, 33*, pp. 239-251.

Louveton, N., Montagne, G., Berthelon, C., & Bootsma, R. J. (2012). Intercepting a moving traffic gap while avoiding collision with lead and trail vehicles: Gap-related and boundary-related influences on drivers' speed regulations during approach to an intersection. *Human movement science, 31*(6), pp. 1500-1516.

Luo, H., Yang, T., Kwon, S., Zuo, M., Li, W., & Choi, I. (2020). Using virtual reality to identify and modify risky pedestrian behaviors amongst Chinese children. *Traffic injury prevention, 21*(1), pp. 108-113.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction, 60*, pp. 225-236.

Mangalmurti, A., Kistler, W. D., Quarrie, B., Sharp, W., Persky, S., & Shaw, P. (2020). Using virtual reality to define the mechanisms linking symptoms with cognitive deficits in attention deficit hyperactivity disorder. *Scientific Reports, 10*(1), p. 529.

Manor, I., Meidad, S., Zalsman, G., Zemishlany, Z., Tyano, S., & Weizman, A. (2008). Objective versus subjective assessment of methylphenidate response. *Child psychiatry and human development, 39*, pp. 273-282.

Maoz, H., Aviram, S., Nitzan, U., Segev, A., & Bloch, Y. (2018). Association between continuous performance and response inhibition tests in adults with ADHD. *Journal of attention disorders, 22*(3), pp. 293-299.

Maples-Keller, J. L., Bunnell, B. E., Kim, S.-J., & Rothbaum, B. O. (2017). The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard review of psychiatry, 25*(3), pp. 103-113.

Martarelli, C. S., Borter, N., Bryjova, J., Mast, F. W., & Munsch, S. (2015). The influence of parent's body mass index on peer selection: An experimental approach using virtual reality. *Psychiatry research, 230*(1), pp. 5-12.

Marusak, H. A., Hehr, A., Bhogal, A., Peters, C., Iadipaolo, A., & Rabinak, C. A. (2021). Alterations in fear extinction neural circuitry and fear-related behavior linked to trauma exposure in children. *Behavioural Brain Research, 112958*, p. 398.

Marusak, H. A., Peters, C. A., Hehr, A., Elrahal, F., & Rabinak, C. A. (2017). A novel paradigm to study interpersonal threat-related learning and extinction in children using virtual reality. *Scientific Reports, 7*(1), p. 16840.

Marusak, H. A., Peters, C., Hehr, A., Elrahal, F., & Rabinak, C. A. (2018). Poor between-session recall of extinction learning and hippocampal activation and connectivity in children. *Neurobiology of learning and memory, 156*, pp. 86-95.

McIntyre, N. S., Solari, E. J., Gonzales, J. E., Solomon, M., Lerro, L. E., Novotny, S., . . . Mundy, P. C. (2017). The scope and nature of reading comprehension impairments in school-aged children with higher-functioning autism spectrum disorder. *Journal of autism and developmental disorders, 47*, pp. 2838-2860.

Meehan, M., Insko, B., Whitton, M., & Brooks Jr, F. P. (2002). Physiological measures of presence in stressful virtual environments. *Acm transactions on graphics (tog), 21*(3), pp. 645-652.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & education, 70*, pp. 29-40.

Merzon, L., Pettersson, K., Aronen, E. T., Huhdanpää, H., Seesjärvi, E., Henriksson, L., . . . Salmi, J. (2022). Eye movement behavior in a real-world virtual reality task reveals ADHD in children. *Scientific reports, 12*(1), p. 20308.

Mesa-Gresa, P., Gil-Gomez, H., Lozano-Quilis, J. A., & Gil-Gomez, J. A. (2018). Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: an evidence-based systematic review. *Sensors, 18*(8), p. 2486.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & education, 56*(3), pp. 769-780.

Mirault, J., Albrand, J.-P., Lassault, J., Grainger, J., & Ziegler, J. C. (2021). Using virtual reality to assess reading fluency in children. *Frontiers in education, 6*, p. 693355.

Mirault, J., Guerre-Genton, A., Dufau, S., & Grainger, J. (2020). Using virtual reality to study reading: An eye-tracking investigation of transposed-word effects. *Methods in Psychology, 3*, p. 100029.

Moïse-Richard, A., Ménard, L., Bouchard, S., & Leclercq, A.-L. (2021). Real and virtual classrooms can trigger the same levels of stuttering severity ratings and anxiety in school-age children and adolescents who stutter. *Journal of Fluency Disorders, 68*, p. 105830.

Moreno, R., & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: Role of methods and media. *Journal of educational psychology, 94*(3), p. 598.

Morris, R. G. (1981). Spatial localization does not require the presence of local cues. *Learning and motivation, 12*(2), pp. 239-260.

Morrongiello, B. A., & Corbett, M. (2015). Using a virtual environment to study child pedestrian behaviours: a comparison of parents' expectations and children's street crossing behaviour. *Injury prevention, 21*(5), pp. 291-295.

Morrongiello, B. A., & Corbett, M. (2021). The effect of a heavy backpack on children's street-crossing performance. *Accident Analysis & Prevention, 150*(105843).

Morrongiello, B. A., Corbett, M. R., & Vander Hoeven, E. (2021). Children's street crossing performance when auditory information about traffic is lacking. *Transportation research part F: traffic psychology and behaviour, 77*, pp. 149-155.

Morrongiello, B. A., Corbett, M., & Stewart , J. (2020). Understanding sex differences in children's injury risk as pedestrians. *Journal of Pediatric Psychology, 45*(10), pp. 1144-1152.

Morrongiello, B. A., Corbett, M., Koutsoulianos, S., & Foster, A. (2019). Effects of peers on child pedestrian behaviors in a virtual traffic context. *Journal of pediatric psychology*, pp. 509-516.

Morrongiello, B. A., Corbett, M., Milanovic, M., & Beer, J. (2016). Using a virtual environment to examine how children cross streets: Advancing our understanding of how injury risk arises. *Journal of pediatric psychology, 41*(2), pp. 265-275.

Morrongiello, B. A., Corbett, M., Milanovic, M., Pyne, S., & Vierich, R. (2015). Innovations in using virtual reality to study how children cross streets in traffic: Evidence for evasive action skills. *Injury prevention, 21*(4), pp. 266-270.

Morrongiello, B. A., Corbett, M., Switzer, J., & Hall, T. (2015). Using a virtual environment to study pedestrian behaviors: how does time pressure affect children's and adults' street crossing behaviors? *Journal of pediatric psychology, 40*(7), pp. 697-703.

Morrongiello, B. A., Seasons, M., McAuley, K., & Koutsoulianos, S. (2019). Child pedestrian behaviors: Influence of peer social norms and correspondence between self-reports and crossing behaviors. *Journal of safety research, 68*, pp. 197-201.

Morrongiello, B. A., Seasons, M., McAuley, K., & Koutsoulianos, S. (2019). Child pedestrian behaviors: Influence of peer social norms and correspondence between self-reports and crossing behaviors. *Journal of safety research, 68*, pp. 197-201.

Mueller, B. A., Rivara, F. P., Lii, S.-M., & Weiss, N. S. (1990). Environmental factors and the risk for childhood pedestrian-motor vehicle collision occurrence. *American journal of epidemiology, 132*(3), pp. 550-560.

Mühlberger, A., Jekel, K., Probst, T., Schecklmann, M., Conzelmann, A., Andreatta, M., . . . Romanos, M. (2020). The influence of methylphenidate on hyperactivity and attention deficits in children with ADHD: a virtual classroom test. *Journal of attention disorders, 24*(2), pp. 277-289.

Muravevskaia, E., & Gardner-McCune, C. (2023). Designing a Virtual Reality Empathy Game framework to create empathic experiences for children. *International Journal of Child-Computer Interaction, 35*, p. 100561.

Myers, K., & Winters, N. C. (2002). Ten-year review of rating scales. I: overview of scale functioning, psychometric properties, and selection. *Journal of the American Academy of Child & Adolescent Psychiatry, 41*(2), pp. 114-122.

Negut, A., Jurma, A. M., & David, D. (2017). Virtual-reality-based attention assessment of ADHD: ClinicaVR: Classroom-CPT versus a traditional continuous performance test. *Child Neuropsychology, 23*(6), pp. 692-712.

Nieuwkamp, R., & Schoeters, A. (2018). *Themadossier Verkeersveiligheid nr. 2-Fietsers.* Brussel, België: Vias institute – Kenniscentrum Verkeersveiligheid.

Norbury, C. F., Gemmell, T., & Paul, R. (2014). Pragmatics abilities in narrative production: A cross-disorder comparison. *Journal of child language, 41*(3), pp. 485-510.

Oh, S. H., Park, J. W., & Cho, S.-J. (2022). Effectiveness of the VR cognitive training for symptom relief in patients with ADHD. *Journal of Web Engineering, 21*(3), pp. 767-788.

Olmos, E., Cavalcanti, J. F., Soler, J.-L., Contero, M., & Alcañiz, M. (2018). Mobile Virtual Reality: A Promising Technology to Change the Way We Learn and Teach. In *Mobile and ubiquitous learning: An international handbook* (pp. 95-106).

Palluel, E., Aspell, J. E., & Blanke, O. (2011). Leg muscle vibration modulates bodily self-consciousness: integration of proprioceptive, visual, and tactile signals. *Journal of neurophysiology, 105*(5), pp. 2239-2247.

Parr, M. N., Tang, H., Mallaro, S. R., Kearney, J. K., & Plumert, J. M. (2021). Do inattention/hyperactivity and motor timing predict children's virtual road-crossing performance? *Journal of pediatric psychology, 46*(9), pp. 1130-1139.

Parrish, D. E., Oxhandler, H. K., Duron, J. F., Swank, P., & Bordnick, P. (2016). Feasibility of virtual reality environments for adolescent social anxiety disorder. *Research on Social Work Practice, 26*(7), pp. 825-835.

Parsons, T. D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in human neuroscience, 9*, p. 660.

Parsons, T. D., & Phillips, A. S. (2016). Virtual reality for psychological assessment in clinical practice. *Practice Innovations, 1*(3), p. 197.

Parsons, T. D., Bowerly, T., Buckwalter, J. G., & Rizzo, A. A. (2007). A controlled clinical comparison of attention performance in children with ADHD in a virtual reality classroom compared to standard neuropsychological methods. *Child neuropsychology, 13*(4), pp. 363-381.

Parsons, T. D., Duffield, T., & Asbee, J. (2019). A comparison of virtual reality classroom continuous performance tests to traditional continuous performance tests in delineating ADHD: a meta-analysis. *Neuropsychology review, 29*(3), pp. 338-356.
Parsons, T. D., Gaggioli, A., & Riva, G. (2017). Virtual reality for research in social neuroscience. *Brain sciences, 7*(4), p. 42.
Paul, G. (1983). Respiration, stress, and cardiovascular function. *Psychophysiology, 20*(3), pp. 284-300.
Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive psychology, 61*(2), pp. 106-151.
Petkova, V., & Ehrsson, H. H. (n.d.). If I were you: perceptual illusion of body swapping. *PloS one, 3*(12). doi:e3832
Pollak, Y., Shomaly, H. B., Weiss, P. L., Rizzo, A. A., & Gross-Tsur, V. (2010). Methylphenidate effect in children with ADHD can be measured by an ecologically valid continuous performance test embedded in virtual reality. *CNS spectrums, 15*(2), pp. 125-130.
Powell, L., Parker, J., & Harpin, V. (2018). What is the level of evidence for the use of currently available technologies in facilitating the self-management of difficulties associated with ADHD in children and young people? A systematic review. *European child & adolescent psychiatry, 27*(11), pp. 1391-1412.
Puder, J. J., & Munsch, S. (2010). Psychological correlates of childhood obesity. *International journal of obesity, 34*(2), pp. 37-43.
Quinn, C. (2003). Detection of malingering in assessment of adult ADHD. *Archives of Clinical Neuropsychology, 18*(4), pp. 379-395.
Ray, A. B., & Deb, S. (2016). Smartphone based virtual reality systems in classroom teaching—a study on the effects of learning outcome. *2016 IEEE eighth international conference on technology for education (T4E)*, (pp. 68-71).
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin, 124*(3), p. 372.
Redmond, S. M. (2004). Conversational profiles of children with ADHD, SLI and typical development. *Clinical linguistics & phonetics, 18*(2), pp. 107-125.
Renati, R., Bonfiglio, N. S., Mascia, M. L., & Penna, M. P. (2022). Exploring the Attentional Profile of Gifted Children Using a VR-Based Assessment Tool. *International Association for Development of the Information Society*.
Risko, E. F., & Kingstone, A. (2015). Attention in the wild: Visual attention in complex, dynamic, and social environments. In R. R. Hoffman, P. A. Hancock, M. W. Scerbo, R. Parasuraman, & J. L. Szalma, *The Cambridge handbook of applied perception research* (pp. 466-487). Cambridge University Press.
Rizzo, A. A., Bowerly, T., Buckwalter, J. G., Klimchuk, D., Mitura, R., & Parsons, T. D. (2009). A virtual reality scenario for all seasons: the virtual classroom. *CNS spectrums, 11*(1), pp. 35-44.
Rizzo, A. A., Bowerly, T., Shahabi, C., Buckwalter, J. G., Klimchuk, D., & Mitura, R. (2004). Diagnosing attention disorders in a virtual classroom. *Computer, 37*(6), pp. 87-89.
Rizzo, A. A., Buckwalter, J. G., Bowerly, T., Van Der Zaag, C., Humphrey, L. A., Neumann, U., . . . Sisemore, D. (2000). The virtual classroom: a virtual reality environment for the assessment and rehabilitation of attention deficits. *CyberPsychology & Behavior, 3*(3), pp. 483-499.
Rizzo, A. S., & Shilling, R. (2017). Clinical virtual reality tools to advance the prevention, assessment, and treatment of PTSD. *European journal of psychotraumatology, 8*(5), p. 1414560.
Rizzo, A. S., Buckwalter, J. G., & Neumann, U. (1997). Virtual reality and cognitive rehabilitation: a brief review of the future. *The Journal of head trauma rehabilitation, 12*(6), pp. 1-15.
Rizzo, A. S., Goodwin, G. J., De Vito, A. N., & Bell, J. D. (2021). Recent advances in virtual reality and psychology: Introduction to the special issue. *Translational Issues in Psychological Science, 7*(3), p. 213.
Rizzo, A., & Koenig, S. T. (2017). Is clinical virtual reality ready for primetime? *Neuropsychology, 31*(8), p. 877.
Roberts, I., Norton, R., Jackson, R., Dunn, R., & Hassall, I. (1995). Effect of environmental factors on risk of injury of child pedestrians by motor vehicles: a case-control study. *Bmj, 310*(6972), pp. 91-94.
Rodriguez, C., Areces, D., García, T., Cueli, M. C., & Castro, G. (2018). Comparison between two continuous performance tests for identifying ADHD: Traditional vs. virtual reality. *International journal of clinical and health psychology, 18*(3), pp. 254-263.
Romero-Ayuso, D., Toledano-González, A., Rodríguez-Martínez, M. d., Arroyo-Castillo, P., Triviño-Juárez, J. M., González, P., . . . Segura-Fragoso, A. (2021). Effectiveness of virtual reality-based interventions for children and adolescents with ADHD: a systematic review and meta-analysis. *Children, 8*(2), p. 70.
Rosenberg-Lee, M., Barth, M., & Menon, V. (2011). What difference does a year of schooling make?: Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *Neuroimage, 57*(3), pp. 796-808.

Roussou, M., & Slater, M. (2017). Comparison of the effect of interactive versus passive virtual reality learning activities in evoking and sustaining conceptual change. *IEEE Transactions on Emerging Topics in Computing, 8*(1), pp. 233-244.

Rufo-Campos, M., Cueto, E., Iriarte, Y., & Rufo-Muñoz, M. (2012). Estudio de sensibilidad de un nuevo método diagnóstico para el TDAH: Aula Nesplora. *Neurol Magazine, 54*(Supl3), pp. 67-93.

Rus-Calafell, M., Garety, P., Sason, E., Craig, T. J., & Valmaggia, L. R. (2018). Virtual reality in the assessment and treatment of psychosis: a systematic review of its utility, acceptability and effectiveness. *Psychological medicine, 48*(3), pp. 362-391.

Russ, S. W., & Dillon, J. A. (2011). Changes in children's pretend play over two decades. *Creativity Research Journal, 23*(4), pp. 330-338.

Sahlberg, H., Lehtonen, E., Rovamo, E., & Summala, H. (2015). Learning game for training child bicyclists' situational awareness for better hazard perception. *Internation Cycling Safety Conference.* Hannover, Germany.

Salomon, R., Lim, M., Pfeiffer, C., Gassert, R., & Blanke, O. (2013). Full body illusion is associated with widespread skin temperature reduction. *Frontiers in behavioral neuroscience, 7*, p. 65.

Satu, P., Minna, L., & Satu, S. (2025). Immersive VR assessment and intervention research of individuals with neurodevelopmental disorders is dominated by ASD and ADHD: A scoping review. *Review Journal of Autism and Developmental Disorders, 12*(1), pp. 50-68.

Schulz, L. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in cognitive sciences, 16*(7), pp. 382-389.

Seesjärvi, E., Puhakka, J., Aronen, E. T., Hering, A., Zuber, S., Merzon, L., . . . Salmi, J. (2023). EPELI: A novel virtual reality task for the assessment of goal-directed behavior in real-life contexts. *Psychological research, 87*(6), pp. 1899-1916.

Seesjärvi, E., Puhakka, J., Aronen, E. T., Lipsanen, J., Mannerkoski, M., Herin, A., . . . Salma, J. (2022). Quantifying ADHD symptoms in open-ended everyday life contexts with a new virtual reality task. *Journal of attention disorders, 26*(11), pp. 1394-1411.

Segawa, T., Baudry, T., Bourla, A., Blanc, J. V., Peretti, C. S., Mouchabac, S., & Ferreri, F. (2020). Virtual reality (VR) in assessment and treatment of addictive disorders: a systematic review. *Frontiers in neuroscience, 13*, p. 1409.

Sergeant, J. A., Oosterlaan, J., & van der Meere, J. (1999). Information processing and energetic factors in attention-deficit/hyperactivity disorder. In *Handbook of disruptive behavior disorders* (pp. 75-104). Boston, MA: Sringer US.

Serino, A., Alsmith, A., Constantini, M., Mandrigin, A., Tajadura-Jimenez, A., & Lopez, C. (2013). Bodily ownership and self-location: components of bodily self-consciousness. *Consciousness and cognition, 22*(4), pp. 1239-1252.

Seyama, J. (2006). Effect of image orientation on the eye direction aftereffect. *Psychological Research, 70*, pp. 367-374.

Siddaway, A. P. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual review of psychology, 70*(10), pp. 747-770.

Silva, C., Bouchard, S., & Bélanger, C. (2022). Children's perception of phobogenic stimuli in virtual reality. *International Journal of Child-Computer Interaction, 32*, p. 100417.

Silverman, W. K., Pina, A. A., & Viswesvaran, C. (2008). Evidence-based psychosocial treatments for phobic and anxiety disorders in children and adolescents. *Journal of Clinical Child & Adolescent Psychology, 37*(1), pp. 105-130.

Slaney, K. (2017). *Validating psychological constructs: Historical, philosophical, and practical dimensions.* Burnaby: Springer.

Smith, A. S. (2019). Virtual reality in episodic memory research: A review. *Psychonomic bulletin & review, 26*, pp. 1213-1237.

Snelson, C., & Hsu, Y.-C. (2020). Educational 360-degree videos in virtual reality: A scoping review of the emerging research. *TechTrends, 64*(3), pp. 404-412.

Soo-Jin, K., Kim, B.-N., Cho, S.-C., Kim, J.-W., Shin, M.-S., Yoo, H.-J., & Wim, H. W. (2010). The prevalence of specific phobia and associated co-morbid features in children and adolescents. *Journal of anxiety disorders, 24*(6), pp. 629-634.

Sovik, R. (2000). The science of breathing—the yogic view. *Progress in brain research, 122*, pp. 491-505.

Stafford, J., Rodger, M., Gómez-Jordana, L. I., Whyatt, C., & Craig, C. M. (2022). Developmental differences across the lifespan in the use of perceptual information to guide action-based decisions. *Psychological research, 86*, pp. 268-283.

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication, 42*(4), pp. 73-93.

St-Jacques, J., Bouchard, S., & Bélanger, C. (2010). Is virtual reality effective to motivate and raise interest in phobic children toward therapy? A clinical trial study of in vivo with in virtuo versus in vivo only treatment exposure. *The Journal of clinical psychiatry, 71*(7), p. 10026.

Stokes, J. D., Rizzo, A., Geng, J. J., & Schweitzer, J. B. (2022). Measuring attentional distraction in children with ADHD using virtual reality technology with eye-tracking. *Frontiers in virtual reality, 3*, p. 855895.

Sumner, C. R., Haynes, V. S., Teicher, M. H., & Newcorn, J. H. (2010). Does placebo response differ between objective and subjective measures in children with attention-deficit/hyperactivity disorder? *Postgraduate Medicine, 122*(5), pp. 52-61.

Tabibi, Z., Schwebel, D. C., & Juzdani, M. H. (2023). How does attention deficit hyperactivity disorder affect children's road-crossing? A case-control study. *Traffic injury prevention, 24*(4), pp. 315-320.

Tan, Y., Zhu, D., Gao, H., Lin, T.-W., Wu, H.-K., Yeh, S.-C., & Hsu, T.-Y. (2019). Virtual classroom: An ADHD assessment and diagnosis system based on virtual reality. *IEEE international conference on industrial cyber physical systems*, pp. 203-209.

Tanaka, J. W., & Sung, A. (2016). The "eye avoidance" hypothesis of autism face processing. *Journal of autism and developmental disorders, 46*(5), pp. 1538-1552.

Teicher, M. H., Ito, Y., Glod, C. A., & Barber, N. I. (1996). Objective measurement of hyperactivity and attentional problems in ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(3), pp. 334-342.

Thelen, E. (1995). Motor development: A new synthesis. *American psychologist, 50*(2), p. 79.

Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Pediatrics, 135*(4), pp. e994-e1001.

Tuena, C., Pedroli, E., Trimarchi, P. D., Gallucci, A., Chiappini, M., Goulene, K., . . . Stramba-Badiale, M. (2020). Usability issues of clinical and research applications of virtual reality in older people: A systematic review. *Frontiers in Human Neuroscience, 14*, p. 93.

Van der Hoort, B., Guterstam, A., & Ehrsson, H. H. (2011). Being Barbie: the size of one's own body determines the perceived size of the world. *PloS one, 6*(5). doi:e20195

Vanryckeghem, M., Matthews, M., & Xu, P. (2017). Speech situation checklist–revised: Investigation with adults who do not stutter and treatment-seeking adults who stutter. *American Journal of Speech-Language Pathology, 26*(4), pp. 1129-1140.

Veiga, G., Neto, C., & Rieffe, C. (2016). Preschoolers' free play: Connections with emotional and social functioning. *The International Journal of Emotional Education, 8*(1), pp. 48-62.

Verhoef, R. E., Alsem, S. C., Verhulp, E., & De Castro, B. O. (2019). Hostile intent attribution and aggressive behavior in children revisited: A meta-analysis. *Child development, 90*(5), pp. e525-e547.

Verhoef, R. E., van Dijk, A., Thomaes, S., Verhulp, E. E., van Rest, M. M., & de Castro, B. O. (2022). Detecting social information processing profiles of boys with aggressive behavior problems: An interactive virtual reality approach. *Development and psychopathology, 35*(4), pp. 1843-1855.

Verhoef, R. E., van Dijk, A., Verhulp, E. E., & de Castro, B. O. (2021). Interactive virtual reality assessment of aggressive social information processing in boys with behaviour problems: A pilot study. *Clinical psychology & psychotherapy, 28*(3), pp. 489-499.

Verhoef, R. E., Verhulp, E. E., van Dijk, A., & de Castro, B. O. (2022b). Interactive virtual reality versus vignette-based assessment of children's aggressive social information processing. *Research on child and adolescent psychopathology, 50*(5), pp. 621-636.

Vlakveld, W. (2011). *Hazard anticipation of young novice drivers: Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations.* University of Groningen.

Wang, H., Gao, Z., Shen, T., Li, F., Xu, J., & Schwebel, D. C. (2020). Roles of individual differences and traffic environment factors on children's street-crossing behaviour in a VR environment. *Injury prevention, 26*(5), pp. 417-423.

Wang, T., Liu, K., Li, Z., Xu, Y., Liu, Y., Shi, W., & Chen, L. (2017). Prevalence of attention deficit/hyperactivity disorder among children and adolescents in China: a systematic review and meta-analysis. *BMC psychiatry, 17*, pp. 1-11.

Waters, R. L., & Mulroy, S. (1999). The energy expenditure of normal and pathologic gait. *Gait & posture, 9*(3), pp. 207-231.

Wechsler, T. F., Kümpers, F., & Mühlberger, A. (2019). Inferiority or even superiority of virtual reality exposure therapy in phobias?—A systematic review and quantitative meta-analysis on randomized controlled trials specifically comparing the efficacy of virtual reality exposure to gold standard in vivo exp. *Frontiers in psychology, 10*, p. 1758.

Weems, C. F., Zakem, A. H., Natalie, C. M., Cannon, M. F., & Watts, S. E. (2005). Physiological response and childhood anxiety: Association with symptoms of anxiety disorders and cognitive bias. *Journal of Clinical Child and Adolescent Psychology, 34*(4), pp. 712-723.

Weijs, M. L., Macartney, E., Daum, M. M., & Lenggenhager, B. (2021). Development of the bodily self: Effects of visuomotor synchrony and visual appearance on virtual embodiment in children and adults. *Journal of Experimental Child Psychology,, 210*, p. 105200.

Weiland-Fiedler, P., Erickson, K., Waldeck, T., Luckenbaugh, D. A., Pike, D., Bonne, O., . . . Neumeister, A. (2004). Evidence for continuing neuropsychological impairments in depression. *Journal of affective disorders, 82*(2), pp. 253-258.

Wesnes, K., & Warburton, D. M. (1984). Effects of scopolamine and nicotine on human rapid information processing performance. *Psychopharmacology, 3*, pp. 147-150.

Wetton, M. A., Hill, A., & Horswill, M. S. (2011). The development and validation of a hazard perception test for use in driver licensing. *Accident Analysis & Prevention, 43*(5), pp. 1759-1770.

Wigal, S. B., & Wigal, T. L. (2006). The laboratory school protocol: its origin, use, and new applications. *Journal of Attention Disorders, 10*(1), pp. 92-111.

Winn, W., Windschitl, M., Fruland, R., & Lee, Y. (2002). When does immersion in a virtual environment help students construct understanding. In *Proceedings of the International Conference of the Learning Sciences* (Vol. 206, pp. 497-503).

Wolitzky-Taylor, K. B., Horowitz, J., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical psychology review, 28*(6), pp. 1021-1037.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines.* Geneva: World Health Organization.

World Health Organization. (2017). *Save lives: a road safety technical package.*

World Health Organization. (2018). *Global status report on road safety 2018.* Geneva: WHO.

World Health Organization. (2018). *Road Traffic Injuries.*

Yaneva, V., Ha, L. A., Eraslan, S., Yesilada, Y., & Mitkov, R. (2020). Detecting high-functioning autism in adults using eye tracking and machine learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28*(6), pp. 1254-1261.

Yeh, S.-C., Lin, S.-Y., Wu, E. H.-K., Zhang , K.-F., Rizzo, A., & Chung, C.-R. (2020). A virtual-reality system integrated with neuro-behavior sensing for attention-deficit/hyperactivity disorder intelligent assessment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28*(9), pp. 1899-1907.

Zeuwts, L. H., Iliano, E., Smith, M., Deconinck, F., & Lenoir, M. (2021). Mental fatigue delays visual search behaviour in young cyclists when negotiating complex traffic situations: A study in virtual reality. *Accident Analysis & Prevention, 161*, p. 106387.

Zeuwts, L. H., Vanhuele, R., Vansteenkiste, P., Deconinck, F. J., & Lenoir, M. (2023). Using an immersive virtual reality bicycle simulator to evaluate hazard detection and anticipation of overt and covert traffic situations in young bicyclists. *Virtual Reality, 27*(2), pp. 1507-1527.

Zeuwts, L. H., Vansteenkiste, P., Deconinck, F. J., Cardon, G., & Lenoir, M. (2017). Hazard perception in young cyclists and adult cyclists. *Accident Analysis & Prevention, 105*, pp. 64-71.

Ziegler, J. C., Pech-Georgel, C., Dufau, S., & Grainger, J. (2010). Rapid processing of letters, digits and symbols: What purely visual-attentional deficit in developmental dyslexia? *Developmental science, 13*(4), pp. 8-14.

Zulueta, A., Díaz-Orueta, U., Crespo-Eguilaz, N., & Torrano, F. (2019). Virtual reality-based assessment and rating scales in ADHD diagnosis. *Psicología Educativa, 25*(1), pp. 13-22.

Zulueta, A., Iriarte, Y., Díaz-Orueta, U., & Climent, G. (2012). Aula Nesplora: avance en la evaluación de los procesos atencionales. Estudio de la validación convergente con el test de percepción de diferencias “caras”(versión ampliada). *ISEP Science, 4*, pp. 4-10.