Usability studies evaluating Virtual Reality treatment in Mental Health Care: a literature review

UNIVERSITY OF TWENTE.

Faculty Behavioral, Management and Social Sciences

Positive Psychology and Technology

Bachelor Thesis

N. Binnenmars

1931334

First supervisor: Dr. C. Bode

Second supervisor: Dr. M. Galetzka

Abstract

Continuous development of technology in mental health care leads to new treatment methods such as Virtual Reality (VR). This literature review aims to determine which usability dimensions are discussed and how these are measured in studies which evaluate the usability of VR interventions for the treatment of mental health symptoms. In three databases, 464 articles were gathered with the search term: Treatment AND "virtual reality" AND (usability OR utility). These articles were screened, and seven exclusion criteria were applied. Three inclusion criteria were exercised on the remaining articles and 18 articles were eventually included in this review. The five usability dimensions (satisfaction, learnability, efficiency, error and memorability) of Nielsen (1994) were checked in the articles. The results show that satisfaction is researched most frequently, and the dimension memorability was not researched in the articles at all. Moreover, completely different measurement instruments were used to measure usability. From this literature review can be concluded that more attention should be paid to the different dimensions of usability and focus on one dimension or the realism of the intervention should be avoided. Also, more attention could be devoted to the methods used to measure usability since most articles did not choose the most fitting instruments to measure usability. Besides, it should be considered whether the dimensions used fit with the intervention evaluated, which will result in an adequate evaluation of the usability rather than the user-experience. Nielsen's framework is not the optimal framework to use when evaluating VR in mental health care and more attention should be devoted to creating a usability framework that can be applied to multiple interventions and allows comparing the usability of these.

Keywords: Virtual Reality, usability, treatment, mental health.

Introduction

Mental health issues are frequently occurring problems with people all over the world. More than one in ten individuals experience a mental illness (Ritchie, 2018). Besides, many people live with mental health problems which cannot be named a disorder but do reduce the quality of life, for example, emotional problems. These problems are often treated by suiting therapies such as exposure therapy, cognitive behavioral therapy, and Eye Movement Desensitization and Reprocessing (EMDR). Technology is making its advance in this domain. For example, the use of apps in mental health is on the rise and the advancement in mobile devices opens up many new opportunities for mental health care such as: managing self-assessment, to determine and monitor symptoms, or virtual coaching, which could be executed via apps (Luxton et al., 2011). The use of technology in mental health care enables interventions to be delivered more flexible and tailored to the needs of the patient group. One of these technologies that could be used in the treatment of mental health problems is Virtual Reality.

Virtual reality (VR) can be defined as: "an artificial environment which is experienced through sensory stimuli (such as sights and sounds) provided by a computer and in which one's actions partially determine what happens in the environment." (Merriam-Webster, n.d.). The system of computers is generating an image which results in a display presenting sensory information and a tracker that is using the users' position and orientation to update the system. This system has already existed for 50 years. However, the equipment we use today can be compared to that used first in the 1980s, which was then only used by specialist laboratories (Slater & Sanchez-Vives, 2016). Currently, the device is moving out of laboratories and is still developing. The newest development is an immersive head-mounted display (HMD). But also non-immersive systems, without the feeling of being immersed in the virtual world, are still being used (Freeman et al., 2017). VR is currently used in multiple medical settings such as education, treatments, diagnosis and thus also in mental health care (Zajtchuk & Satava, 1997). Benefits of using VR in treatment in mental health care are new skills that can be learned in a safe environment; patients have the opportunity to gradually be exposed which cannot be done in the real world. In addition to this, VR has lower therapy administration costs and the intervention can even be used at home. This makes VR an affordable and complementary care device, which can be used in combination with existing methods for optimal results. However, to reach this, sufficient usability is needed to be able to motivate the patients to make use of the intervention.

Usability can be defined as the assessment of how easy the user system is to use and refers to the method of improving easy-to-use during the design process (Nielsen, 2012). Usability of technical products is increasingly more important nowadays since not only experts are using these devices, but broader groups are using them for a larger variety of tasks as well. Therefore, it is important to test the usability for the target group whom you develop the intervention for (Nielsen, 1994). Besides, it is crucial to test the usability, because when your product does not function properly the users cannot get any work done while using the system. Consequently, the system will be rejected, and it will become useless even when the system software can help people. When a sufficient usability level is ensured, less training is needed and the use of the system can be increased (Nielsen, 2012). When looking at the usability of VR, the key is to ensure naturalness of the interaction which also could ensure that the ease of use is in danger. This is the reason why all components of usability should be reviewed (Sutcliffe, & Kaur, 2000).

Usability has multiple dimensions, of which Nielsen (1994) has identified five. These five dimensions are satisfaction, learnability, efficiency, error and memorability. Another leading approach, the International Organization of Standardization (ISO), is only discussing satisfaction, effectiveness and efficiency. The approach of Nielsen will be applied because of its specific focus on the user interface which makes it applicable to our data. Besides this, Nielsen's has a leading approach in the field of usability. The first usability dimension is *satisfaction*, which refers to the pleasantness of the experience users have while using the system. Users should be subjectively satisfied with the system which can be expressed by enjoying interacting with the system and likelihood of reporting it positively to friends. The second attribute is *learnability*, this refers to the easiness of completing basic tasks in the system, the system should be easy to learn so you can start to work easily with the system. The third attribute is *efficiency*, which can be measured once the participants have learned the design, how quickly they can perform the tasks given, and how accurate and complete they are in fulfilling the tasks. The fourth dimension is an *error*, which looks into how many errors the participants make, how severe these mistakes are and whether the participants can recover from these mistakes without additional support. The last attribute is *memorability*, which is the extent to which users can easily reestablish proficiency when returning to the system after some time of not using it. When the system is difficult to remember users should learn the system all over again with every use. These five dimensions can be measured with self-report questionnaires, administrating the time needed, using eye-tracking, observations, think aloud

and the dimension satisfaction can be measured by consulting psychophysiological measures like heart rate, pupil dilation and skin conductivity.

Existing literature reviews on this topic discuss VR treatment for a specific disorder or symptom (Paul et al., 2012; Rus-Calafell et al., 2018), focus on one specific intervention (Glegg et al., 2014) or do not have a specific focus on usability (Rizzo & Koening, 2017; Burridge & Hughes, 2010). The present literature review addresses the question of which dimensions of usability are frequently used to assess the usability of VR application in mental health treatment. And consequently, look whether Nielsen's (1994) dimensions are appropriate for measuring usability or if this framework can be questioned. This was done by reviewing the existing literature and it will add knowledge to the existing literature about how usability is tested, which dimensions are often assessed and whether usability studies accessing VR for treatment in mental health care have done this properly or if there is need for improvement.

Methods

This paper intends to map which dimensions of the usability of virtual reality treatment are researched in mental health care in the last 10 years (2011 - 2020).

The literature needed for this was obtained through searching in three electronic databases, namely, Scopus, PsycINFO and PubMed, which were chosen to include general, healthcare-related and psychology-related articles. In these databases, the same search string was used based on the research question "Which dimensions of usability are frequently used to assess the usability of VR application in mental health treatment?". This resulted in the search string: Treatment AND "virtual reality" AND (usability OR utility). These search terms were selected to include studies that applied virtual reality in the treatment of mental health problems and assessed the usability of these interventions. The more specific term "mental health" was not used in the search string because this led to too much exclusion of useful results since this term was not used in relevant titles, abstracts or keywords. This search string produced 464 articles in the last ten years to start with. In Figure 1 (p.7) displays the search flow.

These articles were examined for duplicates, which existed because multiple databases were consulted. These duplicated were removed by manual comparison and on the remaining articles seven exclusion criteria were applied:

- Focus on physical health symptoms
- Interventions to improve participation in daily activity (except for autism)
- Caregiver or health professional focused, such as training of surgeons
- Non-experimental design
- Conference lectures notes, proceedings.
- Literature reviews
- Abstracts without full locatable texts

The first exclusion criterium is related to all physical diseases, surgery and rehabilitation of them. The second exclusion criterium involves all studies focused on the improvement of daily activities, such as being able to go to the supermarket alone. However, in this, articles about Autism Spectrum disorder (ASD) are an exception since individuals with ASD experience difficulties in daily communication which can only be practised in daily activities like grocery shopping. This exception was made because interventions using daily activities in the treatment of mental a disorder was only experienced with articles about ASD, still, this *mental* disorder should be included. Thirdly, caregiver-focused interventions such as surgeon training devices are excluded, because this study is focused on the treatment of individuals with mental disorders. Fourthly, this study excluded al non-experimental designs, to create homogeneity in the reviewed articles, which gives the direct opportunity to compare the usability. Also, at this point conference lecture proceedings and notes were excluded because these did include summaries from the lectures given and did not involve experimental designs. Fifth, this study excludes literature reviews, to prevent double inclusion. Nonetheless, important articles discussed in these reviews were also included in this research. Sixthly, all abstracts without locatable full texts were excluded.

After the duplicates were removed, the articles were screened based on the title and abstract and the left 55 articles were analyzed based on full text. Abstracts without locatable full English text were excluded at this point. The following inclusion criteria were applied to those articles:

- Treatment focused
- Virtual reality used during treatment
- Usability assessed

The first inclusion criterium was adopted because this study focusses on the treatment of the mental disorders and not on, for instance, the training of the caregiver or the diagnosis of the disorder. Secondly, Virtual reality should be used in the treatment to be able to compare

the different effects of this technology for different mental disorders. Finally, to be able to compare the researched usability dimensions, the usability should be assessed in the included articles. This resulted in 18 articles that met the criteria and, thus, were included in the systematic review, these were examined and displayed in Table 1.

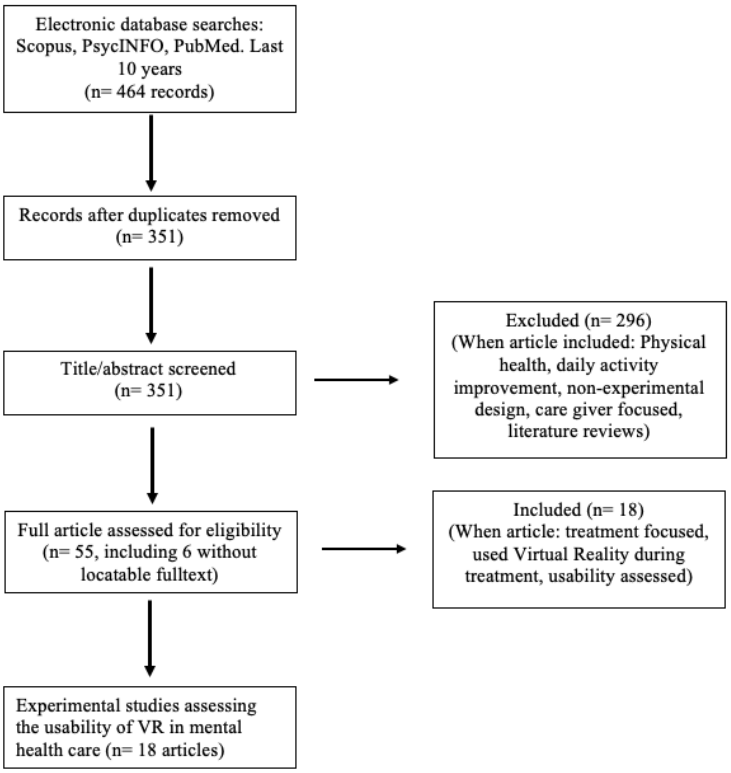

Figure 1. Search Flow

Results

This literature review aims to discover which dimensions of usability are frequently used to assess the usability of VR application in mental health treatment. Eighteen articles are included in this review, which all discussed the usability of VR interventions for several mental health disorders. Of the articles discussed, represented in Table 1, six analyze some sort of anxiety, five discussed Autism Spectrum Disorder (ASD), two discussed stress problems. The other articles all included different mental disorders, namely, obsessive-compulsive disorder (OCD), emotional wellbeing, bulimia nervosa and binge eating disorder, posttraumatic stress disorder, and alcohol abuse disorder.

Table 1

Study characteristics and disorders included

Study	Targeted mental symptoms	Participants	Technology used	IM/NIM	Focus on usability study	Usability framework used
Donker *et al.* (2019)	Fear	193 Dutch participants aged 18-65 with acrophobia symptoms.	Android smartphone	IM	Secondary, the main focus on effectiveness of treatment.	No named
Fowler *et al.* (2019)	Fear	16 Veterans with chronic pain with a mean age of 49.	Oculus rift and Samsung Oculus Gear.	IM	Secondary, the main focus on feasibility.	No named
Paulus, Suryani *et al.* (2019)	Fear	30 undergraduate psychology students with fear level >5	Mobile-Assisted Virtual Reality (MAVR)	IM	Secondary, the main focus on evaluation treatment.	No named
Paulus, Yusuf, *et al.* (2019)	Fear	12 men and 18 women with aged 18-25	Gaming computer and an android smartphone	NI	Primary	Heuristic evaluation Nielsen
Wong Sarver *et al.* (2013)	Social anxiety	11 children with a primary diagnosed with social anxiety disorder, aged 8 – 12	Laptop	NI	Primary	No named
Yuen *et al.* (2019)13	Anxiety	15 adults with PSA	Computer with webcam provided by the research facility.	NI	Secondary, the main focus on feasibility.	No named
Bekele *et al.* (2016)	Difficulties in social communication and interaction	6 children diagnosed with ASD, aged 13-17 and 6 healthy children aged-matched.	24-inch flat panel-monitor.	NI	Primary	No named
Di Mascio *et al.* (2019)	Difficulties in social communication and interaction	6 males with ASD age 21-23	Oculus Rift and HoloLens	IM	Primary	Based dimensions on other studies done on the same subject.
Kuriakose *et al.* (2016)	Difficulties in social communication and interaction	9 adolescents with ASD	Computer	NI	Primary	No named
Lahiri *et al.* (2013)	Difficulties in social communication and interaction	8 adolescents with high-functioning ASD.	Desktop Virtual Reality application	NI	Primary	No named
Tarantino *et al.* (2019)	Difficulties in social communication and interaction	6 males with a diagnosis ASD aged 21-23 year.	Oculus Rift and HoloLens	IM	Primary	Several dimensions based on all different articles.

Table 1

Continued

Gao *et al.* (2014)	Stress	11 participants aged 18-24	Mobile application on IPad. heart rate watches	NI	Primary	No named
Maarsingh *et al.* (2019)	Stress	111 healthy participants and 64 patients from a mental health facility	HTC Vive	IM	Primary	No named
Belloch *et al.* (2014)	Obsessive-Compulsive Disorder symptoms	4 women diagnosed with OCD, aged 22-42.	Television, laptop, Kinect for windows to connect both.	NI	Primary	No named
Flujas-Contreras *et al.* (2020)	Decreased emotional well-being	19 hospitalized participants, aged 8-16 year.	15-inch laptop, headphones and a joystick.	NI	Primary	No named, aimed for adherence, feasibility, us-ability, and acceptability.
Gutiérrez-Maldonado *et al.* (2016)	Food craving, anxiety, change in body image	113 undergraduate students aged 20-36. From which are 90 females, 23 males.	Oculus Rift	IM/NI	Secondary, the main focus on measuring a change in a food craving.	No named
Vollstädt-Klein *et al.* (2019)	Alcohol craving	21 patients, 18 suffering from alcohol abuse disorder and 3 from cannabis, with a mean age of 39.	A lightweight web application which can be used on tablets, laptops and desktop computers.	NI	Primary	Orsmond and Cohn (2015), five objectives.
Tielman *et al.* (2017)	Trauma	4 participants who all followed therapy for PTSD in the past.	Multi-model memory restructuring (3MR) on a computer screen.	NI	Primary	No named

Note. The abbreviations IM/NI means Immersive and Non-Immersive virtual reality used in the study.

Usability dimensions

To be able to answer the research question: "Which dimensions of usability are frequently used to assess the usability of VR application in mental health treatment?", Table 2 was created. This Table show that all, except one of the five dimensions of usability, as stated by Nielsen (2012), were named in the 18 articles included in this literature review. All articles use different methodology and measurement instruments to study these dimensions. Besides,

44.4% of the articles discuss two out of five usability dimensions, 50% discusses three out of five of the dimensions and, lastly, only one article (Yuen, *et al.* 2019) discusses four of the five dimensions. Most of the articles use different measurement methods and do not explicitly state the usability dimensions in their findings but do combine them or use different terms, these methods and terms are discussed in the following section.

The first dimension, satisfaction, refers to the pleasantness of the experience that users have with the system. This was discussed in each of the eighteen papers with similar, but also some different facets. The first facet is immersion, which was named in 13 articles and was expressed in a few ways: engagement in the intervention, the extent to which the virtual environment was realistic, the sense of presence or on the other side awareness of the real and virtual world and the compatibility of the user's task and domain. Secondly, telepresence in the virtual world was discussed in six of the articles, which was expressed by using the terms interaction with the system, the controllability of the system and the coordination of action. Thirdly, in nine of the articles, the motivation of the participants was discussed, which was expressed as the following: involvement with the task, attention and the arousal level of the participants. Motivation can show a sign of satisfaction with the system since individuals can become unmotivated when they unsatisfied with the system. Fourthly, the facet cybersickness was discussed in six of the articles. And lastly, the overall satisfaction was considered in seven of the articles, which included terms as comfort, quality, usable, helpful, participation and statements as: "Would you recommend the game" and "Would you play it on a daily basis?".

The second dimension, learnability, refers to the easiness of completing basic tasks in the system, was discussed in all but two of the articles (88.89%). However, all researchers consider different facets of this dimension. Five of the eighteen articles discussed confidence of the participants, four articles discussed the amount of time needed to complete the assignments and two groups reported the age was a dependent variable in learnability. In five articles it was kept general by saying the system was easy to use or user friendly in terms of learnability, three of those five linked this to the technical knowledge and skills of the participants. Also, the article of Lahiri *et al.* (2013) and Kuriakose *et al.* (2016) both reported that the engagement-based VR system ensured that more progress was made than the performance-based VR system. Lastly, two articles researched the feedback that was given in the VR environment.

The third dimension, efficiency, is measured once the participants have learned the design, how quickly they were able to perform the tasks given, was discussed fewer times,

namely in half of the eighteen articles. The most frequent named facets of this dimension are the intensity, difficulty and impact of the tasks that should be performed. Also, is the predictability of the different scenarios of the interventions discussed in two of the articles. The article of Gao *et al.* (2014) even researched the difference in efficiency between age and gender, which led them to conclude that females and age group 18-24, rather than age group 25-34 performed the tasks more efficiently. This in turn led to reduced stress in these groups. Besides, one of the articles compared the time it took the experts to complete the task to the time the participants needed.

The fourth dimension error looks into how many errors the participants make, which could be caused by bugs in the technology or because of learning mistakes. This dimension was discussed in five of the articles and mostly discussed the weaknesses of the system. In the study of Donker *et al.* (2019), 21.88% of the participants were not even able to participate in the intervention at all because of technical problems. In the study of Wong Sarver *et al.* (2013), there were some problems with installing the software which resulted in 36% of the participants needing technical support. In the articles by Tielman *et al.* (2017) and Yuen *et al.* (2019), only problems while using the software were experienced, such as problems with the sound or bugs in the system. Only one of the articles (Di Mascio, *et al*., 2019) named something positive considering the error rate, namely that the users were autonomous and did not need support during the intervention.

The last dimension, memorability, discusses whether the participants can remember the system after some time, was not discussed in any of the articles, despite its significance for recalling the system after some time of not using it and in creating a good usable intervention. This dimension is, therefore, also left out of the table.

Table 2

Usability dimensions named in the articles

Study	Satisfaction	Learnability	Efficiency	Error
Donker, *et al.* (2019)	Feels like being **present** in the environment 24 out of 193 participants experienced **Cybersickness**			21 out of 96 did not receive the intervention because of **technical problems**.
Fowler, *et al.* (2019)	Minor **Cyber sickness** reported and weight pain of the VR Gear. High session attendance, which showed **motivation**. Higher **immersion** rates in last 6 sessions compared to the first 3.	Participants needed more **time** than prescribed to start.	Increasing **intensity** less incremental than thought, participants learned fast over time.	
Paulus, Suryani, *et al.* (2019)	9 out of 30 didn't finish treatment because of cybersickness. Poor 3D quality causes **cybersickness**. In natural **engagement,** major problems found, namely the scaling of the map was not realistic, the flashlight reflection doesn't fit, and the animation doesn't look real. Also, major problems in **coordination of action** and r**epresentation** were discovered. Only accessory problems in a sense of **presence** and natural expression of action	Only some accessory problems with support for learning. **Realistic feedback** was given. Clear exit and entry points, as well as consistent departure, makes it easy to learn. Contrary, minor problems were found in navigation and orientation. Font size is too small. Two tested respondents were able to finish the task faster than the ideal **time** from the GOMS model.	The game at the next level is **predictable**. The participants can guess where the wood is placed based on experience at the previous level.	
Paulus, Yusuf, *et al.* (2019)	Application **motivates** participants to be braver in facing darkness and get used to the darkness. Good natural **engagement,** natural action. Very good sense of **presence** reported. Very good **compatibility** of the user's task and the domain were reported.	**Realistic feedback** reported. The exit point was hard to find. The environment supported the learning very good with **cues**.		

Table 2

Continued

Wong Sarver, *et al.* (2013)	**Quality** of the virtual environment program rated very good to excellent. 75% of the participants would **recommend** the program to others. The participants rated their comfort good to very good, which they also did to the extent to which they think the program was helpful. However, they rated the extent to which they would use the program again fair to good. At home, the children were not **motivated** to do the assignments which led to only 66 % of the tasks completed.	73% of the participants finished the treatment program. Participants rated their **confidence** good to very good and the helpfulness of the program as well. Learning how to use the intervention required minimal **time** and computer skills.		36% of the participants called when installing the program for **technical support**. Only one child asked for support during the home sessions.
Yuen *et al.* (2019)	87% of the participants reported complete satisfaction with the intervention, the other 13% reported to be mostly **satisfied**. All participants would **recommend** this intervention to a friend. 60% of the participants found the virtual homework was extremely **helpful** and 27 found this moderately. 47% of the participants agreed that the anxiety experienced during the virtual exercise was similar to **real-life** situations.	53% of the participants reported the system **easy** to use, the other 40% reported it as fairly easy to use.	73% of the participants strongly agreed and 27% agreed to the statement that the intervention **increased their ability** to cope with public speaking anxiety. 50% agreed that the intervention decreased their avoidance of public speaking situations.	7% of the participants often couldn't **hear** the facilitators and other participants speak. 27% of the participants sometimes experienced an echo of their voice while speaking.
Bekele, *et al.* (2016)	Decreased **engagement** over time. Good **controllability** and **interaction**.	**Confident** in making choices. Took more **time** than the control group.		The gaze group subjects had difficulty in understanding the occlusion paradigm at first as explicit guidance was not given to not bias the outcome.

Table 2

Continued

Di Mascio, *et al.* (2019)	No **cybersickness**. All participants were enthusiastic and **motivated** to participate in both sessions. Average engagement. The subject felt more **comfortable** facing simplified reality. Subjects had continuous **awareness** of the distinction between the real and the virtual worlds, which isn't necessarily negative because tells us that ASD people see IVE as a safe space. Lower **participation** score when using Oculus rift due to the cable connecting the headset to the computer. HoloLens doesn't require a physical connection.		Completed exercises within a **time** comparable to the time needed by the experts.	All participants were able to mount/dismount the headsets without support. High **autonomy** in managing the devices.
Kuriakose, *et al.* (2016) AS = Anxiety sensitive PS = performance sensitive	Sensors on hand to measure anxiety ensured decrease **immersion**.	6 out of 9 participants made vast **progress** in their performance.	More **difficult** task trials were performed in the AS system. In the PS system, almost no one of the participants came further than the first level.	
Lahiri, *et al.* (2013) ES = engagement-sensitive PS = performance-sensitive	All participants **completed** the sessions. Liked **interaction** with the system problems with wearing the eye-tracking. Improved **engagement** level while interacting with ES than with PS	**Understood** the narratives. Performance **improved** from the PS system to ES system.	Still questions regarding the specific **impact** of this system on learning over time and what specific mechanism may need adjustment to optimize efficient and relevant task adjustment.	
Tarantino, *et al.* (2019)	**Satisfied** facial expression, eager to start the experiment and excitement reported for all participants. Focused **attention** reported. Continuous **awareness** between the virtual and real world was reported.	The participants reported that they couldn't be scared in the virtual world, because they were **aware** of the difference between worlds, which is considered positive. After all, a safe space is created to practice and learn.		

Table 2

Continued

Gao, *et al.* (2014)	Limited **interactivity.**		Easy to learn, no **technician** needed **Confident** in using the system. Controls cumbersome in **use**	Participants in the 18-24 age group **scored** significantly higher in stress reduction than those of ages 25-34. **Females** had greater stress reduction than males.
Maarsingh, *et al.* (2019)	Participants were significantly more than required **involved** with the application.		Between above average and excellence **confidence** reported. User friendliness was reported in terms of learnability.	
Belloch, *et al.* (2014)	Strong sense of **presence**. High **interaction** No **cybersickness.** High **satisfaction** after the trial session. Anxiety is associated with emotional engagement and sense of **presence**. Anxiety and disgust levels increased as the virtual contamination increased, the anxiety produced by the exposure was related to the emotional engagement and sense of **presence.**		Easy to **use**. Participants in their midlife showed a lower score than participants in early adulthood on reality judgement, emotional engagement and sense of presence, this limited the learnability of these older-aged participants. High level of quality and **easiness.**	
Flujas-Contreras, *et al.* (2020)	Greater **motivation** towards therapy because of VR. Strong **presence** feeling. High ease of **interaction**.		High **confidence** scores. Easy to **use** VR treatment. More effective for participants younger than **age** 12.	
Gutiérrez-Maldonado, *et al.* (2016)	IM	60% drop out because of **cybersickness** when using VR glasses. Drop out 60%	**Use** requires significant technical knowledge	
	NI	Cybersickness reduced, dropout 0%		

Table 2

Continued

Vollstädt-Klein, *et al.* (2019)	The game was rated as **realistic**. Both groups rated their **arousal level** as average. From group one 76% would play the game **daily**, whereas in group two only 45% would do this.	The **controllability** was rated above average by the two groups. Of group one, reported 95% and in group two 100% of the participants that the game was **realistic** and included everyday situations.	The **predictability** was rated average and the **efficiency** was rated above average and good by the two groups.	
Tielman, *et al.* (2017)	The system was rated as well **usable**. Very **motivated** in giving feedback.	Bodily **participation** was rated lower for the Oculus rift because cable must be connected to the computer.		A **bug** was experienced by one participant who also as only one rated the usability 55 whereas the other participants rated it between 73 and 75

Note. The dimension memorability is left out of the table since it was not mentioned in any of the articles.

Usability measures

Four techniques were used to measure the usability of the interventions. From the articles discussed, 13 out of 18 articles used a self-report questionnaire. The questionnaires used can be found in Table 3.

Table 3

Questionnaires used to test usability dimensions

Questionnaire	Dimensions measured	Times used
System Usability Scale (SUS) (Lewis & Sauro, 2009)	Satisfaction and learnability	4
Symptoms of cybersickness questionnaire (SSQ) (Bruck & Watters, 2011)	Side effects of the VR technology	1
Visual Analogue Scales (VAS) (Wewers & Lowe, 1990)	Any kind of statement	1
Reality Judgement and Presence Questionnaire (RJPQ) (Baños et al., 2004)	Quality, reality judgment, presence, positive and negative issues, Interaction, Emotional engagement and Emotional Indifference	1
Igroup Presence Questionnaire (IPQ) (Schubert, 2003)	Spatial presence, involvement, and experienced realism	1
Usability and Acceptability Questionnaire (UAQ)	usefulness, reality and immersion	1
Charlestoneston Outpatient Satisfaction Questionnaire (Pellegrin, 2001)	Satisfaction	1
12 heuristic principles	Evaluating the virtual environment	1

In addition to the self-report questionnaires, in four of the articles, a combination of thinking aloud and observation was used to determine the usability of the interventions. In three of the articles, physiological measures were used, such as emotion recognition software, eye tracking, heart rate and skin temperature. And lastly, in four articles other methods were used such as measuring the attendance, homework compliance, the time needed and different difficulty levels. Besides, ten of the eighteen articles only used one method and seven of the articles used a combination of multiple methods, like thinking aloud and questionnaires. In the left article, the article of Gutiérrez-Maldonado et al. (2016), findings of the satisfaction and learnability were named but there was not reported how they researched these dimensions at all.

Limitations of usability methods

As can be seen in Table 3, the authors which are included in this review made choices that put limitations on testing the usability of the interventions. Firstly, in two out of the five articles about ASD self-report questionnaires were used to test the usability, however, this is not a reliable method to use considering that people with ASD have problems with explaining themselves, as was also stated in the research of Frith and Happé (1999). Secondly, four articles used some kind of explanation before starting to work with the intervention, such as a guiding manual, an explanation beforehand, and in the article of Belloch et al. (2014) was even chosen to instruct the users what to do throughout the experiment. Thirdly, in the article of Di Mascio et al. (2019) was chosen only to include intellectually high functioning participants which are attending either high school or University, which excludes a big part of the possible users. Also, in the article of Gutiérrez-Maldonado et al. (2016) only healthy participants were used, but still, they wanted to generalize the results to the patient group and in the article of Gutiérrez-Maldonado et al. (2016) a lack of variability in Bulimia scores in the patients was reported. In these three articles, it is not possible to generalize the results because a random sample of the whole patient population is required for a valid generalization.

Table 4

Study goals and limitations

Study	Goal	Limitations Methodology
Donker *et al.* (2019)	Examine the effectiveness of ZeroPhobia, a fully self-guided app-based VR cognitive behavioural therapy.	Because of following up after 3 months and dropout long term effects stay unknown.
Fowler *et al.* (2019)	Assess the feasibility of VR use through a pilot test.	The user manual was used.
Paulus, Suryani *et al.* (2019)	Evaluate the use of MAVR to treat fear of darkness based on usability, time consumption and its ability to decrease fear.	Received relaxation before going into MARV system.
Paulus, Yusuf, *et al.* (2019)	See if Night Forest application can be an alternative for treating Nyctophobia treatment.	
Wong Sarver *et al.* (2013)	Evaluating the feasibility, acceptability, and credibility of an interactive virtual school environment for the treatment of social anxiety disorder in preadolescent children.	Both children and their parents were provided with a 10-minute training to use the at-home software during the first generalization session.

Table 4

Continued

Yuen *et al.* (2019)13	Assess the feasibility and efficacy of a weekly group video conference intervention.	Ecological validity is limited because used in the lab and not on own home, as the intervention aimed to use it in own environment.
Bekele *et al.* (2016)	Test the usability of an adaptive multimodal virtual reality-based social interaction platform for children with ASD.	
Di Mascio *et al.* (2019)	Evaluating acceptability, usability, and engagement of the two HMDs	Only high functioning participants used which are attending either high school or University. Self-reported data used, could be difficult for people with ASD.
Kuriakose *et al.* (2016)	Assess the usability of a novel VR-based interactive system with Anxiety-Sensitive adaptive technology	
Lahiri *et al.* (2013)	Present the design and development of a VR-based engagement-sensitive system (ES) which can be applied to social communication tasks for children with ASD and present the results of a usability study designed to analyze the difference between the ES and PS system.	
Tarantino *et al.* (2019)	Evaluating the 2 VR glasses in terms of engagement to potential utilization of VR and MR in ASD-oriented applications.	Self-report used, difficult for people with ASD
Gao *et al.* (2014)	Analyzing the efficacy of using a mobile application to deliver relaxing virtual reality for psychological stress reduction	Participants were guided through the features of the app.
Maarsingh *et al.* (2019)	Examining whether StressJam is related to improvement in the stress mindset and examine the attractiveness.	Only self-motivated participants used, and only healthy participants used for the usability study.
Belloch *et al.* (2014)	Examining the utility of Virtual reality exposure therapy with OCD patients.	One patient did not receive exposure and response prevention, whereas others did. The therapist guides the patient through the activities, indicating what must be done moment-by-moment.
Flujas-Contreras *et al.* (2020)	Describe and compare the usability and acceptability of face-to-face treatment and a VR treatment for children and adolescents for long-term hospitalization.	
Gutiérrez-Maldonado *et al.* (2016)	Assessing whether exposure to food-related virtual environments could decrease food craving in a non-clinical sample.	Exclusion criteria based on self-report. Only healthy participants included but wanted to generalize it to patients with Bulimia nervosa or Binge eating disorder. And lack of variability in Bulimia scores.

Table 4

Continued

Vollstädt-Klein *et al.* (2019)	Evaluate the feasibility of a web-based gaming tool for individuals with AUD.	The application was initially designed for AUD patients but also CUD patients were involved.
Tielman *et al.* (2017)	Evaluating the usability of the system, how much the systems contribute to the therapy and to which extent traumatic memories are recollected appropriately.	

Note. In the empty spaces, no limitations were found in the text.

Discussion

This study aimed to investigate which dimensions of usability are frequently used to assess the usability of VR application in mental health treatment. The results obtained during the literature review show that there certainly is a difference to which extent the five usability dimensions of Nielsen (1994), satisfaction, learnability, efficiency, errors and memorability are discussed in mental health care VR studies, which is going to be discussed in the following sections.

Usability dimensions

Five usability dimensions were studied in this review and show some important findings. First, it was found that all researchers consider the usability dimension *satisfaction* as most important since all of the articles discussed this dimension and this dimension was discussed most thoroughly. In all articles, most information could be found on satisfaction. This could be explained by the fact that the satisfaction of the users determines the user experience and whether they are going to use and recommend the treatment, which is what you aim for as a researcher. Satisfaction certainly is an important dimension, however, such a strong focus on one dimension does not ensure a good usable end product. To achieve good usability all dimensions should be considered because when high standard in the other dimensions are upheld this will ensure that the users like interacting with the system and so the satisfaction will be rated high.

Secondly, the dimension *learnability* was researched to a great extent in the articles which indicates that it is seen as an important dimension as well. For this dimension was named that different age groups learn differently, and that confidence differed per individual, which shows the importance of a good learnable design of the VR environment for all individuals despite their age. This complies with the existing body of literature which states that older adults have more difficulties with learning how to interact with VR environments (Schraepen, 2017). These findings show that a random sample should be included in which individuals with a wide range of age and characteristics rate the usability, in this way, the system can be used to help as many individuals as possible. Because as named in the literature earlier (Nielsen, 1994), the system should be easy to use and learn, which should apply to most of the users. However, some studies show a lack in variability, as only intellectually high functioning participants were included, or the group participants lacked variability in the strength of the disorder. This causes an impossibility to generalize the results and to make

judgements about the applicability of the VR treatment. Undoubtedly there could still be paid more attention to this dimension in the remaining articles to tailor the design to all its users.

Thirdly, the dimension efficiency was researched to an average extent. Much attention was paid to the predictability and intensity of the VR environments, also, there was remarked a difference in age again. As mentioned before, this could be explained by the affinity individuals in their early adulthood have with a new technology which ensures that they be able to manage the technology better and be able to work more efficient. Besides, this could be hard to change since people over 60 years have more difficulty with working with new technologies or systems and this makes efficient working harder to achieve (Becker, 2004).

Fourthly, the dimension error was considered little in the articles discussed, only five articles mentioned the problems participants experienced during the intervention, besides, in many articles the dimension was not even mentioned in terms of no problems, which we are left to assume when the dimension is not mentioned in the results. In all four dimensions, the technical side of the VR interventions was mentioned and nothing about the VR environment itself, which could indicate that more attention is paid to the working of the equipment rather than the actual treatment itself. Although, the fact that the technology should be working, the interaction with the virtual environment is as important for the treatment of the mental disorder and so should also be evaluated.

The last dimension, memorability, is the most neglected dimensions of usability studies into the use of VR treatment in mental health, considering that it was not mentioned in any of the articles. Besides, all of the articles used multiple treatment moments at which the VR intervention is used, and memorability could be measured, in the articles was not given an explanation for this choice, this indicates that this dimension is not seen as important and thus highly neglected in the studies presented. A possible explanation for this could be that the researched did not expect any difficulties with the participants returning to the system after some time of not using it and expected that the system should be easy to remember. Moreover, it could be argued that VR is so engaging that memorability is not an important aspect for this sort of technology. The dimension memorability is not seen as important in VR usability studies which definitely must change for long term satisfaction of the users.

Usability measures

In the results the measurement methods of the articles were discussed, which mostly included self-report questionnaires, but also physiological measures and other measurement methods were used. The first intention of this review was to research how each different

dimension of usability was measured, however, when reading the articles, it was noticeable that the way of measuring the dimensions was deemed inappropriate. This was found in the lack of information about the method of measurement in the articles, only the results were mentioned. Besides, in one article the SUS was used only to measure satisfaction, whereas this questionnaire measures satisfaction and learnability of the intervention. Also, in the article of Flujas-Contreras et al. (2020) the usability is tested with 10 items, whereas, in the article of Belloch et al. (2014), the usability is measured with 56 items, which is a big difference. Besides, the factors dimensions measured with these questionnaires do not fit with dimensions of Nielsen. In the researcher's opinion, the difference and unclarity express that there still is not a one-way ticket for efficiently testing the usability into mental health VR use. However, this could be caused by the difference in study characteristics, such as studies into ASD targeted intervention for which self-reported measures are not applicable, which also can be recognized in the limitations reported in Table 3. A better option would be to pay systematic attention to the dimensions that fit the application evaluated and how to measure these dimensions. Besides, as can be seen in Table1, most of the authors did not include or name a framework in their articles, which could also explain the difference in measurement methods. Because when not knowing much about the different dimensions of usability study and not using a framework, this can lead to not all dimensions being evaluated.

Limitations of usability methods

The limitations discussed present the difficulties that are experienced while choosing the measurement methods of a usability study, namely, that every study is different and so the best fitting choices must be made in testing the usability and that there is not a one-way ticket as said earlier. Not just one method can be applied, which can be seen in the fact that studies evaluating VR treatment for individuals with ASD use self-report methods, which is mostly used. However, this is not a reliable method to use considering that people with ASD have problems with explaining themselves, as was also stated in the research of Frith and Happé (1999). Also, when testing the usability you want to know whether the participants can use the interventions by themselves or if one of the usability dimensions need to be changed, if you explain beforehand or during the whole experiment what should be done, whereas in real life they need to figure this out on their own, in this way, the real usability problems will not be exposed.

Besides, some factors do not fit, in the framework of Nielsen, for example, immersion is named in this review under the dimension satisfaction whereas this factor needs an own

dimension because of its importance for VR use. Also, the factor motivation is named under the same dimension, which also could be named under the dimension learnability since it was not named in the articles clearly and the framework of Nielsen also does not give a statement on this. This fits with the existing literature which states that Nielsen (1994) uses general dimensions and therefore lacks critical elements of a specific application (Quiñones, & Rusu, 2017). This shows that even though Nielsen's framework is meant for user interfaces, it does not entirely fit with measuring the usability of VR technology.

Additional findings on the topic usability

Based on the findings it can be stated that in most of the articles much attention is devoted to creating something similar to the real world of their target group and as realistic as real-life to provide the ultimate learning experience and elicit "normal" behaviour. However, from most of the articles, it can be concluded that this is not feasible because the design is not made for one individual but different kind of individuals. For example, in the study of Gao *et al.* (2019) participants in adolescence, in first four years of early adulthood, and females were able to work more efficient and so get more out of the treatment. This goal also ensures that when doing usability studies, the virtual world is being compared to the real world, which is a standard that is hard to achieve.

In the study of Mascio *et al.* (2019) and Tarantino *et al.* (2019), the participants reported that they could not be scared because they knew what they saw was not real, this shows that how real you make the environment the patients you are treating will never have the same experience as in real life. In the case of the previously mentioned articles, however, this was seen as positive since it creates a safe space to practice for the real world.

The study of van Gisbergen (2019) also suggested that realism is not the key to VR interventions, in this study was researched what difference the realism makes for the behaviour showed by the participants. The study showed that the level of realism does not make a difference in the natural behaviour the users expressed. Furthermore, as mentioned in the introduction, Sutcliffe and Kaur (2000) also state that this focus on "naturalness" can not only ensure positive results but can ensure that the ease of use is in danger. Because so much focus is directed to making the environment realistic other important usability dimensions are often not fully discussed, as can be seen in Table 2, only one article discussed four of the five usability dimensions, so not even all dimensions. Which could be because few of the authors followed a usability framework and when they did the most did not include the framework of Nielsen as can be seen in Table 1. Consequently, could be asked, whether such strong focus

on satisfaction and no handling of a usability framework suggests that the authors are measuring user experience rather than the usability of the intervention. For a better-quality intervention and article, the preferable option would be to consider all the usability dimensions and not some in particular.

Also, can be noticed that there is no information about all mental disorders, but a very specific group of disorders. In the article of Eichenburg and Walters (2012) they discuss the current state of research about VR treatment in mental health care, which shows that the disorders discussed in this review are not random but merely all we know until now. The article discusses specific phobias (also social), panic disorder, OCD, eating disorders, schizophrenia, attention deficits and addictions. The only disorders that lack in this review are attention deficits and schizophrenia which can be caused by the VR treatment of these disorders still coming up, whereas more attention is devoted to the other disorders. There is not said why these disorders are studied, only for phobias is said that it is treated with the aid of exposure, this makes it suitable for VR treatment. However, in the researcher's opinion, this aid of exposure also can be suitable for the other currently researched disorders treated with VR.

When taking all the information into account and being well educated in the field, the researcher would suggest a new usability framework which leads to an optimal combination of the dimensions. In this four of the dimensions of Nielsen (1994) would be included, namely, *satisfaction, learnability, efficiency* and *error,* a new dimensions "*immersion*" should be created, and the dimension memorability should be addressed under the topic efficiency. This combination of dimensions is chosen based on the importance of the dimensions. The dimensions satisfaction, learnability and efficiency were already named in the articles often, however, in my opinion, the dimension error should be given more priority since this involves an important part of the functioning of the intervention itself and, accordingly, determines a part of the treatment outcome. Also, the dimension memorability was not given any priority and could be assessed in the dimension learnability because of its similarity. And lastly, there should be included a dimension "immersion" which is a specific attribute of VR which cannot and should not be included in one of the other dimensions. The "Binnenmars framework" including the five dimensions satisfaction, learnability, efficiency, error and immersion could be an opportunity for better usability testing of VR not only in mental health treatment but all VR use, which will lead to better treatment.

Conclusion

The use of VR in the treatment of mental health disorders is increasing. Nowadays, this technology is not only being used by experts but by a larger population. This literature review shows that usability studies evaluating these VR treatments did not pay attention to al the five different dimensions which Nielsen (1994) identified, especially since one dimension was not mentioned in any of the articles. This causes doubts about whether the five dimensions of Nielsen are the ultimate dimensions for testing usability of VR treatment used in mental health care. The dimensions satisfaction and learnability are unjustifiably more prominent discussed in comparison with the other three dimensions, efficiency, errors and memorability. The lack of discussion of the dimension memorability suggests that this dimension is not relevant for this technology, still, this could be important for more difficult technologies. This should be considered when studying further into the usability of VR treatment. Also, the methods that are used leave, some questions unanswered, for example, why is there so much difference in the methods and number of items used to measure the dimensions? Why do some studies guide the users through the system when you want to know how they experience the usability themselves? This shows again that there should be paid more attention to all the different dimensions and how these can be measured most Reliable. However, it is still difficult to make the system optimally usable for all patients since the difference in age and intelligence makes it hard to take everything into account before putting it into practice. But as mentioned at the beginning of this literature review; one in ten individuals experience a mental illness (Ritchie, 2018), because of this, it is important that new systems with good usability are designed so more people can experience the benefits.

Several limitations should be addressed in this review. As all the articles also mentioned themselves, the sample size in the articles is too small and lack diversity to make generalizations to the whole group of individuals with these disorders. Nonetheless, do these separate small studies on the same disorder, like the six on anxiety and five on ASD, make a bigger collective contribution to the knowledge of VR use in the treatment of these disorders. Another limitation is the fact that in much articles the dimensions are researched but not mentioned in the results or any tables which makes this literature review to some extent. incomplete. Besides, this study only focused on experimental designs to create homogeneity and because the interventions discussed are further in the design process which makes that the usability can be tested better since in the designing phase startup problems are already improved in this process. However, this could be seen as a limitation since much usability tests are also performed in a non-experimental setting and because these startup problems are

already fixed, and certain dimensions are not given a high priority anymore. Another limitation is that five articles were excluded because the full text could not be located, namely, Moller, et al., 2017; Rivera, et al.,2015; Krysta et al., 2017; Quero et al., 2017; and Cárdenas et al., 2016.

Future research should be done into finding the right framework to study the usability of VR treatment. In this, terms specific to VR, such as immersion, should be included and dimensions not relevant for VR technology, as memorability, should be excluded. However, there should be done more research into how to measure the usability efficiently, because in this literature review came forward that there is not a one way to measure the different dimensions of the usability and often multiple different self-report questionnaires are used. This led to rather evaluating the user experience rather than the usability itself. A possible answer to this could be a usability guide for different interventions with their specific focus of attention, which would make it able to compare the usability of different types of interventions.

This literature review distinguishes itself in the field because of its specific focus on usability and mental health, which gives an overview of the disorder treatable with VR until now. From this literature review can be concluded that more systematic attention should be paid to the different dimensions of usability and focus on one dimension of the intervention should be avoided. More attention could be devoted to the methods used to measure usability since most articles did not choose fitting instruments to measure usability. Most importantly, it should be considered whether the dimensions used do fit with the intervention you are evaluating, which will result in an adequate evaluation of the usability rather than the user-experience. The framework of Nielsen (1994) is not the optimal usability framework, as some dimensions are still questionable. This difference and unclarity in usability studies into mental health care treatment shows there still is not a one-way ticket for efficiently testing the usability. Could the “Binnenmars framework” perhaps be this one-way ticket?

References

Baños, R. M., Botella, C., Liaño, V., Guerrero, B., Rey, B., & Alcañiz, M. (2004). Sense of presence in emotional virtual environments. *Proceedings of presence*, 156-159. https://www.researchgate.net/profile/Cristina_Botella/publication/228880180_Sense_of_presence_in_emotional_virtual_environments/links/0912f50826d1e32ed9000000.pdf

Becker, S. A. (2004). A study of web usability for older adults seeking online health resources. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *11*(4), 387-406. https://doi.org/10.1145/1035575.1035578

Bekele, E., Wade, J., Bian, D., Fan, J., Swanson, A., Warren, Z., & Sarkar, N. (2016). Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD). *IEEE Virtual Reality,* 121-130.. https://doi.org/10.1109/vr.2016.7504695

Belloch, A., Cabedo, E., Carrió, C., Lozano-Quilis, J. A., Gil-Gómez, J. A., & Gil-Gómez, H. (2014). Virtual reality exposure for OCD: is it feasible? *Revista de Psicopatología y Psicología Clínica*, *19*(1), 37-44. https://doi.org/10.5944/rppc.vol.19.num.1.2014.12981

Bruck, S., & Watters, P. A. (2011). The factor structure of cybersickness. *Displays*, *32*(4), 153-158. https://doi.org/10.1016/j.displa.2011.07.002

Burridge, J. H., & Hughes, A. M. (2010). Potential for new technologies in clinical practice. *Current opinion in neurology*, *23*(6), 671-677. https://doi.org/10.1097/wco.0b013e3283402af5

Cárdenas, G., Botella, C., Quero, S., Baños, R., Durán Baca, X., De la Rosa, A., Flores, L., & Molés Amposta, M. (2016). Efectividad del programa de tratamiento virtual flight con población mexicana. *Fundación Aiglé, XXV (2),* 145-156. http://repositori.uji.es/xmlui/handle/10234/165657

Di Mascio, T., Tarantino, L., De Gasperis, G., & Pino, C. (2019). Immersive Virtual Environments: A Comparison of Mixed Reality and Virtual Reality Headsets for ASD Treatment. *International Conference in Methodologies and intelligent Systems for Technology Enhanced Learning* (pp. 153-163). https://doi.org/10.1007/978-3-030-23990-9_19

Donker, T., Cornelisz, I., Van Klaveren, C., Van Straten, A., Carlbring, P., Cuijpers, P., &

Van Gelder, J. L. (2019). Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. *JAMA psychiatry*, *76*(7), 682-690. https://doi.org/10.1001/jamapsychiatry.2019.0219

Eichenberg, C., & Wolters, C. (2012). Virtual Realities in the treatment of mental disorders: A Review of the current state of research. Virtual reality in psychological, medical and pedagogical applications, 2, 35-64. http://dx.doi.org/10.5772/2607

Flujas-Contreras, J. M., Ruiz-Castañeda, D., & Gómez, I. (2020). Promoting Emotional Wellbeing in Hospitalized Children and Adolescents with Virtual Reality: Usability and Acceptability of a Randomized Controlled Trial. *CIN: Computers, Informatics, Nursing*, *38*(2), 99-107. https://doi.org/10.1097/cin.0000000000000586

Fowler, C. A., Ballistrea, L. M., Mazzone, K. E., Martin, A. M., Kaplan, H., Kip, K. E., Ralston, K., Murphy, J. L., &Winkler, S. L. (2019). Virtual reality as a therapy Adjunct for Fear of Movement in Veterans With Chronic Pain: Single-Arm Feasibility Study. *JMIR Formative Research*, *3*(4), 1-12. https://doi.org/10.2196/11266

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological Medicine*, *47*(14), 2393-2400. https://doi.org/10.1017/S003329171700040X

Frith, U., & Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind & Language*, *14*(1), 82-89. https://doi.org/10.1111/1468-0017.00100

Gao, K., Boyd, C., Wiederhold, M. D., & Wiederhold, B. K. (2014). VR mobile solutions for chronic stress reduction in young adults. *Annual review of Cybertherapy and Telemedicine 2014*, 88-93. doi:10.3233/978-1-61499-401-5-88

Van Gisbergen, M., Kovacs, M., Campos, F., van der Heeft, M., & Vugts, V. (2019). What we don't know. the effect of realism in virtual reality on experience and behaviour. *Augmented Reality and Virtual Reality* (pp. 45-57). https://doi.org/10.1007/978-3-030-06246-0_4

Glegg, S. M., Tatla, S. K., & Holsti, L. (2014). The GestureTek virtual reality system in rehabilitation: a scoping review. *Disability and Rehabilitation: Assistive Technology*, *9*(2), 89-111. https://doi.org/10.3109/17483107.2013.799236

Gutiérrez-Maldonado, J., Pla-Sanjuanelo, J., & Ferrer-García, M. (2016). Cue-exposure software for the treatment of bulimia nervosa and binge eating disorder. *Psicothema*, *28*(4), 363–369. https://doi.org/10.7334/psicothema2014.274

ISO 9241: Ergonomics Requirements for Office Work with Visual Display Terminals (VDTs)

International Standards Organisation, Geneva (1997).

Krysta, K., Krzystanek, M., Cubała, W. J., Wiglusz, M. S., Jakuszkowiak-Wojten, K., Gałuszko Wegielnik, M., & Janas-Kozik, M. (2017). Telepsychiatry and virtual reality in the treatment of patients with intellectual and developmental disabilities. *Psychiatria Danubina*, *29*(3), 656-659. https://pubmed.ncbi.nlm.nih.gov/28953847/

Kuriakose, S., & Lahiri, U. (2016). Design of a physiology-sensitive VR-based social communication platform for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *25*(8), 1180-1191. https://doi.org/10.1109/tnsre.2016.2613879

Lahiri, U., Bekele, E., Dohrmann, E., Warren, Z., & Sarkar, N. (2012). Design of a virtual reality based adaptive response technology for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*(1), 55-64. https://doi.org/10.1109/tnsre.2012.2218618

Lewis, J. R., & Sauro, J. (2009, July). The factor structure of the system usability scale. In *International conference on human centered design* (pp. 94-103). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-02806-9_12

Luxton, D. D., McCann, R. A., Bush, N. E., Mishkind, M. C., & Reger, G. M. (2011). mHealth for mental health: Integrating smartphone technology in behavioral healthcare. *Professional Psychology: Research and Practice*, *42*(6), 505. https://doi.org/10.1037/a0024485

Nielsen, J. (1994). What is usability? In *Usability engineering* (pp. 23–43). Morgan Kaufmann. https://doi.org/10.1016/b978-0-08-052029-2.50007-3

Nielsen, J. (2012, January 3). Usability 101: Introduction to Usability. Retreived from https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Maarsingh, B. M., Bos, J., Van Tuijn, C. F., & Renard, S. B. (2019). Changing Stress Mindset Through Stressjam: A Virtual Reality Game Using Biofeedback. *Games for health journal*, *8*(5), 326-331. https://doi.org/10.1089/g4h.2018.0145

Merriam-Webster. (n.d.). Virtual reality. In *Merriam-Webster.com dictionary*. Retrieved February 29, 2020, Retrieved from https://www.merriamwebster.com/dictionary/virtual%20reality

Moller, H. J., Saynor, L., Bal, H., Sudan, K., & Jones, L. (2017). Putting immersive therapies into PRAXIS: Towards holistic wellbeing multisensory meditation environments. *Journal of Alternative Medicine Research*, *9*(2), 153.

https://search.proquest.com/openview/9887f053f446ae7061aabd8e4c9b6493/1?pq-origsite=gscholar&cbl=2034852

Orsmond, G. I., & Cohn, E. S. (2015). The distinctive features of a feasibility study: Objectives and guiding questions. *OTJR: occupation, participation and health*, *35*(3), 169-177. https://doi-org.proxy-ub.rug.nl/10.1177/1539449215578649

Paul, L. A., Hassija, C. M., & Clapp, J. D. (2012). Technological advances in the treatment of trauma: a review of promising practices. *Behavior modification*, *36*(6), 897-923. https://doi.org/10.1177%2F0145445512450733

Paulus, E., Suryani, M., Wijayanti, P. A. K., Yusuf, F. P., & Iskandarsyah, A. (2019). The use of mobile-assisted virtual reality in fear of darkness therapy. *Telkomnika*, *17*(1), 282-290. https://doi.org/10.12928/telkomnika.v17i1.11614

Paulus, E., Yusuf, F. P., Suryani, M., & Suryana, I. (2019, June). Development and Evaluation on Night Forest Virtual Reality as Innovative Nyctophobia Treatment. In *Journal of Physics: Conference Series* (Vol. 1235, No. 1, p. 012003). IOP Publishing. https://doi.org/10.1088/1742-6596/1235/1/012003

Pellegrin, K. L., Stuart, G. W., Maree, B., Frueh, B. C., & Ballenger, J. C. (2001). A brief scale for assessing patients' satisfaction with care in outpatient psychiatric services. *Psychiatric Services*, *52*(6), 816-819. https://doi.org/10.1176/appi.ps.52.6.816

Quero, S., Andreu-Mateu, S., Moragrega, I., Baños, R. M., Molés, M., Nebot, S., & Botella, C. (2017). Acognitive-behavioural therapy program utilizing virtual reality for the treatment of adjustment disorders : Acase series. Revista Argentina de Clínica Pisicológica, 26(I), 5. Https://doi.org/10.24205/03276716.2017.1001

Quiñones, Daniela, and Cristian Rusu, (2017) How to develop usability heuristics: A systematic literature review. *Computer Standards & Interfaces* 53 (2017): 89-122. https://doi.org/10.1016/j.csi.2017.03.009

Ritchie, H. (2018). Mental Health. https://ourworldindata.org/mental-health

Rivera, R. M. B., Arbona, C. B., García-Palacios, A., Castellano, S. Q., & López, J. B. (2015). Treating emotional problems with virtual and augmented reality. *The handbook of the psychology of communication technology*, *32*, 548. https://doi.org/10.1002/9781118426456.ch25

Rizzo, A., & Koenig, S. T. (2017). Is clinical virtual reality ready for primetime? *Neuropsychology*, *31*(8), 877. https://doi.org/10.1037/neu0000405

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), 362-391. https://doi.org/10.1017/S0033291717001945

Schraepen, B., Huygelier, H., Van Ee, R., Vanden Abeele, V., & Gillebert, C. (2017). Attitudes towards and Usability of Head-mounted Virtual Reality in Older Adults. *Amsterdam, The Netherlands*. https://lirias.kuleuven.be/1674377?limo=0

Schubert, T. (2003). The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness. Zeitschrift fur Medienpsychologie, 15, 69-71. https://doi.org/10.1026//1617-6383.15.2.69

Slater M, Sanchez-Vives M (2016). Enhancing our lives with immersive virtual reality. Frontiers in Robotics and AI, 3. https://doi.org/10.3389/frobt.2016.00074

Sutcliffe, A. G., & Kaur, K. D. (2000). Evaluating the usability of virtual reality user interfaces. *Behaviour & Information Technology*, *19*(6), 415-426. https://doi.org/10.1080/014492900750052679

Tarantino, L., De Gasperis, G., Mascio, T. D., & Pino, M. C. (2019, November). Immersive applications: what if users are in the autism spectrum? *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry* (pp. 1-7). https://doi.org/10.1145/3359997.3365696

Tielman, M. L., Neerincx, M. A., Bidarra, R., Kybartas, B., & Brinkman, W. P. (2017). A therapy system for post-traumatic stress disorder using a virtual agent and virtual storytelling to reconstruct traumatic memories. *Journal of medical systems*, *41*(8), 125. https://doi.org/10.1007/s10916-017-0771-y

Vollstädt-Klein, S., Mildner, P., Bumb, J. M., Karl, D., Ueberle, C., Shevchenko, Y., ... & Effelsberg, W. (2019). The training game SALIENCE for the therapy of alcohol use disorder. *Health informatics journal*, 1460458219839612. https://doi.org/10.1177/1460458219839612

Yuen, E. K., Goetter, E. M., Stasio, M. J., Ash, P., Mansour, B., McNally, E., Sanchez, M., Hobar, E., Forte, S., Zulaica, K., & Watkins, J. (2019). A pilot of acceptance and commitment therapy for public speaking anxiety delivered with group videoconferencing and virtual reality exposure. *Journal of Contextual Behavioral Science*, *12*, 47-54. https://doi.org/10.1016/j.jcbs.2019.01.006

Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in nursing & health*, *13*(4), 227-236. https://doi.org/10.1002/nur.4770130405

Wong Sarver, N., Beidel, D. C., & Spitalnick, J. S. (2014). The feasibility and acceptability of virtual environments in the treatment of childhood social anxiety disorder. *Journal of Clinical Child & Adolescent Psychology*, *43*(1), 63-73. https://doi.org/10.1080/15374416.2013.843461

Zajtchuk, R., & Satava, R. M. (1997). Medical applications of virtual reality. *Communications of the ACM*, *40*(9), 63-64. https://doi.org/10.1145/260750.260768