Universität Basel

Fakultät für Psychologie

Beyond the screen –The potential of smartphone apps and immersive technologies in exposure-based interventions for phobias

A cumulative dissertation submitted to the Faculty of Psychology, University of Basel, in partial fulfillment of the requirements for the degree of Doctor of Philosophy by

Anja Zimmer

from Basel

Basel, 2022

Universität Basel

Fakultät für Psychologie

Approved by the Faculty of Psychology at the request of

Prof. Dr. med. Dominique J.-F. de Quervain

Prof. Dr. Tanja Michael (Uni Saarland)

Date of the doctoral examination:

Dean of the Faculty of Psychology

Universität Basel

Fakultät für Psychologie

Declaration of Scientific Fairness

I, **Anja Zimmer**, hereby declare that the present work was written independently without the help of third parties and without the use of any means other than those indicated. Sources used for help are marked as such. The manuscripts published or submitted for publication in journals were prepared in cooperation with the co-authors and were not published elsewhere by any of the participants, submitted for publication, or submitted to any other examination authority as a qualification paper. These are the following manuscripts:

A) Bentz, D., Wang, N., Ibach, M. K., Schicktanz, N. S., **Zimmer, A.**, Papassotiropoulos, A., & de Quervain, D. J. (2021). Effectiveness of a stand-alone, smartphone-based virtual reality exposure app to reduce fear of heights in real-life: A randomized trial. *NPJ digital medicine*, *4*(1), 1-9.

B) Mueller, F. D., Fehlmann, B., Wang, N., Ibach, M. K., Schlitt, T., Bentz, D., **Zimmer, A.**, Papassotiropoulos, A. & de Quervain, D. J. (2022). Virtual reality-based gaze exposure treatment reduces fear of public speaking. *PNAS (submitted for publication)*

C) **Zimmer, A.**, Wang, N., Ibach, M. K., Fehlmann, B., Schicktanz, N. S., Bentz, D., Michael, T., Papassotiropoulos, A. & de Quervain, D. J. (2021). Effectiveness of a smartphone-based, augmented reality exposure app to reduce fear of spiders in real-life: A randomized controlled trial. *Journal of anxiety disorders*, *82*, 102442.

Basel, 2.5.2022

Anja Zimmer

Abstract

Specific phobias are extremely common among adults. They are characterized by strong emotional reactions and avoidance behavior when exposed to the feared stimuli. Specifically fears concerning heights or animals such as spiders are highly prevalent, followed by fear of social situations such as fear of public speaking. The gold standard in treating specific phobias is exposure-based therapy. However, exposure-based therapy is limited in its practicability in clinical routine and poses a high hurdle for affected individuals. Virtual and augmented reality (VR/AR) smartphone apps offer attractive platforms to simulate exposure situations and by that increase the accessibility of mental health services in general. Thus, novel smartphone-based treatments hold the potential to facilitate the dissemination of exposure-based treatments for specific phobias. The studies presented as part of this thesis aimed at investigating three newly developed interventions for fear of heights, fear of public speaking and fear of spiders, using the currently available advanced technologies.

In the first study (Bentz et al., 2021), a stand-alone, automated and gamified VR exposure app *Easyheights* was developed using 360° images. The app's effectiveness to reduce fear of heights and avoidance behavior was investigated in a randomized controlled trial in an adult population with clinical and subclinical fear of heights. The repeated use of the app led to reduced fear and avoidance behavior in a real-life situation on a tower.

For the second study (Müller, Fehlmann et al., 2022), the developed stand-alone, automated and gamified VR exposure app *Fearless Speech* aimed at reducing public speaking anxiety (PSA) and avoidance of eye contact. A virtual audience with 360° videos was used for the exposure and gaze control for the eye contact training. The app was investigated in a randomized controlled trial in healthy adults with subclinical PSA. After the repeated use of the app, participants showed reduced fear and improved eye contact in a real-life speech situation.

The third study (Zimmer et al., 2021) examined the developed stand-alone, automated and gamified AR exposure app *Phobys.* In comparison to VR, AR has only recently been introduced to clinical research. The app was designed to reduce fear, disgust and avoidance behavior in adults with clinical and subclinical fear of spiders. The results of the randomized controlled trial showed that repeatedly using the app led to reduced fear, disgust and avoidance behavior in a real-life situation with a real spider.

The results of these studies support the potential of stand-alone, automated VR and AR interventions delivered through smartphone apps. The developed apps allow for a high-quality user experience with a highly realistic environment, gaze control for an easy navigation as well as the possibility of interaction. In addition, gamification elements foster engagement with the apps. All three investigated apps offer low-threshold and low-cost treatment for individuals affected by specific phobias. Testing the effectiveness of these newly developed apps in real-life settings sets them apart from previous studies. Hence, this thesis highlights the potential of using smartphone apps with immersive technologies to advance and disseminate exposure-based treatments for specific phobias.

Acknowledgements

During my four years as a PhD candidate I was able to work in a interdisciplinary and collaborative team, alongside with excellent researchers and inspiring people to grow as a researcher and psychologist.

My deepest gratitude goes to my two supervisors Prof. Dr. Dominique de Quervain and Prof. Dr. Tanja Michael. Their expertise, guidance and appreciation have encouraged and supported me throughout my whole doctorate. Prof. de Quervain has continually supported me and my ideas since day one and even before. He trusted and challenged me as a researcher and study team leader, allowed me to experiment, learn and flourish. I thank Prof. Dr. Tanja Michael from the University of Saarland, Germany for her expertise, feedback and support from afar. I could not be happier to have had such warmhearted and supportive supervisors and it was an honor to work with them.

Further thanks go to the Swiss National Science Foundation (SNSF) for financially supporting my PhD with the Doc.CH grant: P0BSP1_181761 / 1.

I want to thank the whole team of the Transfaculty Research Platform of Molecular and Cognitive Neurosciences (MCN) of the University of Basel for its collaborative and multidisciplinary environment. Especially, Dr. Nan Wang for programming and Merle Ibach for the UI/UX design of the apps. It has been such an enriching and challenging experience, to work in a small, interdisciplinary and creative team to bring my ideas to (virtual/augmented) reality. Special thanks also go to Dr. Bernhard Fehlmann for supporting me in the application process of the SNF grant and within my own project. Thank you for your commitment, intelligence and humor. It was the Masters project of Dr. Dorothée Bentz that sparked my interest in research on immersive technologies. It has been a pleasure to have her as my colleague and teaching partner. In addition, I thank all the students, assisting in the studies and allowing me to grow as a supervisor and team leader. One of the happiest moments was to receive the highly competitive Doc.CH grant from the SNF. Not only for allowing me to pursue my research project but especially to be able to work alongside my co-PhD student and friend Dr. Priska Zuber. Thank you for all the coffee chats, the walks with your dog, the lunch breaks, career planning and inspirational conversations in person and over zoom during the pandemic. Additionally, I am grateful for the opportunity to do a research stay the Freie Universität in Berlin to get enriching insights into the research projects of Prof. Dr. Christine Knavelsrud and for the exhilarating environment of the city of Berlin to write my dissertation.

Next to Dr. Priska Zuber and Dr. Bernhard Fehlmann, I also want to thank the rest of my PhD family with Fabian Müller, Galya Iseli, Ehssan Amini and Silvan Pless for all the tea and coffee breaks, the dinner tables and supportive conversations. Although the work environment played a huge role in this process, I would have not been able to start, continue and finish this journey without the support of my family and friends, of which I'd like to name a few: my parents, my brother Fabian and my beloved friends Lesley, Cécile, Sophie, Laura, Caroline and Anat. A special thanks also goes to the student organization Mind-Map, a side but passionate project of mine, allowing me to develop myself as an activist for mental health.

Table of contents

Abbreviations

AR	Augmented Reality
ARET	Augmented Reality Exposure Therapy
APA	American Psychological Association
GPS	Global Positioning System
HMD	Head-mounted display
NIMH	National Institute for Mental Health
MR	Mixed Reality
OCD	Obsessive and Compulsory Disorder
PSA	Public Speaking Anxiety
PTSD	Post-Traumatic Stress Disorder
RCT	Randomized Controlled Trials
SAD	Social Anxiety Disorder
UI	User Interface
UX	User Experience
VR	Virtual Reality
VRET	Virtual Reality Exposure Therapy

"As technology evolves so does the need for humans to do the same. We accomplish this by overcoming things that keep us reigned in, things like fear." (Mubarak, 2021)

1 Introduction

In today's society there is an urgent need to pay more attention to mental health. One in five adults suffers from a mental health disorder at some point in their life (Steel et al., 2014). Worldwide, mental health disorders climb in their ranking for years lived with disability, leading to a higher need of health care resources (Vigo et al., 2016; Sporinova et al., 2019). The COVID-19 pandemic, among other factors, has highlighted the existing gaps in our health care system, especially in the resources available to deal with mental health problems (Baños et al., 2022). Although evidence-based psychological treatments are available for the general population, access to them remains difficult due to stigma, financial aspects, geographical or time limitations (Henderson et al., 2013; Harvey & Gumport, 2015; Thyloth et al., 2016).

Digital mental health tools provide the opportunity to make psychological support accessible and engaging. Specifically, smartphone applications (apps) are an attractive platform to deliver personalized interventions with instant feedback and to process large amounts of data to share with clinicians and researchers (Alqahtani & Orji, 2020; Bhugra et al., 2017). Evidence on the efficacy of mental health apps has created a wave of enthusiasm and support, highlighting their potential and scalability (Chandrashekar, 2018; Firth et al., 2017a, 2017b; Naslund et al., 2017). However, approximately 20'000 mental health apps currently available for download (Clay et al., 2021; Lagan et al., 2021; Torous & Roberts, 2017) are often untested, unreviewed and unsupported (Leigh & Flatt, 2015). Furthermore, most available apps on the market provide similar primary functions such as symptom trackers, mindfulness exercises or brain trainings (Neary & Schueller, 2018), suggesting the need to expand to new functionalities and features (Wisniewski et al., 2019).

With the current consumer hardware available, immersive technologies like virtual and augmented reality allow such an expansion to create engaging, realistic and high-quality experiences (Lindner, 2021; Riva, 2022). Virtual reality (VR) has been used in the clinical setting for decades and research has accumulated to support its effectiveness as a successful intervention for several disorders, e.g. anxiety and eating disorders as well as addiction, aggression or pain management (Riva, 2022). Although the clinical application of augmented reality (AR) is investigated to a smaller extend, the results for its' effectiveness in treating small animal phobia and substance abuse are promising, suggesting a potential as treatment for other disorders (Vinci et al., 2020a).

The intervention research for specific phobias such as fear of animals, objects and situations has been of particular interest to explore the potential of immersive technologies in exposure-based treatment approaches (Baus & Bouchard, 2014; Eaton et al., 2018). Exposure in vivo experiences a dissemination problem despite its high efficacy and recommendation by several treatment guidelines in cognitive behavioral therapy (CBT) (Pittig & Hoyer, 2017). Smartphone apps using immersive technologies have the potential to become a game changer in this matter.

Therefore, the goal of this dissertation is to shed light on the current standards, potential and challenges of using advanced immersive technology in clinical research. Three studies contribute to the field of research utilizing the currently available technology for VR and AR smartphone apps for the treatment of three common specific phobias in adults. Two original research papers are presented in regard to VR. Study **A)** addresses fear of heights, entitled *"Effectiveness of a stand-alone, smartphone-based virtual reality exposure app to reduce fear of heights in real-life: A randomized trial*". For this study my contribution covers the assistance in conceptualization of the VR training and the study design, collection of behavioral data and collaborative writing of the original manuscript. Study **B)** is devoted to fear of public speaking, a common form of social anxiety, yet similar to specific phobias in symptoms. My contribution for this study entitled *"Virtual reality-based gaze exposure treatment reduces fear of public speaking"* covers the assistance in conceptualization of the VR training and the study design, collection of behavioral data and collaborative writing of the original manuscript. The third presented original research paper **C)** showcases the potential of AR to reduce fear of spiders, entitled *"Effectiveness of a smartphone-based, augmented reality exposure app to reduce fear of spiders in real-life: A randomized controlled trial"*. My contribution for this study covers the conceptualization of the AR training, collection and curation of the behavioral data, their formal analysis and writing of the original manuscript.

2 Mental health apps

Smartphone apps represent an innovative way to increase the availability of mental health services by overcoming barriers to traditional treatments such as cost, access, and stigma (Neary & Schueller, 2018). They offer an attractive platform to deliver mental health information and interventions to people in need for support and daily life data to clinicians and researchers (Bhugra et al., 2017). Psychotherapists and patients are open and interested to use apps to assist self-management of mental health (Proudfoot, 2013; Torous et al., 2014). Also, industry and technology have recognized the potential of apps delivering support on mental health and well-being. The approximately 20,000 mental health apps currently available for immediate download from the app stores (Clay et al., 2021; Lagan et al., 2021) speak to their easy availability, as well as to the high interest (Torous & Roberts, 2017).

Some mental health apps are designed to be used stand-alone without any further support or guidance. This can increase the user's health literacy (Farrer et al., 2012; Taylor-Rodgers & Batterham, 2014), self-management skills and self-efficacy (Christensen et al., 2006). Other apps can be combined with face-to-face therapy in a blended care setting (Wilhelm et al., 2020) or assist clinicians in their diagnosis and treatment suggestions (Labrique et al., 2013; Torous & Powell, 2015). It has been shown, that app-based interventions in blended care are more impactful in reducing symptoms than stand-alone (Huckvale et al., 2020). However, waiting-lists for psychotherapy are long, not meeting the need for immediate support. For example, in Germany, 28-63 % of people with various mental health disorders remain untreated (Mack et al., 2014; Wang et al., 2007). Therefore, stand-alone apps could bridge the long waiting-lists and additionally, have the potential to reach non-treatment-seeking patient groups and could reduce stigmatization (Eichenberg, 2011; Musiat & Tarrier, 2014; Neary & Schueller, 2018).

Even so, mental health apps still have a few hurdles to overcome before they can be widely integrated into clinical practice. Their real-world usage is typically not sustained over time and there is little research regarding which features increase engagement (Wu et al., 2021). Further, concerns about data security, privacy and interoperability for psychotherapy settings remain (Schueller et al., 2016). The regulations concerning mental health apps are challenging with different health care systems and laws worldwide (Vincent et al., 2015). Platforms evaluating apps or guidelines regarding their development and integration differ in their focus on e.g. treatment outcome, usability or market value, making it difficult to find the appropriate evaluation tool (Lagan et al., 2020, 2021). Further, there is a lack of evidence from app-specific studies with randomized controlled trials investigating their intended effects, e.g. reduction of specific symptoms or improved self-management. Existing studies generally show low quality in study design and differences in methodic design, which makes it difficult to compare them in order to evaluate and rank the investigated apps (Tønning et al., 2019). With the time-lag for the translation of evidence-based research into health care practice (Kristensen et al., 2016), current research is not always reflected in the apps available in app stores. Furthermore, it has been described that there is no correlation between app ratings as well as popularity and the presence of evidence-based features (Kertz et al., 2017). Additionally, many publicly available mental health apps in app stores use techniques with little to no evidence (Larsen et al., 2019). Although the research around mental health apps in general may expand the quality assurance of mental health treatments, the efficacy of the apps remains contested (Baños et al., 2022; East & Havard, 2015). Thus, the app marketplace demonstrates a high availability but a low evidence base (Leigh & Flatt, 2015), meaning that consumers often use unreviewed, unsupported apps for mental health care.

Available apps offer a variety of functionalities, emphasizing on the potential to be used for self-help or in combination with face-to-face therapy. The most prominent functionalities target self-management, cognition improvement, skills-training, social support, symptom tracking or passive data collection addressing and amplifying all stages of clinical care (NIMH, 2017; Nahum-Shani et al., 2018). In response to active data collection with self-reports e.g. regarding mood, sleep, hallucinations, medication intake, substance use or social interactions, apps can deliver personalized, adaptive and immediate crisis interventions, preventions, diagnoses or treatments. The same can be done for passive data collection, also called *digital phenotyping* (Insel, 2017). For that, behavioral data is collected through smartphone sensors and device interactions, e.g. global positioning system (GPS), physical movement, text or scrolling activity, voice and speech analysis. This poses a promising field especially for prevention and personalized intervention (Bhugra et al., 2017; Huckvale et al., 2020; Price et al., 2014).

In sum, smartphone apps with psychological information and interventions increase the availability and accessibility of mental health support. However, the large amount of mental health apps on the market are rarely backed up by scientific evidence, emphasizing the need for further clinical investigation of their indicated effectiveness. Additionally, new functionalities and features are needed to exploit their full potential to address the variety of mental health disorders (Wisniewski et al., 2019; Wu et al., 2021).

3 Immersive technologies

The development and maturation of data processing speed and camera quality in smartphones have led to the possibility of displaying immersive technologies such as virtual and augmented reality (Lindner, 2021). These technologies have also been described as mixed, digital or hybrid realities, where the user can dive into the simulation of an alternative environment of visuals and sounds, hence being immersed. Research is accumulating support for the potential of VR and AR in mental health and attracted the interest of investors and the general public (Luckerson, 2014; Castelvecchi, 2016) as well as the scientific community for clinical research (Riva, 2022).

VR generates an immersive experience in a digital environment that replaces the user's real-world environment with either computer-generated images (CGI) or 360° pictures and videos. This is typically achieved through wearing a head-mounted display (HMD) with integrated headphones and speakers, and continuously tracking the users head rotation and movement to interact with the environment (Scarfe & Glennerster, 2019). Modern VR HMDs allow navigation and interaction in the VR environment through sensors and trackers following the head and body movement, using joysticks or controllers and also gaze control through built-in eye-tracking (Park et al., 2019; Riva, 2022). The release of consumer VR platforms such as Oculus, HTC Vive, PlayStation VR, Google Daydream or Samsung Gear VR led to a paradigm shift in the capabilities and scalability of VR for mental health. Consumer VR was quickly adapted for the use in clinical research and mobile VR, also with low-cost cardboard solutions offered a unique platform for home trainings (Lindner, 2021; Riva, 2022). Additionally, the development of VR in general with increasing display resolution and reactivity to movement have reduced the side effect known as cybersickness, similar to motion sickness with symptoms of dizziness, nausea, headaches, disorientation and oculomotor symptoms (LaViola, 2000). Although consumer VR platforms became relatively popular, mass adoption is missing. Big companies stopped working on VR, arguing with high costs and low display solution leading to a drop for the development of games for VR in the gaming market (Korolov, 2014; Ebert, 2015; Castelvecchi, 2016).

Augmented Reality (AR) is a variant of VR. While VR substitutes the existing physical environment with a virtual one, AR uses virtual elements to project into the existing environment in real-time (Azuma, 1997, 2001). Headsets developed like the Google Glass or Microsoft HoloLens are capable of overlaying information and integrating virtual objects into the physical world, creating a form of mixed reality (MR) (Lungu et al., 2021). However, the development of smartphones benefits and drives the evolution, use and dissemination of AR, as it works without a headset but via a smartphone or a tablet (Vinci et al., 2020a). A wave of interest and media attention for this technology occurred with the mobile game Pokémon Go, which was publicly played by people around the globe, attracting over 20 million daily users in the United States of America only a week after its release (Allan, 2016). It demonstrated the ability to integrate virtual objects into our real environment and aroused interest in researchers and clinicians for mobile AR (Baranowski & Lyons, 2020; de Souza e Silva, 2017; Vinci et al., 2020a).

The currently available consumer hardware of headsets and smartphones therefore offers advanced methods to further evolve mental health treatments with VR and AR. Additionally, the COVID-19 pandemic fostered not only the requirement of digital solutions for communication and work environments

but also for mental health support leading to another spark of interest in virtual platforms (Lindner, 2021; Riva, 2022; Roth et al., 2021).

3.1 User experience in immersive technologies

The advanced technologies available for VR and AR as well as their gaming and engaging character allow a high-quality user experience, which makes it a key feature to consider for both the consumer market as well as the clinical application (Lindner, 2021). Experiencing VR in comparison to AR is different, as VR substitutes the physical environment with a virtual one and AR projects virtual elements into the physical environment. Nevertheless, the instruments to qualify a user's experience are the same in regard to three intertwined concepts, namely the feeling of presence, the level of realism and the degree of reality (Baus & Bouchard, 2014).

Presence refers to the feeling of *being there* and is a key concept in virtual environments (Slater, 2018). It is proposed to ensure that the experiences and responses generated in the virtual environment are similar to those in the real world (Sanchez-Vives & Slater, 2005). The level of realism describes the perceived overlap between the expectations of the user and the actual experience (Baños et al., 2000). Reality refers to the immersion and response to the stimuli being authentic. Thus, a higher level of realism should be associated with a higher level of reality (Baños et al., 2000) and is also a necessary component of presence (Schubert et al., 2001).

Most studies have utilized VR stimuli consisting of CGI. However, CGI has inherently low representational quality and realism in VR environments is crucial to elicit greater levels of emotions and natural behavioral responses (Alsina-Jurnet et al., 2011; Slater et al., 2009). Moreover, the use of CGI avatars as a representation of humans in VR may increase the risk of the *uncanny valley effect* (Mori, 1970). This effect describes the discomfort in viewers when exposed to artificial faces of avatars with a strong, yet imperfect resemblance to humans, which are being rated as unbelievable and uncanny (Burleigh et al., 2013). Nowadays, the smartphones used in combination with consumer VR platforms offer the possibility to display 360° images and videos for the virtual environment. Preliminary findings indicate that virtual environments with 360° material are rated as more realistic than CGI environments (Melo et al., 2018; Tarnawski, 2017).

Even with the technological advances, it is still a challenge to create the VR environments as realistic as possible while still guaranteeing real-time interactivity. Here, AR offers new possibilities, eliciting a greater degree of realism by letting users see their own body and allowing a different immersive experience, since the user can interact with the virtual stimuli while remaining in the real world (Baus & Bouchard, 2014; Botella et al., 2005; Juan et al., 2005).

As real-world usage of mental health apps declines over time and user engagement is low (Baumel et al., 2019), thus assuring a good user experience proves to be crucial. A recent meta-analysis demonstrates that implementing gamification features, such as points, badges, levels and avatars could increase compliance and engagement (Wu et al., 2021). Immersive technologies are well suited for gamification elements which could increase engagement and may even enhance the apps intended

effects (Cheng et al., 2019) as apps with a greater number of engagement features have larger clinical effects (Wu et al., 2021).

3.2 Clinical application of immersive technologies

The field of immersive technologies has shown how developments in technology have fueled clinical progress (Lindner, 2021). Virtual environments have been utilized to simulate complex situations for psychoeducational and interventional purposes in a variety of disorders, such as phobias, post-traumatic stress disorder (PTSD), eating and weight disorders, obsessive-compulsive disorder (OCD), psychosis, substance abuse disorders, social skills training as well as aggression or pain management with rapidly being expanded to other disorders (Lindner, 2021; Riva, 2022).

VR has been used in CBT since the 1990s, accumulating an impressive evidence base showing clinically relevant effect sizes in reducing symptoms and severity of several disorders (Freeman et al., 2017; Lindner, 2021). AR has only recently been applied for therapeutic purposes and already the potential has been recognized (Cipresso et al., 2018; Vinci et al., 2020a). Although the dissemination of AR and VR varies for the treatment of different disorders, both immersive technologies have been used and investigated for decades to treat specific phobias, which will be the focus of the next chapters and also the field of research where our studies contribute with innovative and effective solutions.

4 Specific phobias

Anxiety disorders rank among the most frequent mental disorders (Jacobi et al., 2014; Kessler et al., 2012). Specific phobias in particular are one of the most common ones with an estimated lifetime prevalence ranging from 3% to 15% (Eaton et al., 2018). They present themselves as irrational fears about situations, creatures, places or objects, categorized in fear of animals (e.g. spiders, insects, dogs), blood-injections-injuries (e.g. taking blood, dentists), natural environments (e.g. heights, storms, water), situations (e.g. flying, elevators, enclosed spaces) and other (e.g. vomiting, suffocating, buttons, costumed people) (APA, 2013). Specific phobias can be diagnosed with only a few questions asked by a clinician or survey with one criterion being that the individuals recognize their phobias as unreasonable. Self-reported measures are thus a valid assessment method (Eaton et al., 2018). However, not all affected fulfill the criteria for a clinical diagnosis. From the general population 50% to 70% report having at least one unreasonably strong fear, indicating a much higher prevalence of these fears than their consequent diagnoses (Curtis et al., 1998; Eaton et al., 2018; Stinson et al., 2007).

When exposed to the feared stimuli, affected individuals experience an immediate emotional and physiological reaction with symptoms such as intense fear or disgust, panic, accelerated heart rate and sweating. This often results in a strong avoidance behavior of the feared stimuli to reduce stress and impairment (APA, 2013). Even though individuals with specific phobias realize the irrationality and even absurdity of their fears, the symptoms as well as the avoidance behavior are difficult to control, hindering them from doing a variety of daily life activities. This results in functional impairment, a negative impact on interpersonal interactions and quality of life in general (Bandelow & Michaelis, 2015).

The occurrence of specific phobias is more prevalent in Europe and Northern America than in Asia, Africa and Latin America (APA, 2013), with women being overall more affected, especially by animal phobias (Eaton et al., 2018; Schienle, 2021). There are several theories for their origin, which can differ from person to person. Additionally, the evolutionary point of view sees fear (and disgust) as a necessary emotion for protection against harm and danger, leading to a natural avoidance of certain objects, creatures or situations. Phobias can be caused by one or several traumatic events with the feared stimuli resulting in a conditioned fear response through a learning association between the stimuli and the fear reaction. Most phobias show an early onset with an average age between 7 and 10 years (Schienle, 2021; APA 2013) and can therefore also be explained by the model of behavioral learning, indicating the fear was learned by observing other people, e.g. parents (Eaton et al., 2018; Schienle, 2021).

The most common specific phobias concern either heights or animals (Stinson et al., 2007), closely followed by fear of social situations (Michael et al., 2007). Fear of heights shows a lifetime prevalence of 20-30% in the general population with 5% meeting the criteria for a clinical diagnosis (Depla et al., 2008; Eaton et al., 2018; Huppert et al., 2013). The triggers are widespread and not only limited to natural environments such as mountains or cliffs, but also bridges, externally attached stairs or apartments, balconies and offices in high buildings (Davey et al., 1997). Fear of small animals is one of the most prevalent forms of specific phobias, especially in women. Numbers between studies highly vary, the lifetime prevalence is overall estimated to 5-12% (Becker et al., 2007; Curtis et al., 1998; Depla et al., 2008; Fredrikson et al., 1996). Spiders especially trigger a strong fear in many people, highly intertwined with the feeling of disgust (Davey, 2011; Polák et al., 2020). Among one in 10 individuals will meet diagnostic criteria of social anxiety at some point in their lives (Kessler et al., 2012). A frequently reported situation where people experience a form of social anxiety also without a clinical diagnosis is public speaking anxiety (PSA) with prevalence estimates of 6.5% to 30% (Ruscio et al., 2008; Pull, 2012; Seim & Spates, 2019). PSA although belonging to the spectrum of social anxiety disorders (SAD) shows similar symptoms to a specific phobia (APA, 2013). In addition, PSA is associated with fear of negative evaluation and self-perception resulting in little participation in or avoidance or oral communication activities and speeches. Social interactions cannot be completely avoided and speeches and presentations are especially common in schools, universities and work places, leading to a variety of safety behaviors and coping strategies in affected individuals. Avoiding eye contact is a prominent safety behavior in socially anxious and even healthy adults when giving public speeches. Therefore, it has been postulated to be an important risk and maintenance factor in PSA (Chen et al., 2020; Moukheiber et al., 2010; Schulze et al., 2013).

Experiencing symptoms of specific phobias present a burden and an impairment in certain situations in daily life. Avoidance of the feared stimuli and applying safety behaviors can lead to an acute reduction of stress and impairment. However, if not treated, specific phobias can also take a chronic form, leading to a permanent presence of symptoms. Moreover, people affected by specific phobias have a higher risk for other anxiety disorders, secondary depression or substance abuse (APA, 2013).

4.1 Exposure therapy

There are effective treatment approaches to reduce the phobic feelings and increase quality of life for those affected. Exposure in vivo is currently the gold standard for the treatment of specific phobias (APA, 2013), describing a method which consists of confrontation with the feared stimuli until distress has decreased. Many studies have demonstrated that exposure-based treatments are among the most effective treatments for reducing fear and avoidance in specific phobias (Choy et al., 2007; Wolitzky-Taylor et al., 2008). Exposure therapy can be performed in gradual stages through a *fear hierarchy* starting from the least fearful or with *flooding* starting with the most fearful stimulus or as *systematic desensitization*, which includes relaxation techniques (APA-SCP, 2017). The exposure process can be stretched over several sessions. Additionally, there are protocols for intensive one-session exposure therapies for several specific phobias, which last three to four hours and show highly effective results (Davis et al., 2012). The general aim is to create new learning experiences to help those affected manage their fear. Although manualized protocols exist, the treatment needs to be tailored to each individual in order to obtain positive long-term results (Bandelow et al. 2014.; Schienle, 2021; Davis et al., 2012).

The fear reduction accomplished with exposure therapy can be explained through multiple psychological mechanisms. Habituation models propose that repeatedly entering and remaining in fear-activating situations results in a reduced fear response (Benito & Walther, 2015; Groves & Thompson, 1970). The emotional processing theory states that confrontation with fearful situations provides additional information, restructuring the cognitive representations of the feared stimuli as being harmful or dangerous (Foa & Kozak, 1986). The model of extinction learning – also called inhibitory learning – builds on the theory of fear conditioning. The fear reaction to the stimuli learned during conditioning is not eradicated through repeated exposure, but a secondary inhibitory association is developed, resulting in a decline in the conditioned fear response (Craske et al., 2014). Studies have also shown that the main predictors of treatment outcome in exposure-based therapies are the diminution of dysfunctional beliefs and the increase of self-efficacy (Côté & Bouchard, 2009; Tardif et al., 2019). For a significant reduction of fear, each of these mechanisms rely on situations optimally activating an individual's fear-network (Choy et al., 2007; Wolitzky-Taylor et al., 2008).

Although the success of exposure-based treatments is supported by several studies and recommended in most international CBT guidelines for the treatment of specific phobias, exposure in vivo is rarely used in therapeutic settings due to several limitations, leaving many affected individuals untreated. Reasons most commonly entail issues regarding the practicability with unpredictable time management, location scouting or getting hold of the specific fear stimuli such as animals or an audience for social situation and negative beliefs, such as concerns about the superficial effectiveness of exposure and that the procedure might be too stressful for the patients (Cook et al., 2010; Pittig et al., 2019). Additionally, exposure can also be challenging for therapists, especially for trainees and the relatively unexperienced (Broicher et al., 2017). Even though the focus on training therapists in routine care of specific phobias has been associated with more frequent self-reported utilization of exposure (Broicher et al., 2017; Pittig & Hoyer, 2017) exposure is rarely utilized in non-university settings (Becker et al., 2004; Böhm et al., 2008; Cook et al., 2010) and psychotherapists often have reservations to conduct exposure.

Consequently, a substantial number of therapists do not incorporate exposure into their practice in the long run (Becker et al., 2004; Böhm et al., 2008). In addition, surveys over the last few decades show that approximately 60-80% of people suffering from specific phobias do not seek professional help (Boyd et al., 1990; Essau et al., 2000; Magee et al., 1996; Agras et al., 1969). They might have adapted their daily lives to their fear or are not willing to expose themselves to the feared stimuli voluntarily (Bandelow & Michaelis, 2015). Approximately 25% of the people who do seek treatment, refuse to do exposure therapy after learning what it entails (Garcia-Palacios et al., 2001, 2007). Further, there is a high drop-out rate of up to 45% for in vivo exposure treatment of adults due to low acceptance in the patients (Choy et al., 2007). In addition, return of fear is a common risk even after successful exposure interventions (Côté & Bouchard, 2009; Mystkowski et al., 2002; Tardif et al., 2019).

These various barriers illustrate the need for novel exposure-based treatment approaches that circumvent the raised limitations of conventional in vivo treatment, targeting individual, practical and systemic barriers.

4.2 Virtual reality exposure therapy for specific phobias

Many limitations of in vivo exposure can be addressed or minimized by virtual reality exposure therapy (VRET). VRET aims to transfer the patient's exposure to the feared stimuli into a virtual environment, enabling certain advantages over exposure in vivo. Conducting the exposure treatment in VR leads to a reduction of the amount of logistic planning e.g. the duration of the exposure within a psychotherapy session as well as choosing or scouting the appropriate location and stimuli, which comes in hand with reduced cost. Additionally, VRET is more practicable for the psychotherapists by offering multiple, repeatable and adaptable scenarios as well as high controllability of the environment and stimuli (Maples-Keller et al., 2017; Neudeck & Einsle, 2012). Moreover, those affected by specific phobias and fears have rated VRET as a more preferable treatment option than in vivo exposure (Garcia-Palacios et al., 2007), possibly lowering the hurdle to seek treatment.

Recent research has shown that VR exposure can also be packaged as automated interventions relying on gamification components rather than a real-life therapist directing treatment (Donker et al., 2019; Freeman et al., 2018; Lindner et al., 2020; Miloff et al., 2015) or incorporating a virtual therapist for guidance (Lambe et al., 2020; Miloff et al., 2020). These interventions have the potential to make a significant impact by addressing the large treatment gap (Alonso et al., 2018) and the delay in treatment-seeking for anxiety disorders such as specific phobias (Wang et al., 2020). There is clear evidence for the effectiveness of therapy guided VRET in treating specific phobias (e.g., fear of heights, spiders or flying) and social anxiety disorder with comparable effect sizes to those of traditional CBT (Botella et al., 2017; Chou et al., 2021; Dellazizzo et al., 2020). Recent RCTs show promising results with significant symptom reductions in fear of heights and spiders as well as PSA with automated and gamified VRET (Donker et al., 2019; Freeman et al., 2018; Lindner et al., 2020, 2021b; Miloff et al., 2019). However, many studies on automated VRET do not investigate the effectiveness in a real-life setting, even if the transfer of the effects of VRET to real-life situations have been shown in the past (Morina et al., 2015). Additionally, standardization for VR is currently lacking, both in terms of development as well as in

treatment protocols. This makes direct comparisons between studies difficult, since different labs may have used different hardware, software or therapeutic procedures. Moreover, the majority of research on the efficacy of VRET have been conducted in an academic context. As a result, the effectiveness of VR psychotherapies in clinical practice is still unclear, which contributes to psychotherapists' reluctance in adopting such novel treatments in their routine practice (Becker & Jensen-Doss, 2013).

Surveys have shown that clinicians and psychotherapists have a generally favorable view of using VR in therapy, although they also reported fears about required training, handling the equipment, financial costs, and reported an overall low degree of acquaintance with the technology (Schwartzman et al., 2012; Segal et al., 2011). With the upcoming of consumer VR, these fears have decreased, although clinical use of VR is still rare and knowledge of VRET is still low (Lindner et al. 2019). The integration of VR into devices such as smartphones further increase the availability and accessibility of VR in everyday life. All the while the cost of equipment is becoming more affordable. All these factors could benefit the development and implementation of VR psychotherapies (Ma et al., 2021) presenting a paradigm shift also for clinical and public health applications of VR (Lindner, 2021).

In sum, VRET offers an effective and feasible alternative to exposure in vivo. However, even with affordable consumer VR available the developed interventions have yet to be integrated into clinical routine, which indicates further research on their practicability and engagement. In our studies we investigated two VR smartphones apps for the highly prevalent specific phobia of fear of heights and the very common fear of public speaking. Study **A)** for fear of heights aimed at reducing fear and avoidance and additionally investigating the usability of the developed app *EasyHeights*. The investigation of the developed app called *Fearless Speech* in study **B)** is devoted to the reduction of PSA as well as the increase in eye contact when exposed to an audience. Both interventions offer fully automated and gamified exposure procedures with no further cognitive elements. They are designed as stand-alone home-trainings, without the need for additional guidance or support. Integrated gaze control is used for an easy navigation within the virtual environment and in study **B)** even as an intervention to increase eye contact. The studies additionally set themselves apart from previous studies on automated VRET by two features: The virtual environments are created by 360° images and videos instead of CGI to increase realism and in both studies the effectiveness of the VR apps is investigated in real-life settings in addition to subjective measures.

4.2.1 Effectiveness of a stand-alone, smartphone-based virtual reality exposure app to reduce fear of heights in real-life: a randomized trial

npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

Check for updates

Effectiveness of a stand-alone, smartphone-based virtual reality exposure app to reduce fear of heights in real-life: a randomized trial

Dorothée Bentz[1,2], Nan Wang[1,2], Merle K. Ibach[1,2], Nathalie S. Schicktanz[1,2], Anja Zimmer[1,2], Andreas Papassotiropoulos[2,3,4,5] and Dominique J. F. de Quervain[1,2,5]✉

Smartphone-based virtual reality (VR) applications (apps) might help to counter low utilization rates of available treatments for fear of heights. Demonstration of effectiveness in real-life situations of such apps is crucial, but lacking so far. Objective of this study was to develop a stand-alone, smartphone-based VR exposure app—*Easy Heights*—and to test its effectiveness in a real-life situation. We performed a single-blind, parallel group, randomized controlled trial. We recruited 70 participants with fear of heights, aged 18–60 years. Primary outcome was performance in a real-life Behavioral Avoidance Test (BAT) on a lookout tower after a single 1-h app use (phase 1) and after additional repeated (6 × 30 min) app use at home (phase 2). After phase 2, but not phase 1, participants in the *Easy Heights* condition showed significantly higher BAT scores compared to participants in the control condition (Cohen's $d = 1.3$, $p = 0.0001$). Repeated use of our stand-alone, smartphone-based VR exposure app reduces avoidance behavior and fear, providing a low-threshold treatment for fear of heights.

npj Digital Medicine (2021)4:16 ; https://doi.org/10.1038/s41746-021-00387-7

INTRODUCTION

Fear of heights is a common problem with a lifetime prevalence of around 20–30% and with around 5% of the general population meeting diagnostic criteria of the American and international classification for specific phobia (natural-environmental type: heights)[1–5]. For those affected, exposure to height situations sets off strong emotional and physiological reactions, such as intense fear or panic and accelerated heart rate, often resulting in avoidance of specific height triggers[1,4]. Compared to other specific phobias, these triggers are widespread, including not only unusual encounters as high mountains or cliffs, but also daily life encounters, such as stairs, terraces, bridges, apartments, and offices located in high buildings[6]. This might lead to a profound impact on daily life and can result in functional impairment for the sufferers. Affected people report a considerable impact on interpersonal interactions and quality of life in general[7].

Nonetheless, treatment seeking and uptake in clinical practice is still limited[3], despite existing treatment options with high success rates for acute symptom improvement in up to 80% of patients. The current gold-standard to treat fear of heights is in vivo exposure therapy, where patients expose themselves with the feared stimuli[8,9]. The lack of dissemination seems to be partly rooted in the core element of exposure treatment, namely 'exposure to the feared stimulus'. Not all individuals with specific fears are willing to expose themselves to the feared stimuli voluntarily[10] or patients drop out of exposure treatment due to low acceptance of in vivo exposure[8]. Additionally, psychotherapists often have reservations to conduct exposure[11,12]. Reasons include liability issues and concerns that exposure might be too stressful for their patients[12]. Moreover, exposure sessions are often time intensive in preparation (e.g., selecting appropriate height triggers) and conduction (e.g., travelling to reach height triggers). Therefore, new modes of delivery for exposure that circumvent the raised issues are warranted.

The implementation of virtual reality (VR), where individuals can expose themselves to the feared stimuli in VR, has the potential to counter many of the raised problems of in vivo exposure. Triggering stimuli can be simulated in the therapy room of a therapist at any time, which first reduces the preparation time for exposure and second makes its use more flexible and for example not dependent on time of day or weather conditions[13].

Furthermore, patients have the possibility to expose themselves in the comfort of the therapy room and there is evidence for a higher willingness to expose themselves with virtual than real triggers[14,15]. Therapists also see potential in the VR technology and state various advantages as for example heightened accessibility and control over fearful triggers[16]. Since the first published study to treat fear of heights with VR in 1995[17], evidence accumulated in favor of exposure in VR to treat specific phobias with large effect sizes compared to control conditions and a comparable efficacy of exposure in VR to in vivo exposure[18]. Despite its good efficacy including transfer to real-life situations[19] and high acceptability within patient and therapist populations, exposure in VR is still mostly restricted to laboratories and experimental studies[13]. Fear of potential technical difficulties and monetary expenses for the VR equipment and software might be reasons that only a minority of therapists offer VR treatment[16,20].

Smartphone-based VR relying on a portable VR headset and a conventional smartphone might be the solution for the current dissemination problem of exposure in VR. Smartphone-based VR is highly accessible due to low costs of necessary VR headsets and the widespread use of smartphones in the general population. Furthermore, digital marketplaces are already in place to enable dissemination of VR exposure apps to practitioners or as self-help

[1]Division of Cognitive Neuroscience, Faculty of Psychology, University of Basel, Basel, Switzerland. [2]Transfaculty Research Platform, University of Basel, Basel, Switzerland. [3]Division of Molecular Neuroscience, Faculty of Psychology, University of Basel, Basel, Switzerland. [4]Life Sciences Training Facility, Department Biozentrum, University of Basel, Basel, Switzerland. [5]University Psychiatric Clinics, University of Basel, Basel, Switzerland. ✉email: dominique.dequervain@unibas.ch

Published in partnership with Seoul National University Bundang Hospital

npj nature partner journals

tool directly to sufferers. Smartphone-based exposure has all the benefits mentioned for stationary VR and, in addition, it can be conducted both in the therapy office without cost-expensive VR gear and as stand-alone add-on in form of home-work in between-sessions (blended treatment)[21]. Three studies (two of them with smartphone-based interventions) are in favor of the idea that stand-alone applications (apps) with exposure elements are beneficial to reduce fear of heights in sufferers[22–24]. However, in these studies, fear of heights was only assessed by self-reported measures (i.e., fear of heights questionnaires), but not in real-life situations. Based on a meta-analysis showing that VR effects on subjective fear as assessed by questionnaires generally translate to real-life situations[19], one could speculate that the interventions of the published stand-alone studies might also lead to fear reduction in real-life situations. However, to convince sufferers with fear of heights of the effectiveness of VR treatment, demonstration of fear reduction in real-life situations is crucial.

Our study is a randomized controlled trial that sets itself apart from the published smartphone-based interventions in the following aspects: (1) We measure avoidance behavior and subjective fear in a real-life height situation. (2) Our approach is solely based on exposure without cognitive elements. (3) Our study includes both individuals with either subclinical or clinical (DSM-5) fear of heights. Our primary outcome is performance in a Behavioral Avoidance Test (BAT) in a real-life situation, which is considered an objective measure of fear[25]. Based on the published studies using VR exposure treatment[23,24], we expected large treatment effects in the BAT and also in the secondary outcome measures, such as subjective fear on the tower and fear of heights questionnaires, in our smartphone-based VR intervention condition directly after a 1-h session and after an additional prolonged home-treatment (6 × 30 min) assessed at 3–5 weeks after app use as compared to the control condition.

RESULTS

Participant's characteristics (study phase 1)

One hundred and six individuals were screened for trial participation, of whom 29 were excluded after screening (Fig. 1). Consequently, 77 individuals were enrolled and underwent randomization, of whom 39 were allocated to use the *Easy Heights* app (intervention condition) and 38 were allocated to the control condition. Seventy participants (42 fulfilling DSM-5 criteria for specific phobia) completed study phase 1 as planned and were analyzed. Participant's baseline characteristics were balanced across conditions (Table 1). In study phase 1, two participants (one in the intervention condition, one in the control condition) dropped out due to VR side effects.

Effects of acute use of *Easy Heights* (study phase 1)

The duration for *Easy Heights* app use for all participants of study phase 1 was 60 min in total. Uptake of the VR heights exposure was high in study phase 1 (100%).

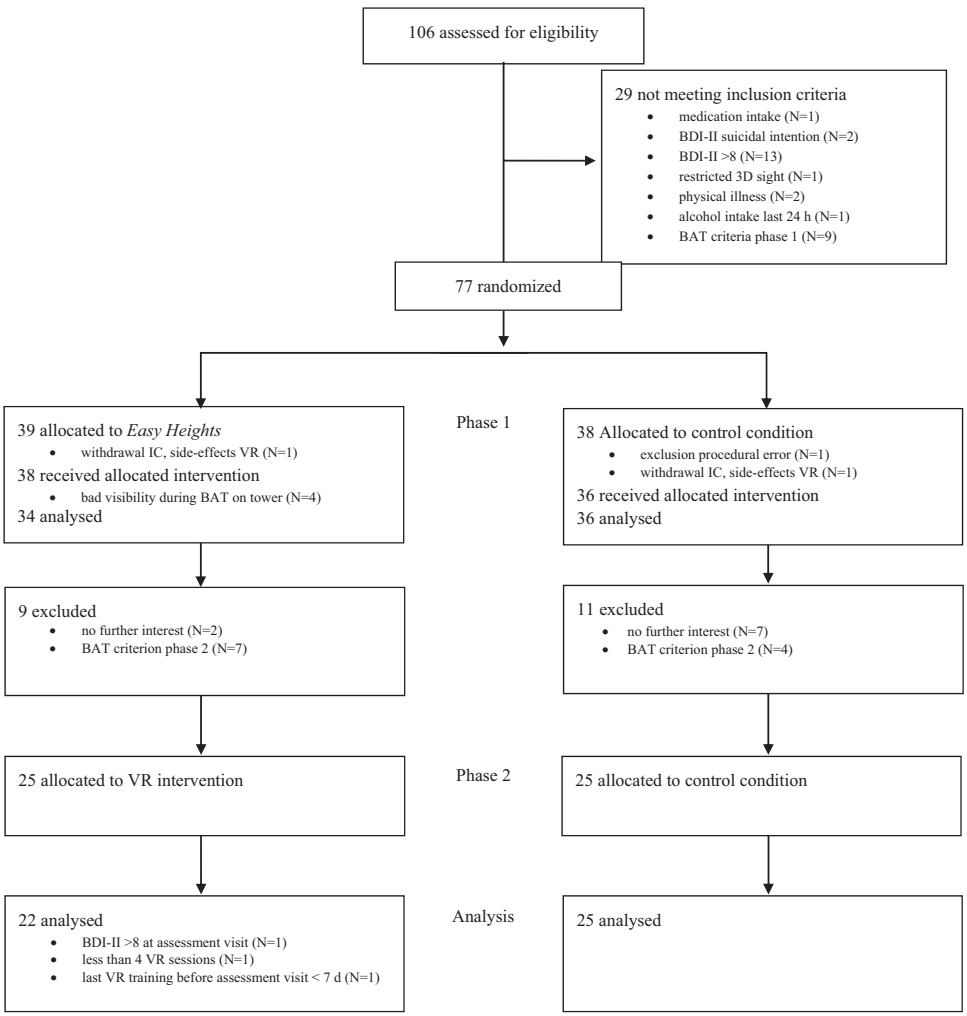

Fig. 1 Flowchart of participants—CONSORT. BDI-II Beck Depression Inventory, BAT Behavioural Avoidance Test, VR Virtual Reality, IC Informed Consent.

Table 1. Baseline characteristics.

	Easy Heights		Control condition	
	Phase 1 ($n = 34$)	Phase 2 ($n = 22$)	Phase 1 ($n = 36$)	Phase 2 ($n = 25$)
Age (years)	30.2 (9.8)	31.6 (11.4)	32.8 (11.3)	31.8 (11.1)
Women	15 (44%)	10 (45%)	19 (53%)	15 (60%)
Education				
Master, master equivalent or higher degree	9 (27%)	5 (23%)	11 (31%)	8 (32%)
Bachelor degree	10 (29%)	6 (27%)	10 (28%)	6 (24%)
Vocational education	4 (12%)	2 (9%)	3 (8%)	2 (8%)
High school education	11 (32%)	9 (41%)	11 (31%)	9 (36%)
Obligatory schooling	0 (0%)	0 (0%)	1 (3%)	0 (0%)
DSM-5 diagnosis	21 (62%)	15 (68%)	21 (58%)	16 (64%)

Data are numbers of participants or mean (SD/percentages).

Table 2 summarizes the mean scores of the primary and secondary outcomes at post VR heights exposure in study phase 1 (acute use) with respective baseline calculated with values of all participants of study phase 1, outcome data were missing from one participant for our secondary outcomes AQ, DES, AES, ATHQ, and mean subjective fear on the tower during the BAT.

The *Easy Heights* app users compared with the control condition did not show significantly higher BAT scores immediately after acute VR heights-exposure ($F_{(1,64)} = 0.74$, $p = 0.392$, Cohen's d = 0.21). For two secondary outcomes, the acute use of *Easy Heights* showed beneficial effects: AQ ($F_{(1,63)} = 9.86$, $p = 0.003$, Cohen's d = 0.77) and self-reported change of fear of heights ($F_{(1,65)} = 8.46$, $p = 0.005$, Cohen's d = 0.71). These beneficial effects on fear questionnaires in study phase 1 were independent of sex, age, diagnosis, and baseline values (no significant interactions between sex and condition, age and condition, diagnosis and condition, baseline values and condition: all $p > 0.08$). No other significant two-way interactions or main effects of condition on secondary outcomes were detected (all two-way interactions $p > 0.014$; all main effects of condition $p > 0.063$).

Participant's characteristics (study phase 2)

Of the 70 participants from study phase 1, 59 were eligible to take part in study phase 2 (Fig. 1). Ninety-six percent completed the full intervention course of study phase 2 with at least 4 VR exposure trainings with a minimum training duration of 20 min. Uptake of the VR heights exposure was high in study phase 2 (93%).

Effects of repeated use of *Easy Heights* (study phase 2)

After the additional home-treatment (mean total *Easy Heights* app use in minutes: 170.59, SD 35.49) in study phase 2, spanning on average over 15.00 days (SD 5.50), the use of the VR *Easy Heights* app showed a beneficial effect on our primary outcome BAT score. The intervention condition showed higher BAT scores compared to the control condition 29.91 days [SD 13.20] after the last use of the *Easy Heights* app ($F_{(1,41)} = 18.45$, $p = 0.0001$, Cohen's d = 1.28, see Fig. 2a). The intention to treat analysis based on 76 participants (39 participants in the control condition and 37 participants in the intervention condition) confirmed the results with higher BAT scores at assessment visit after repeated use of the *Easy Heights* app in study phase 2 in the intervention compared to the control condition ($F_{(1,70)} = 5.23$, $p = 0.025$, Cohen's d = 0.53, $p_{perm} = 0.022$).

Further, the intervention condition indicated less mean subjective fear on the tower during the BAT ($F_{(1,41)} = 18.13$, $p = 0.0001$, Cohen's d = 1.27, see Fig. 2b), as well as significantly higher self-reported change of fear of heights ($F_{(1,41)} = 19.08$, $p = 0.00008$, Cohen's d = 1.32), and less fear of heights in our questionnaires AQ ($F_{(1,40)} = 16.96$, $p = 0.0002$, Cohen's d = 1.25, see Fig. 2c), and ATHQ ($F_{(1,40)} = 17.45$, $p = 0.0002$, Cohen's d = 1.26, see Fig. 2d). The intervention effects were independent of sex, age, diagnosis, and baseline variable (interactions between sex and condition, age and condition, diagnosis and condition, baseline values and condition: all $p > 0.047$), with the exception of the AES questionnaire (for further information about the AES and deltas for all primary and secondary outcomes see Supplementary Methods 1). No significant two-way interactions or main effects of condition on DES were detected (all $p > 0.115$).

Table 2 summarizes the mean scores of the primary and secondary outcomes at post VR heights exposure in study phase 2 after repeated use, with respective baseline calculated with values of solely of participants of study phase 2, outcome data was missing from one participant of for our secondary outcomes AQ, DES, AES, and ATHQ).

DISCUSSION

We showed that our stand-alone, smartphone-based virtual reality exposure app *Easy Heights* is highly effective in the reduction of avoidance behavior and subjective fear in a real-life height situation after repeated use. Furthermore, we found a reduction of fear of heights in self-report measures already after a single 1-h session with the app. Intervention uptake in the *Easy Heights* condition was high in study phase 1 as well as the continuation rate in study phase 2, indicating that the app was well accepted. We assessed symptoms of simulation sickness in VR and found them to be slightly higher in our *Easy Heights* condition compared to the control condition. Nevertheless, with only 15% of the maximal score of the simulation sickness questionnaire[26], they were still very low and due to the overlap between common side effects of VR (simulation sickness) and fear symptoms the score of the simulation sickness questionnaire in the exposure situation is likely to be confounded by fear symptoms.

Findings of a meta-analysis on in vivo treatments of specific phobia indicate an effect size of d = 1.1[9] and a meta-analysis on VR treatments of specific phobia found a comparable effect size[18]. With our stand-alone, smartphone-based VR exposure app *Easy Heights* we found an effect size of d = 1.3 for the repeated use. In this sense it compares well with the current gold-standard to treat fear of heights, the in vivo exposure therapy, and with the stationary therapist-guided VR exposure. It is also in line with other stand-alone VR apps reporting large effect sizes as assessed by questionnaires[23,24]. The strength of our study is that we showed the benefits of our intervention in a real-life height situation on the behavioral as well as the subjective level. We

Table 2. Outcome measures and differences between conditions.

	Easy Heights	*n*	Control condition	*n*	Adjusted group difference (95% CI)[a]	Effect size (Cohen's d)	*p* value
BAT score (primary outcome)							
Phase 1							
Baseline	11.1 (7.8)	34	9.6 (6.8)	36	—	—	—
Post intervention	15.2 (8.4)	34	12.7 (7.8)	36	0.9 (−1.2 to 2.9)	0.2	0.392
Phase 2							
Baseline	8.6 (5.3)	22	7.2 (3.0)	25	—	—	—
Post intervention	14.4 (8.1)	22	7.1 (2.9)	25	6.7 (3.6 to 9.9)	1.3	0.0001
Mean subjective fear on the tower during the BAT (secondary outcome)							
Phase 1							
Baseline	4.1 (1.9)	34	4.1 (1.7)	35	—	—	—
Post intervention	2.3 (1.9)	34	2.8 (1.9)	35	−0.5 (−1.1 to 0.1)	0.4	0.081
Phase 2							
Baseline	3.8 (2.0)	22	3.8 (2.0)	25	—	—	—
Post intervention	1.9 (1.7)	22	3.5 (1.8)	25	−1.6 (−2.4 to −0.9)	1.3	0.0001[b]
AQ anxiety subscale (secondary outcome)							
Phase 1							
Baseline	47.1 (19.1)	33	49.5 (17.4)	36	—	—	—
Post intervention	31.8 (19.0)	34	44.5 (22.3)	36	−10.5 (−17.2 to −3.8)	0.8	0.003[b]
Phase 2							
Baseline	49.4 (19.8)	21	54.4 (14.4)	25	—	—	—
Post intervention	34.0 (15.1)	22	52.3 (17.3)	25	−15.6 (−23.2 to −7.9)	1.2	0.0002[b]
ATHQ total (secondary outcome)							
Phase 1							
Baseline	41.7 (7.8)	33	41.2 (8.9)	36	—	—	—
Post intervention	35.8 (10.7)	34	38.2 (12.4)	36	−2.5 (−6.1 to 1.1)	0.3	0.175
Phase 2							
Baseline	41.3 (7.6)	21	42.6 (8.5)	25	—	—	—
Post intervention	35.1 (8.2)	22	42.2 (6.9)	25	−7.3 (−10.9 to −3.8)	1.3	0.0002[b]
AES total (secondary outcome)							
Phase 1							
Baseline	33.7 (5.2)	33	34.8 (5.7)	36	—	—	—
Post intervention	27.8 (6.7)	34	30.7 (8.0)	36	−2.6 (−5.5 to 0.4)	0.4	0.086
Phase 2[c]							
Baseline							
Females	36.1 (4.5)	9	37.5 (3.5)	15	—	—	—
Males	32.0 (5.4)	12	33.0 (7.0)	10	—	—	—
Subclinical	29.7 (4.9)	7	32.6 (7.1)	9	—	—	—
Clinical	35.7 (4.5)	14	37.4 (3.5)	16	—	—	—
Post intervention							
Females	29.2 (9.1)	10	36.9 (5.4)	15	−7.4 (−12.8 to −2.0)	0.4	0.010
Males	29.7 (7.6)	12	31.9 (5.9)	10	−2.0 (−6.6 to 2.6)	1.3	0.374
Subclinical	27.9 (9.9)	7	31.6 (7.4)	9	−4.2 (−11.4 to 3.0)	0.7	0.224
Clinical	30.2 (7.4)	15	36.8 (4.3)	16	−5.6 (−9.5 to −1.7)	1.1	0.007[b]
DES total (secondary outcome)							
Phase 1							
Baseline	17.7 (4.2)	33	17.5 (5.1)	36	—	—	—
Post intervention	14.7 (4.7)	34	15.9 (5.8)	36	−1.5 (−3.2 to 0.1)	0.5	0.063
Phase 2							
Baseline	17.1 (4.4)	21	18.9 (4.5)	25	—	—	—
Post intervention	15.0 (4.8)	22	16.6 (4.6)	25	−0.7 (−3.1 to 1.7)	0.2	0.564
Self-reported change of fear of heights (secondary outcome)							
Phase 1							
Post intervention	64.4 (12.7)	34	54.3 (13.5)	36	9.2 (2.9 to 15.5)	0.7	0.005[b]

Table 2 continued

	Easy Heights	*n*	Control condition	*n*	Adjusted group difference (95% CI)[a]	Effect size (Cohen's d)	*p* value
Phase 2							
Post intervention	65.9 (14.0)	21	51.5 (8.3)	25	14.6 (7.8 to 21.3)	1.3	<0.0001[b]
SSQ total (safety outcome)							
Phase 1							
Baseline	3.5 (3.7)	34	3.5 (3.4)	35	—	—	—
Post intervention	7.2 (4.2)	34	5.0 (4.2)	36	2.4 (0.5 to 4.3)	0.5	0.013

Data are mean (SD), unless otherwise indicated. Phase 1 = after 1-h *Easy Heights* vs. 1-h virtual reality (VR) control intervention, Phase 2 = after 1-h and 6 × 30 min *Easy Heights* vs. 1-h VR control intervention and no further intervention.
BAT Behavioral Avoidance Test, *AQ* Acrophobia Questionnaire, *ATHQ* Attitudes Toward Heights Questionnaire, *AES* Anxiety Expectancy Scale, *DES* Danger Expectancy Scale, *SSQ* simulation sickness questionnaire.
[a]Adjusted for condition, diagnosis, sex, age and baseline measure (BAT score, Mean subjective fear on the tower during the BAT, AQ anxiety subscale, ATHQ total, AES total, DES total). The difference was assessed by linear models.
[b]Significant after Bonferroni correction (significance threshold $p < 0.008$ for secondary outcomes).
[c]Means (SD) are displayed separately for sex and diagnosis, because of significant interaction between condition and sex as well condition and diagnosis.

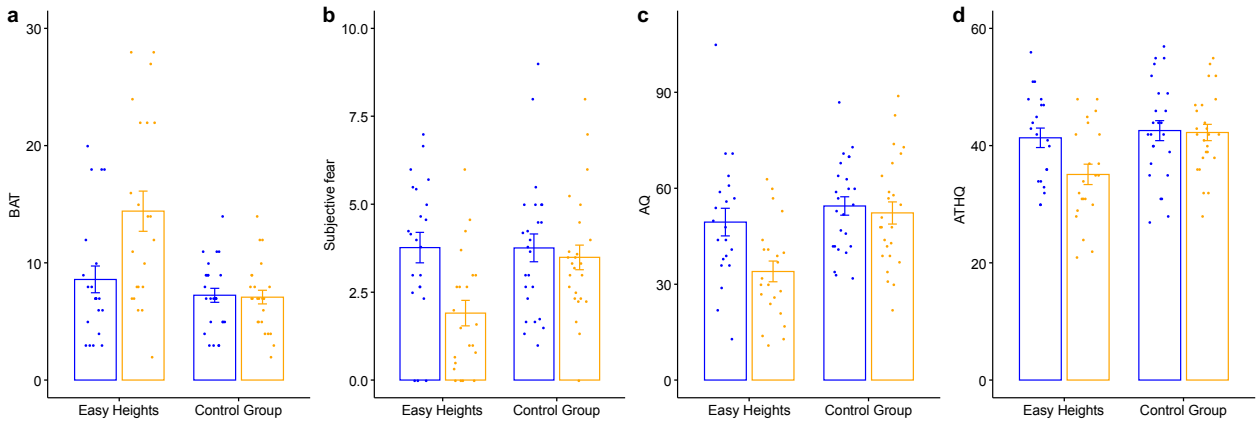

Fig. 2 Primary (Behavioral Avoidance Test, BAT) and secondary outcome measures (mean subjective fear on the tower during BAT, Acrophobia Questionnaire, AQ, Attitudes Towards Heights Questionnaire, ATHQ at baseline (represented with blue bars) and at phase 2 post intervention (represented with yellow bars). a Behavioral Avoidance Test (BAT): The range of the BAT score is 0–28 (1 point was given for reaching each platform and 1 point for looking down on each platform for 10 seconds). **b** Mean subjective fear on the tower during BAT: Mean subjective fear was calculated from the fear levels assessed on the reached platforms after looking down for 10 seconds. The range of the score is 0–10 (0 = no fear to 10 = maximum fear). **c** Acrophobia Questionnaire (AQ): The range of the AQ score is 0–120 with higher scores indicating higher severity. **d** Attitudes Towards Heights Questionnaire (ATHQ): The range of the ATHQ is 0–60 with higher scores indicating a more negative attitude towards heights. Means and standard errors are displayed.

implemented an approach that was solely based on exposure with no cognitive elements and included clinically diagnosed (DSM-5) as well as subclinical individuals with fear of heights.

Our trial has several limitations. First, we recruited specifically for a smartphone-based intervention to treat fear of heights that might have led to a selection bias of participants willing to use modern technologies for treatment purposes. Therefore, we do not know how representative our study population is for the general population. Second, study participation was only possible for the German speaking population of Switzerland or neighboring Germany. Consequently, our app was solely tested on this specific population with fear of heights. Nevertheless, we suppose that the broad dissemination of smartphones worldwide, the resulting familiarity with mobile technologies in combination with the easy handling of the setup that we observed during the study conduction (especially during the home-training without assistance from the study team) are in favor of the generalizability of our results to other populations with fear of heights. Additionally, our *Easy Heights* app will be adapted to current VR systems and made available at no costs in the English language. Third, we only assessed fear of heights around 3–5 weeks after the last app use and not at a later timepoint. Fourth, our intervention duration and regime of first 1-h in study phase 1 and later 6 × 30 min was predefined and not compared to other intervention regimes. Therefore, we have no information about optimal dose-response relationship of our *Easy Heights* app. And last, we have no experience on how well our results translate to clinical practice or how well our *Easy Heights* app will be accepted in the general population. We can only extrapolate from the feedback of our participants that acceptability was high, but the treatment uptake as a stand-alone intervention (downloadable app) or integration in a blended treatment has to be further scrutinized.

To conclude, our results indicate that the repeated use of a smartphone-based, stand-alone virtual reality exposure app leads to large improvements in avoidance behavior and subjective fear of heights both in a clinical and subclinical population. Low costs of the necessary setup and easy accessibility of the app qualify it as a useful addition to the current mental health care services as well as a self-help option for people with subclinical fear of heights.

METHODS

Study design and participants

We performed a single-blind, parallel group, randomized controlled trial comparing a smartphone-based VR height exposure app with a VR condition without height exposure in study phase 1 and with no intervention in study phase 2. For trial participation we recruited physically healthy participants with clinical and subclinical fear of heights between age 18–60 years from the German speaking general population of Switzerland by print, radio and online advertisements. Participants were enrolled in the study between October 16, 2018 and November 26, 2018. We included individuals with fear of heights (subclinical: criteria A-E and G but not F (distress/impairment), clinical: A–G of the DSM-5[1] criteria for specific phobia, natural-environmental type: heights). We excluded individuals if they were not fluent in German, received concurrent psycho- or pharmacotherapy, were ever in treatment for fear of heights or participated in another study, showed signs of at least mild depression (Beck Depression Inventory II, BDI-II[27] total score > 8) or suicidal ideation (BDI-II item 9 > 0), had physical illnesses, restricted 3D sight or chronic medication intake (except intake of oral contraceptives) and females if they were pregnant. Participants were instructed to abstain from alcohol and medication intake for 12 h and psychoactive substances (including benzodiazepines) for 5 days before study days. Furthermore, to counteract a possible ceiling effect in study phase 1, we excluded people who have reached the highest possible platform of the tower and have given a fear rating of 6 or smaller on a scale between 0 and 10 during our baseline Behavioral Avoidance Test (BAT). For study phase 2 we excluded all participants who have reached the highest platform during our post VR intervention BAT (BAT score 27 or higher) in study phase 1, since there was no further improvement possible. The study protocol including the definition of primary and secondary outcomes and statistical analysis plan was approved by the Ethic Committee of North-West and Central Switzerland (EKNZ) before start of the study (October 16 2018). On October 20 2018 the enrolment criteria concerning BAT performance were updated and an interim analysis was included in the protocol. As the study protocol was first set out to only study the acute use of the *Easy Heights* app, we had to add study phase 2 to the study protocol to investigate the repeated use. The adapted protocol version was approved by the EKNZ before start of study phase 2 (February 17, 2019) (for further information see Supplementary Methods 2). As the Swiss law (Ordinance on Clinical Trials in Human Research) foresees the possibility of retrospective registering to prevent that the registration of a trial (along with the disclosed information) interferes with later patent filing, this option has been chosen per default. The trial was registered on ClinicalTrials.gov on July 1, 2019. However, the protocol (including primary and secondary outcome measures and statistical analysis plan) has been predefined and accepted by the official ethics committee (https://www.eknz.ch) before the start of the study phases. Final data was collected on May 24, 2019.

All research has been performed in accordance with the Declaration of Helsinki. All participants gave written informed consent for trial participation. Participants received a compensation of CHF 150 for their participation in study phase 1 and CHF 300 in study phase 2.

Randomization and masking

After study inclusion, participants were randomly (stratified for the presence of a DSM-5 diagnosis of fear of heights and sex) allocated to the two treatment conditions (intervention condition: VR heights exposure app vs. control condition: fear-unrelated VR tasks in phase 1, and no treatment in phase 2, Fig. 1). Each eligible participant was allocated to one of the four randomization lists (two lists for participants with subclinical fear of heights (male/female) and two for clinical fear of heights (male/female)). The first author of the manuscript prepared the randomizations lists by means of random number tables. In these randomization lists treatment conditions were block-randomized in blocks of four. Every block of four included two times the allocation to each condition (intervention/control condition).

Allocation concealment was given for the experimenter who enrolled participants, as treatment allocation was made by a different experimenter. Therefore, the experimenter who enrolled participants did not know in advance which treatment the next person gets. The experimenter who collected our primary and secondary outcome measures in the real-life height situation was unaware of the group assignment of participants (single-blind).

Procedures

After a potential participant contacted the study team, more detailed information about the study along with the main inclusion and exclusion criteria was sent by email. People who showed an exclusion criterion during the subsequent online screening carried out via SoSci Survey[28] were directly informed that they are not eligible for participation. Eligible participants were contacted and scheduled for the study. Study phase 1 took part in the facilities and on the lookout tower of the Uto Kulm AG on the Uetliberg near Zurich, Switzerland. The Uto Kulm AG provided the minimum technological infrastructure necessary for study conduction and use of the *Easy Heights* app namely electricity to charge the smartphones and headphones. Before study enrolment, a study team member checked all inclusion and exclusion criteria and collected basic demographic data, including a baseline BAT. Subsequently, participants filled out questionnaires to collect baseline measures for their fear of heights. Afterwards another study team member allocated the participants to one of the two treatment conditions by filling in one of the four randomization lists (depending on the presence of a DSM-5 diagnosis of fear of heights and sex). Then participants filled out the Simulator Sickness Questionnaire (SSQ)[29] before starting the VR intervention accompanied by another study team member (for detailed information on implemented questionnaires and tests see Supplementary Methods 1 and information provided under outcomes below).

For VR height exposure, we used *Easy Heights*, a stand-alone smartphone-based VR height exposure app (for in-app content see Fig. 3a, b). The app was designed to be used without further assistance or accompanying therapist, but it can also be integrated in a blended treatment approach in a clinical context. The content of the VR exposure app is based on a graduated behavioral exposure approach and includes no psycho-educative elements and specific cognitive interventions (as e.g., challenging of cognitive distortions). According to the German evidence-based guideline for the treatment of specific phobias exposure in vivo is the treatment of choice for specific phobias, exposure in VR is evaluated as second best option if in vivo is not available or possible[30]. Once the user opens the *Easy Heights* app all the information on how to use the app is given in written in 2D. The information starts with a short description of the app content, it is explained that the 3D part of the *Easy Heights* app consists of three different scenarios (rural mountain, cloudy weather, urban town) in which the user is standing on a virtual platform (starting on the ground level). The VR scenarios are based on 360° panoramic photos taken by a drone at different heights and accompanied by sounds characteristic to each VR scenario and level (e.g., sound of birds at lower levels of the rural mountain scenario and wind sounds at higher levels were played in). In each of the three scenarios 16 different height levels are available (corresponding to a range of heights between 0 and 75 m). Users proceed from ground level to further levels according to a predefined exposure scheme based on Subjective Units of Distress Scales (SUDS, "How big is your fear at this level?", scale 0 = no fear to 10 = maximum fear) (Fig. 3a). Users have to stay at each level until their SUDS are 3 or below for two consecutive ratings. After completing one level the users are reinforced with a yellow balloon for each level they completed (one balloon up to 15 balloons with completing the last level of each scenario) as well sound effects accompanying the movement of the virtual platform upwards (gamified reinforcement elements). SUDS are assessed continuously throughout the three exposure sessions. The first rating is prompted after 10 seconds at each level followed by at least two more SUDS in each situation. SUDS are given by the user via gaze selection. Each exposure session is terminated by the time limit of 20 min, irrespective of achieved level, for study phase 1 and 30 min for study phase 2.

During study conduction achieved levels, SUDS, date and time of *Easy Heights* app use were stored locally for later analysis. The stored data were deidentified and only the experimenter was able to link the data on a smartphone to a specific participant. Each smartphone was numbered and the allocation to a specific participant was recorded before handing it out to the participant. Data were only collected for study purposes and there will be no data collection in the *Easy Heights* app that will be made publicly available later on.

For the VR height exposure intervention participants were given Samsung smartphones with a preinstalled *Easy Heights* app, noise cancelling headphones and a Google Daydream View version 2 VR headset (Fig. 3c shows a member of our division wearing the setup) to enable 100-degrees stereoscopic view and a controller for the headset (for further information of the VR setup see Supplementary Methods 3).

For a smooth integration of the *Easy Heights* app use into our study course an experimenter gave some assistance for the use of the material

Fig. 3 Virtual Reality exposure app and real-life testing. a In-App content of the *Easy Heights* mountain scenario on the ground level and **b** an advanced level from the perspective of the user. **c** Equipment of the study worn by a team member (written informed consent for reprint is given): an android smartphone with a preinstalled *Easy Heights* app, noise cancelling headphones and a Google Daydream View version 2 VR headset. **d** The lookout tower where our Behavioural Avoidance Test (BAT) was conducted.

and some verbal instructions that are also given in written within the app. Participants were assisted with putting on the portable VR headset and were instructed to stand still on a chosen spot and only move their upper part of the body and their head while in VR. Between each 20 min exposure sessions participants had a 5 min break and were offered something to drink. Participants of the control condition received the same devices. Their task was to use the Google Street View app that was preinstalled on the smartphones given to them and to explore three predefined virtual scenarios (Iglou visitor center, Versailles, cubic houses) in VR[31]. The three chosen Google Street View scenarios were selected, because they did not include any height stimuli. Participants of the control condition were not prompted to give SUDS and were allowed to explore each scenario at their own pace by teleporting themselves with the controller.

After completion of all three VR sessions, participants filled in a second SSQ. Afterwards they completed a second BAT on the Uetliberg lookout tower (Fig. 3d) and filled out the same questionnaires on their fear of heights and indicated on a scale self-reported change of fear of heights their subjective improvement after app use. Additionally, they filled out a questionnaire on presence in VR and a VR app acceptability and usability scale. At the end of study phase 1, participants were assessed for adverse events and sent home, if no safety concerns were present.

Participants of study phase 1 that did not climb the highest platform during our post VR intervention BAT were offered to take part in study phase 2. Study phase 2 comprised for the intervention condition an additional home-treatment spanning over two weeks concluding with an assessment visit at the Uetliberg 3–5 weeks after cessation of the home-training and for the control condition no further intervention and only an assessment visit at the Uetliberg. Eligible participants that showed interest in participation were reassessed with a second online screening for main inclusion and exclusion criteria concerning their health status and gave written informed consent to take part in study phase 2. Afterwards participants of the intervention condition received via mail a Samsung smartphone with a preinstalled *Easy Heights* app, as well as standard accessory charger and headphones and a headset with controller.

Participants were instructed to use the *Easy Heights* app six times within 14 days. Participants were allowed to train on any day they wanted with the restriction to train only once a day. The sequence of the scenarios was predetermined (2× rural mountain, 2× cloudy weather, and 2× urban town). Each scenario lasted for 30 min. Participants were instructed to stand still on a chosen spot and only move their upper part of the body and their head while in VR.

At the assessment visit in study phase 2 at the Uetliberg Uto Kulm facilities, we first checked for alcohol, medication and psychoactive substances intake, and depressive symptomatology as well as suicidal ideation. Afterwards, we conducted another BAT similar to the first two BATs in study phase 1 and participants again filled out the same questionnaires about their fear of heights as well as the scale on self-reported change of fear of heights.

Outcomes

Our primary outcome was performance in the real-life BAT on a lookout tower with 14 platforms. During the BAT, participants were instructed to walk up the Uetliberg Tower as far as their current fear allowed and to look down to ground level on each platform for 10 seconds (for further information about procedure of the BAT see Supplementary Methods 1). The BAT score ranged between 0 and 28 (1 point was given per platform reached and 1 point for looking down on each platform for 10 seconds).

Our secondary outcomes were mean subjective fear on the tower during the BAT as calculated from the fear levels (indicated by participants after looking down on each platform for 10 seconds based on SUDS, Subjective Units of Distress Scale) assessed on the reached platforms during the BAT (for further information about the calculation of mean subjective fear on the tower during the BAT or the other secondary outcomes see Supplementary Methods 1) (range 0–10 with higher scores indicating higher subjective fear), the Acrophobia Questionnaire (AQ) (range of 0–120 with higher scores indicating higher severity)[25], the Attitudes Towards Heights Questionnaire (ATHQ) (range of 0–60 with higher scores indicating a more negative attitude)[32], the Anxiety and Danger Expectancy scales (AES/DES) (range of 10–50 for AES and range of 5–25 for DES with higher scores indicating higher severity)[33], and self-reported change of fear of heights measured by a single visual analogue scale (range 0–100, 0 = a lot worse, 50 = no change and 100 = a lot better).

Primary and secondary outcomes were evaluated before the first (baseline) and directly after the single VR intervention in study phase 1, and on the assessment visit, scheduled 3–5 weeks after the last use of the *Easy Heights* app during study phase 2.

The Simulator Sickness Questionnaire (SSQ) was implemented to assess side effects of the VR exposure (range 0–48 with higher scores indicating higher severity)[29].

Statistical analyses

We applied a per-protocol analysis, our data were analyzed with R studio version 3.6.2[34] and validated by a second statistician.

We applied linear models (nlme-package)[35] in combination with ANOVA (SS II). Study phase 1 and 2 were analyzed with separate linear models.

Dependent variables were our primary outcome BAT score and our secondary outcomes mean subjective fear on the tower during the BAT, AQ, ATHQ, AES, DES, and self-reported change of fear of heights. Independent variable was the between-subject factor condition (intervention or control). Baseline measures of our primary (BAT score) and secondary outcomes (mean subjective fear on the tower during the BAT, AQ, ATHQ, AES, DES) from study phase 1 were included as covariate. Further, sex, age, and diagnosis (clinical/subclinical) were entered as covariates/cofactors. Covariates/cofactors were included as main effects, and as two-way interactions with condition. In case of no significant interactions between covariates/cofactors and the factor condition on dependent variables, the two-way interactions were removed from the statistical model.

Furthermore, to account for loss of participants between study phase 1 and study phase 2, we conducted an intention to treat (ITT) analysis for our primary outcome BAT score using the same linear model specified above. For missing outcomes, we applied the method Last Observation Carried Forward (LOCF) including the last available value of every subject that was reliably assessed. Group assignment was maintained according to randomization for participants entering the ITT[36,37].

We present results as mean (SD) for the intervention and control condition, and associated two-sided p values, as well as adjusted group difference with 95% CIs (emmeans-package). Due to our six secondary outcomes, we set the significance threshold to $p < 0.008$ (Bonferroni correction for six independent tests) for the secondary outcomes. We estimated Cohen's d as effect size measurement. The estimate of d was based on t values of the linear models. Therefore, d is corrected for the effects of all confounding variables included in the linear model. By convention, $d = 0.2$ is considered to be a small, $d = 0.5$ to be an intermediate and $d = 0.8$ to be a large effect[38]. According to previous VR exposure studies to treat fear of heights we expect large effect sizes[39]. The estimation of $N = 80$ is based on a power analysis using an ANCOVA with fixed effects assuming to detect a large effect size ($f = 0.5$) with a power of 80% at $\alpha = 0.05$ (software: G-power 3).

No data monitoring committee oversaw the study. A clinical trial monitor oversaw data collection and entry according to a written monitoring plan approved by the IEC before trial conduction. The trial is registered at ClinicalTrials.gov with the Identifier: NCT04003753.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

DATA AVAILABILITY

Deidentified data generated during and/or analyzed for the current study are available from the corresponding author on reasonable request.

CODE AVAILABILITY

The *Easy Heights* software code is available from the corresponding author upon reasonable request for academic purposes. The *Easy Heights* app will be made publicly available at no costs in the English language.

Received: 24 August 2020; Accepted: 6 January 2021;
Published online: 08 February 2021

REFERENCES

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders* 5th edn. (APA Press, Washington, 2013).
2. Depla, M., Have, M., Balkom, A. & Graaf, R. Specific fears and phobias in the general population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Soc. Psychiatry Psychiatr. Epidemiol.* **43**, 200–208 (2008).
3. Huppert, D., Grill, E. & Brandt, T. Down on heights? One in three has visual height intolerance. *J. Neurol.* **260**, 597–604 (2013).
4. World Health Organization. *The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research* (WHO, Geneva, 1993).
5. Eaton, W. W., Bienvenu, O. J. & Miloyan, B. Specific phobias. *Lancet Psychiatry* **5**, 678–686 (2018).
6. Menzies, R. G. in *Phobias: A Handbook of Theory, Research and Treatment* (ed. Davey, G. C.), pp. 129-138 (Wiley, Chichester, 1997).
7. Schäffler, F. et al. Consequences of visual height intolerance for quality of life: a qualitative study. *Qual. Life Res.* **23**, 697–705 (2014).
8. Choy, Y., Fyer, A. J. & Lipsitz, J. D. Treatment of specific phobia in adults. *Clin. Psychol. Rev.* **27**, 266–286 (2007).
9. Wolitzky-Taylor, K., Horowitz, J., Powers, M. & Telch, M. Psychological approaches in the treatment of specific phobias: a meta-analysis. *Clin. Psychol. Rev.* **28**, 1021–1037 (2008).
10. Öst, L. One-session treatment for specific phobias. *Behav. Res Ther.* **27**, 1–7 (1989).
11. Cook, J., Biyanova, T., Elhai, J., Schnurr, P. & Coyne, J. What do psychotherapists really do in practice? An Internet study of over 2,000 practitioners. *Psychotherapy (Chic.)* **47**, 260 (2010).
12. Pittig, A., Kotter, R. & Hoyer, J. The struggle of behavioral therapists with exposure: self-reported practicability, negative beliefs, and therapist distress about exposure-based interventions. *Behav. Ther.* **50**, 355–363 (2018).
13. Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A. & Baños, R. Recent progress in virtual reality exposure therapy for phobias: a systematic review. *Curr. Psychiatry Rep.* **19**, 42 (2017).
14. Garcia-Palacios, A., Hoffman, H., See, S., Tsai, A. & Botella, C. Redefining therapeutic success with virtual reality exposure therapy. *Cyberpsychol. Behav.* **4**, 341–348 (2001).
15. Garcia-Palacios, A., Botella, C., Hoffman, H. & Fabregat, S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychol. Behav.* **10**, 722–724 (2007).
16. Segal, R., Bhatia, M. & Drapeau, M. Therapists' perception of benefits and costs of using virtual reality treatments. *Cyberpsychol. Behav. Soc. Netw.* **14**, 29–34 (2011).
17. Rothbaum, B., Hodges, L. & Kooper, R. Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *Am. J. Psychiatry* **152**, 626–628 (1995).
18. Carl, E. et al. Virtual reality exposure therapy for anxiety and related disorders: a meta-analysis of randomized controlled trials. *J. Anxiety Disord.* **61**, 27–36 (2018).
19. Morina, N., Ijntema, H., Meyerbröker, K. & Emmelkamp, P. Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behav. Res. Ther.* **74**, 18–24 (2015).
20. Schwartzman, D., Segal, R. & Drapeau, M. Perceptions of virtual reality among therapists who do not apply this technology in clinical practice. *Psychol. Serv.* **9**, 310–315 (2012).
21. Lindner, P. et al. Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: a cross-sectional survey study in the era of consumer VR platforms. *Front. Psychol.* **10**, 1–10 (2019).
22. Hong, Y., Kim, H., Jung, Y., Kyeong, S. & Kim, J. Usefulness of the mobile virtual reality self-training for overcoming a fear of heights. *Cyberpsychol. Behav. Soc. Netw.* **20**, 753–761 (2017).
23. Donker, T. et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. *JAMA Psychiatry* **76**, 682–690 (2019).
24. Freeman, D. et al. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. *Lancet Psychiatry* **5**, 625–632 (2018).
25. Cohen, D. C. Comparison of self-report and overt-behavioral procedures for assessing acrophobia. *Behav. Ther.* **8**, 17–23 (1977).
26. Bouchard, S., St-Jacques, J., Renaud, P. & Wiederhold, B. Side effects of immersions in virtual reality for people suffering from anxiety disorders. *J. Cyber Ther. Rehabil.* **2**, 127–137 (2009).
27. Beck, A. T., Steer, R. A., Ball, R. & Ranieri, W. Comparison of Beck Depression Inventories-IA and -II in psychiatric outpatients. *J. Pers. Assess.* **67**, 588–597 (1996).
28. Leiner, D. J. *SoSci Survey (Version 3.1.06) [Computer software]*. https://www.soscisurvey.de (2019).
29. Kennedy, R. S., Lane, N. E., Berbaum, K. S. & Lilienthal, M. G. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.* **3**, 203–220 (1993).
30. Bandelow, B. et al. *Deutsche S-3 Leitlinie zur Behandlung von Angststörungen* (Springer, 2014).
31. Google (n.d.). *Google Street View (Version 2.0.0.257517656). [Mobile application software]*. https://play.google.com/store/apps (2019).
32. Abelson, J. L. & Curtis, G. C. Cardiac and neuroendocrine responses to exposure therapy in height phobics: desynchrony within the 'physiological response system'. *Behav. Res. Ther.* **27**, 561–567 (1989).
33. Gursky, D. M. & Reiss, S. Identifying danger and anxiety expectancies as components of common fears. *J. Behav. Exp. Psychiatry* **18**, 317–324 (1987).
34. R Development Core Team. *R: a language and environment for statistical computing*. (R Foundation for Statistical Computing, Vienna, 2012).
35. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Core Team, R. nlme: linear and nonlinear mixed effects models. *R. Package Version* **3**, 1–143 (2019).

36. Streiner, D. & Geddes, J. Intention to treat analysis in clinical trials when there are missing data. *Evid. Based Ment. Health* **4**, 70–71 (2001).
37. Gupta, S. K. Intention-to-treat concept: a review. *Perspect. Clin. Res.* **2**, 109–112 (2011).
38. Cohen, J. A power primer. *Psychol. Bull.* **112**, 155–159 (1992).
39. de Quervain, D. J.-F. et al. Glucocorticoids enhance extinction-based psychotherapy. *Proc. Natl Acad. Sci. USA* **108**, 6621–6625 (2011).

ACKNOWLEDGEMENTS

We thank Amanda Aerni for monitoring the study. Thomas Schlitt for the data management of the study. Fabian Müller for designing sound effects for the app, Alexander de Quervain for 360-degree picture material for the app, Caroline Nicola Beer, Tamara Flury, Christiane Gerhards, Galya Iseli, Irena Kovacic, Cécile Longoni, Fabian Müller, Anna Lena Seitz, Saskia Schnydrig, Jerôme Schwarz for helping in trial conduction. The study has been funded by the Transfaculty Research Platform Molecular and Cognitive Neurosciences of the University of Basel.

AUTHOR CONTRIBUTIONS

D.B. and D.Q. designed the exposure app and the trial and drafted the paper. M.I. and N.W. programmed and visually designed the VR app. D.B. and A.Z. supervised data collection. A.P. commented on the design of the exposure app and trial. D.B. and N.S. analyzed the data. All authors commented on and gave final approval of the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-021-00387-7.

Correspondence and requests for materials should be addressed to D.J.F.d.Q.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

4.2.2 Virtual reality-based gaze exposure treatment reduces fear of public speaking

PNAS

Main Manuscript for

Virtual reality-based gaze exposure treatment reduces fear of public speaking

Fabian D Mueller[a,e], Berhard Fehlmann[a,e], Nan Wang[a,e], Merle K Ibach[a,e], Thomas Schlitt[a,e], Dorothée Bentz[a,e], Anja Zimmer[a,e], Andreas Papassotiropoulos[b,c,d,e], and Dominique JF de Quervain[a,d,e]

[a]Division of Cognitive Neuroscience, Department of Psychology, University of Basel, 4055 Basel, Switzerland; [b]Division of Molecular Neuroscience, Department of Psychology,University of Basel, 4055 Basel, Switzerland; [c]Life Sciences Training Facility, Department Biozentrum, University of Basel, 4056 Basel, Switzerland; [d]University Psychiatric Clinics, University of Basel, 4002 Basel, Switzerland; [e]Transfaculty Research Platform, University of Basel, 4055 Basel, Switzerland

* Prof. Dr. Dominique J.-F. de Quervain, University of Basel, Department of Psychology, Division of Cognitive Neuroscience, Birmannsgasse 8, CH-4055 Basel, Switzerland, Phone: +41 61 207 02 37

Email: dominique.dequervain@unibas.ch

Author Contributions: Fabian D. Mueller (FM), Bernhard Fehlmann (BF), Dominique De Quervain (DQ) and Andreas Papassotiropoulos (AP) conceived and designed the study. Nan Wang and Merle K Ibach programmed and visually designed the VR app. FM, BF and Anja Zimmer collected and analysed the data. Thomas Schlitt supervised data collection, storage and processing. Dorothée Benz provided clinical advice and contributed to trial design. FM, BF and DQ wrote the manuscript, with substantial input from the other authors. DQ and AP provided critical oversight and feedback of the work.

Competing Interest Statement: all authors declare no competing interests.

Classification: Biological Science; Psychological and Cognitive Sciences.

Keywords: Eye contact, gaze avoidance, public speaking anxiety, social anxiety, virtual reality exposure therapy

[1] FM and BF contributed equally to this work.

This PDF file includes:

Main Manuscript

- Figure 1: Schematic study procedure

- Figure 2. Flowchart of participants

- Figure 3. Effects of the gaze exposure treatment on perceived fear and relative dwell time on faces during public speaking

- Figure 4. Virtual reality-based gaze exposure application

- Figure 5. Example of face detection

- Table 1. Demographic characteristics of participants

- Table 2. Primary and main secondary outcome measures and differences between conditions

Supplementary Information

- Table S1. Further outcome measures and differences between conditions.

- Table S2. Level settings of the gaze exposure application.

Abstract

Public speaking anxiety (PSA) is the most widespread social fear, with prevalence estimates up to 30%. Whereas gaze avoidance is a common feature of PSA, it is not known if an exposure treatment aimed at reducing gaze avoidance would lead to a reduction of fear of public speaking. The objective of this study was to develop a stand-alone, virtual reality (VR)-based gaze exposure treatment and to test its effectiveness to reduce fear of public speaking. We performed a single-blind, parallel-group, randomized controlled trial in 89 healthy adults with PSA. The primary outcome was subjective fear in a real-life public speaking test. Secondary outcomes included the relative dwell time on faces during public speaking, as measured by eye-tracking. Assessments were done at baseline, after a single 1-h app use (phase 1), and after additional repeated (9 × 20 min) app use at home (phase 2). After phase 2, but not phase 1, participants in the treatment condition showed significantly reduced fear as well as increased relative dwell time on faces compared to participants in the control condition with no intervention (fear: Cohen's $d = 1.07$, $p < 0.0001$; relative dwell time on faces: Cohen's $d = 0.97$), $p < 0.0001$). The repeated use of a VR-based gaze exposure treatment intervention leads to large reductions in fear of public speaking and gaze avoidance in a real-life speech situation, suggesting effectiveness of fully automated treatment options with a focus on gaze behavior.

Significance Statement

Fear of public speaking is widespread and affected individuals typically avoid eye contact with the audience to avoid experiencing non-verbal signs of negative evaluation. The present research reveals that a two-week virtual reality (VR) gaze exposure treatment reduces fear of public speaking as well as gaze avoidance in a real-life public speech situation. The VR gaze exposure treatment is appealing for people with public speaking anxiety because it offers an easy-to-use and widely accessible self-help option that requires neither exposure to real people nor verbal interaction.

Introduction

Eye gaze serves a crucial role in regulating human social interaction (1, 2). Avoiding gaze in social situations indicates submission, fear, or refraining from further interaction across many species including humans (3–5). Notably, gaze aversion is also characteristic of several psychological disorders involving fear, anxiety, or attention modulation in social interactions, including social anxiety disorder, depression, and autism spectrum disorder (6–8). In particular, socially anxious people report fear of direct eye contact (9) and avoid mutual gaze when interacting with others (10, 11). Therefore, gaze behavior and attention processing of facial social cues (i.e., the eye region) are extensively discussed in social anxiety research, with the avoidance of mutual gaze in social situations thought to reflect a submissive gesture to reduce the anticipated social threat (12–15).

People with social anxiety typically avoid social situations due to fear of public scrutiny and evaluation (16). Avoiding mutual gaze in social situations may be an intuitive and unintentional anxiety-reduction technique, but may ultimately proof maladaptive (17). Although it may lead to a short-term reduction in anxiety because the experience of non-verbal signs of negative evaluation is reduced (18), gaze avoidance may contribute to the maintenance of anxiety in the long term by eliciting adverse reactions from others that nourish the existing fear (19) or by preventing anxious individuals from learning corrective information about the feared stimulus (20).

Gaze avoidance in social anxiety is well-documented (8, 21–24). Eye-tracking studies demonstrate that socially anxious people spend less time looking at others' eyes and make fewer fixations towards the eye region than do non-anxious individuals (22, 25). These findings are consistent across emotional expressions (25, 26) and types of social situations (8, 27–29), suggesting gaze avoidance as a bio-behavioral marker of social anxiety (13, 15). In public speaking situations in particular, both individuals with clinical and subclinical social anxiety exhibit higher rates of gaze avoidance than their non-anxious peers (21, 23, 30). In fact, public speaking anxiety (PSA) is the most widespread social fear, with prevalence estimates ranging between 6.5% to 30% (31–34).

Here we aimed at investigating if a gaze exposure treatment in adults with PSA is effective in reducing gaze avoidance and, more importantly, in reducing fear in a public speaking situation. We conducted a single-blind, parallel-group, randomized controlled trial. For the gaze exposure treatment, we developed a stand-alone smartphone application (app), implementing 360° panoramic videos of real audiences in VR. The gaze exposure focused on the maintenance of eye contact with virtual audiences across increasingly difficult social situations in VR. We tested the effectiveness of the gaze exposure treatment in a real-life public speaking test (PST) compared to a control condition. The primary outcome was the subjectively perceived fear during the PST. The main secondary outcomes were face gaze, measured as the relative dwell time on faces during public speaking, and speech quality assessed by the PST evaluation committee. Outcomes were tested at baseline, directly after a short (1-h app use) intervention, and after a prolonged home treatment intervention (9x 20 min) assessed 3-5 weeks after last use (Fig. 1).

Results

Participant's characteristics (study phase 1). Two hundred and twenty-six individuals were screened for trial participation (Fig. 2). The inclusion criteria were met by 89 individuals, of which 43 were randomly allocated to the treatment group and 46 to the control group. Study phase 1 was completed as planned by 86 participants (treatment group: 41, control group: 45). Participants' baseline characteristics were balanced across groups (Table 1).

Effects of 1-h use of VR treatment app (study phase 1). Uptake of the VR exposure to social threat was high in study phase 1 (100%). Table 2 summarizes the mean scores of the primary and main secondary outcomes at post VR gaze exposure treatment intervention in study phase 1 (1-h app use) with respective baseline calculated with values of all participants of study phase 1. Outcome data were missing for two participants in study phase 1 (1 of the control group and 1 of the treatment group) with regards to the global improvement and the usability of the app (see supplementary information; SI). Also, for two participants (1 of the control group and 1 of the treatment group), eye-tracking data was missing from study phase 1 due to a technical error. For the primary outcome measure, the subjectively perceived fear during the PST, one hour of VR gaze exposure with the training app showed no beneficial effect compared to the VR control intervention in study phase 1 ($F(2, 154) = 23.32, p = 0.2$, Cohen's $d = 0.27$). Similarly, for the main secondary outcomes, the relative dwell time on faces and the global performance, assessed by the committee, we observed no significant effects of the 1-h intervention compared to the control intervention in study phase 1 ($p \geq 0.05$).

Participant's characteristics (study phase 2). Of the 86 participants from study phase 1, 72 participants completed study phase 2 (treatment group: 28, control group: 44). In the treatment group, 65% correctly completed the full intervention course with at least six training sessions at home with a minimum training duration of 14 minutes. Uptake of the VR exposure to social threat was high in study phase 2 (90%).

Effects of repeated use of treatment app (study phase 2). After the additional home-treatment, the use of the treatment app (mean total app use in minutes: 168.84 [*SD 9.36*] spanning over an average of 11.96 days (*SD 2.30*) led to a significant reduction of subjective fear during the PST in study phase 2, 34.00 days [*SD 4.44*] after the last use of the app ($F(2, 154) = 23.32$, $p < 0.0001$, Cohen's $d = -1.07$; Fig. 3a and Table 2). The intention to treat (ITT) analysis based on 89 participants (46 participants in the control group and 43 participants in the treatment group) confirmed these results, with only marginally smaller effect size estimates ($p < .0001$; Cohen's $d = -1.04$; Table 2).

For the main secondary outcomes, the home treatment with the app had beneficial effects on the relative dwell time on faces but not on the global performance. The repeated use of the app led to an increase in the relative dwell time on faces during the PST in study phase 2 ($F(2, 151) = 8.71$, $p < 0.001$; Fig. 3b and Table 2). The ITT analysis confirmed these results, with slightly smaller effect size estimates at post-intervention phase 2 ($p < 0.001$; Cohen's $d = 0.83$; Table 2). For the global performance assessed by the committee, there was no effect of the home treatment ($F(2, 154) = 024$, $p = 0.79$). Mean scores of repeated training effects of all further outcome measures are summarized in the SI Table S1.

Discussion

While gaze avoidance is known to be a common feature of PSA (23, 30), it was not known if an intervention aimed at reducing gaze avoidance would be sufficient to reduce fear of public speaking. The present single-blind randomized controlled trial provides evidence that a VR-based gaze exposure treatment is effective in reducing fear and gaze avoidance in a real-life public speaking situation after repeated use. With a large effect size in fear reduction (d = -1.07) for the intervention group after only approximately four hours of training over two weeks, the effectiveness of the gaze exposure treatment is comparable to conventional exposure treatments that require patients to speak in front of other people in vivo (35) or in VR (36–38).

Interestingly, we observed a decrease in fear after the acute intervention in both the treatment and the control group. A possible explanation for this decrease might be a short-term habituation to the repeated PST on the same day, an effect that was obviously lost in the control group after several weeks. In the treatment group, the decrease in fear after the acute intervention was not greater than in the control group and not paralleled by a change in relative dwell time, suggesting that in the treatment group the decrease in fear was not the result of the single gaze exposure intervention and that specific effects only evolve after repeated interventions.

Importantly, we show that the repeated gaze exposure treatment aimed at increasing eye contact with a virtual audience is sufficient to reduce the fear of public speaking without requiring verbal interactions, added cognitive-behavioral elements, or input from a therapist. For people with PSA who find it already too frightening to verbally perform in front of a VR audience, this could further lower the threshold to overcome initial fears. Relying on mere exposure to gaze sets the current study apart from the conventional VRET approach, suggesting high effectiveness of behavioral treatment interventions with a focus on gaze behavior. Furthermore, we tested the effectiveness in a real-life public speaking situation, indicating that gaze exposure in VR generalizes to real life.

Our study has several limitations. First, we evaluated the effectiveness of the gaze exposure treatment intervention in a sample of adults with PSA but not with a full clinical presentation of social anxiety. Although the present findings may have potential for treating socially anxious individuals with a clinical diagnosis, the safety and effectiveness of this approach for clinical populations remain to be tested. Second, although the explicit task instructions during the gaze exposure treatment was to hold eye contact with the virtual audience members, the limitations of the used eye-tracking equipment and face detection algorithms did not allow us to accurately discriminate between gaze directed to the eyes versus other salient facial features (e.g., mouth, nose). However, the subjective perception of mutual gaze may be a product of mutual face gaze rather than mutual eye contact per se (39) and other facial features also provide essential social cues for interpreting social feedback (40). Finally, the intervention did not increase the global quality assessment of the speech performance as externally rated by the PST evaluation committee. Arguably, the addition of cognitive or interactive elements to the VR training sessions (e.g., instruction to perform speeches) could increase the training effects in terms of speech performance.

To conclude, our results indicate that the repeated use of a VR-based gaze exposure intervention leads to large reductions in the fear of public speaking as well as gaze avoidance in subsequent real-life speech situations, suggesting high effectiveness of fully automated treatment options with a focus on gaze behavior. Such stand-alone, widely accessible, and scalable training tools offer evidence-based solutions and alternatives for people with PSA at a low threshold for initiation, countering the dissemination problem of traditional in-vivo treatment. In the clinical setting, stand-alone apps could provide a valuable add-on to guided standard exposure therapy. Based on the promising results in a population with PSA, the full clinical potential of targeting gaze avoidance in social anxiety and other populations with constrained social gaze behavior remains to be investigated.

Materials and Methods

Study design and participants. We performed a single-blind, parallel-group, randomized controlled trial comparing a VR-based gaze exposure treatment app with a VR condition without social exposure in study phase 1 and with no intervention in study phase 2. For trial participation, we recruited from Switzerland's German-speaking population via print and online advertisements. Participants were enrolled in the study between June 18, 2019, and September 19, 2019. We included physically healthy individuals aged between 18-40 years that were fluent in German and indicated high fear and avoidance of social situations with the potential of being evaluated by others. We excluded individuals suffering from clinically relevant social anxiety, based on the criteria of the corresponding section of the structured Diagnostic Interview for Mental Disorders for DSM-5 (41). Further, individuals were not allowed to participate if they received concurrent psycho- or pharmacotherapy, were ever in treatment for anxiety disorders or participated in another study, showed signs of clinically relevant depressive symptoms (Beck Depression Inventory-II, BDI-II total score > 20) or suicidal ideation (BDI-II item 9>0), had serious psychological or medical conditions (including epilepsy and migraine), restricted 3D sight or chronic medication intake (except intake of oral contraceptives) and females if they were pregnant. Participants were instructed to abstain from alcohol and medication intake for 12h and psychoactive substances (including benzodiazepines) for 5 days before study days.

The study protocol, including the definition of primary and secondary outcomes and statistical analysis plan, was approved by the Independent Ethics Committee (IEC; Ethics Committee of North-West and Central Switzerland) before the start of the study. There were no deviations from the protocol after the trial start. The trial was registered at ClinicalTrials.gov before the start of the study with the identifier NCT03970187 on May 31, 2019. Final data was collected on September 19, 2019. All research was performed in accordance with the Declaration of Helsinki. All participants gave written informed consent for trial participation. Participants received compensation of CHF 25/h for their participation and CHF 50 for the successful completing of the study.

Randomization and masking. After study inclusion, participants were randomly (stratified for sex) allocated to the treatment group (VR gaze training in study phases 1 and 2) or the control group (a fear-unrelated VR task in study phase 1, no intervention in study phase 2, Fig. 2). Each eligible participant was allocated to one of two randomization lists (male/female). Within each list, groups were randomized according to a maximum tolerated imbalance (MTI-) procedure implemented in R ('RandomizeR'-package; MTI across lists = 4).
Allocation concealment was given for the experimenter who enrolled participants, as treatment allocation was made by a different experimenter. Therefore, the experimenter who enrolled participants did not know in advance which treatment the next person gets. All experimenters who collected outcome measures were unaware of participants' group assignment (single-blind).

Procedures. Upon contacting the study team, potential participants received detailed information about the study including the inclusion and exclusion criteria by email. Individuals who met an exclusion criterion during the subsequent online screening via SoSci Survey (42) were directly informed that they were not eligible for participation. Eligible individuals were scheduled for two visits at the Division of Cognitive Neuroscience at the University of Basel, Switzerland. Before study enrolment, we rechecked inclusion and exclusion criteria and collected basic demographic data. If the enrolment criteria were still met, participants filled in the German versions of the Social Phobia Inventory (SPIN) and the Brief Fear of Negative Evaluation Scale-Revised (FNE-K) to evaluate their general social anxiety and fear of negative evaluation by others and were allocated to one of the two intervention groups.
We then collected salivary cortisol with a saliva-sampling device (Salivette® from Sarstedt, Rommelsdorf, Germany) and conducted the baseline PST. The PST represents a real-life assessment of public speaking anxiety in a socially threatening situation. The participant's task was to perform three semi-improvised speeches in front of an evaluation committee consisting of three experimenters trained to maintain eye contact and a neutral facial expression and body posture. Participants had to choose three out of five predefined general topics and prepared the speeches for 10 minutes. Before presenting the first topic, participants rated their subjective fear – the primary outcome measure – as well as their fear to hold eye contact on the Subjective Units of Distress Scales (SUDS, "How high is your fear-level at this moment?"; scale of 0 = no

fear to 10 = maximum fear). The maximum duration of each given topic was 3 min. After each speech, the participants again rated their subjective fear and their fear of holding eye contact on the SUDS. The PST terminated when the participant indicated feeling too uncomfortable to proceed or when the full 9 minutes period (3×3 min) was over. After the last speech, participants and committee members independently and covertly rated the speech quality using visual analog scales (VAS). VAS were used to assess the global performance as well as eight specific subscales covering different aspects of the performance (verbal fluency, verbal expression, vocal modulation, tempo, posture, facial expression, eye contact, nervousness; scale of 0 = very bad, 100 = very good). Speeches were video and audio recorded, and gaze behavior was tracked throughout the PST using a mobile eye-tracking system (for further information see section 'Eye-tracking'). Following the PST, participants provided a second cortisol sample and filled out the Simulator Sickness Questionnaire (SSQ) (43) before starting the VR intervention accompanied by another experimenter.

For VR exposure to public speaking situations, we developed a stand-alone VR-based gaze exposure app for smartphones (for in-app content, Fig. 4). The app consists of an initial written description of the app content in 2D and three VR gaze exposure scenarios (i.e., proximity, classroom, lecture hall scenario). The exposure scenarios were based on 360° panoramic video clips of small to large audiences filmed in two different sized classrooms and a large lecture hall of the facilities of the University of Basel. In all three scenarios, the task was to maintain eye contact with the present virtual audience across six levels of increasing difficulty. The training procedure was based on a graduated behavioral exposure scheme and included no psycho-educative elements or specific cognitive interventions (e.g., challenging of cognitive distortions), but a minimal number of gamified reinforcement elements (i.e., clapping sounds after successful level completion).

At the beginning of each scenario, users are standing in an empty room (level 0). Following acclimatization to the VR, the room was increasingly filled with the virtual audience directing their eyes towards the user (level 1–6). Users are asked to briefly introduce themselves, which is played back via the microphone of the smartphone, and then to maintain eye contact with specific audience members, each indicated by an arrow. Advancing to further levels follows a predefined exposure scheme based on SUDS ratings ("How high was your subjective fear in this level?", scale of 0 = no fear to 10 = maximum fear) indicating the perceived fear at a given level and the maintenance of eye contact with the audience members, measured by gaze tracking (Fig.4b). Each level is repeated until the user indicates low subjective fear (> 3) on the SUDS rating, and the mutual gaze maintained exceeds a predefined time threshold ranging from 56 to 96 seconds, depending on the level (for a detailed description of the level settings, see SI Table S2). After completing one level, users are reinforced with applause sound effects (gamified reinforcement element). Each scenario contains six levels with a net duration of 3 minutes each (i.e., without introductory slides and SUDS-ratings). The levels are gradually increasing in difficulty, operationalized by the emotional valence of the audience's facial expression (level 1–2: positive, level 3–4: neutral, level 5–6: negative), the size of the audience (proximity scenario: 2–12, classroom scenario: 2–21, lecture hall scenario: 2–100) and the time required to keep eye contact (see Fig.4a,c and SI Table A2). Each exposure scenario is terminated by the time limit of 20 minutes, irrespective of the achieved level.

For the VR intervention in study phase 1, participants received Samsung smartphones with the preinstalled app, noise-canceling headphones, and a Google Daydream View version 2 VR headset (Fig. 4d) to enable 100-degrees stereoscopic view and a controller for the headset (for further information of the VR setup see SI Methods). For smooth integration of the VR intervention into our study course, an experimenter assisted in using the VR equipment and gave verbal instructions that were also provided within the app. Each 20 minutes scenario was followed by a 5 minutes break, and the first training session ended after completing all three scenarios. Participants of the control group received the same Samsung devices with Google Street View (v2.0.0.25751656) installed. Their task was to explore three pre-defined virtual scenarios in VR (Nautilus House, Mexico; Montpelier, Orange County; Wethersfield, Connecticut) for 20 minutes each and to answer several basic questions (e.g., "describe the weather in the scenario"). The control scenarios did not contain any stimuli of potential social threat. Participants were not prompted to indicate their fear level and could explore each scenario at their own pace by teleporting themselves using the controller.

After completion of the VR training session, symptoms of cybersickness were reassessed by the SSQ, followed by the Igroup Presence Questionnaire assessing presence in VR (44). Subsequently, participants

completed a second PST following a procedure identical to the first, but with five new speech topics to choose from. At the end of study phase 1, participants rated the VR app acceptability and usability by questionnaire, as well as the subjectively perceived improvement after app use regarding their fear, eye contact, and speech performance. Participants of the treatment group then received smartphones with the treatment app preinstalled, standard accessory charger and headphones, a Google Daydream View VR headset with controller, and detailed instructions about the home treatment.

For study phase 2, participants of the treatment group were instructed to use the treatment app for nine sessions within a maximum duration of 14 days, while the control group did not receive any task. The home treatment was requested to start 1-5 weeks after study phase 1. Participants were allowed to schedule the nine sessions freely but train only once a day. The sequence of the scenarios was predetermined (3× proximity scenario, 3× classroom scenario, and 3× lecture hall scenario). Each scenario lasted for 20 minutes and followed the procedure of the training on-site. The completion of a minimum of two-thirds of the planned VR home training (i.e., 6 completed scenarios with a minimum duration of 14 min) was defined per protocol as mandatory to be further included in the study.

The second assessment visit took place approximately two months after study phase 1 (mean [*SD*] number of days, treatment condition, 57.46 [3.46], control condition, 56.55 [6.41]) and one month after training completion of the treatment group (mean number of days: 34.00 [*SD* 4.44]). We rechecked inclusion and exclusion criteria and reassessed participants' fear of public speaking by the SPIN and the FNE-K. Afterward, we conducted a third PST identical to the first two PSTs in study phase 1, accompanied by cortisol sampling shortly before and afterward. Finally, participants indicated their subjectively perceived improvement of fear, eye contact and speech performance, as well as the amount of self-exposure to social situations between the two visits. The treatment group additionally filled in a VR app acceptability and usability scale and a questionnaire concerning their feeling of immersion during app use (44).

Outcomes

The primary outcome measure was subjective fear during the PST (SUDS, scale of 0 = no fear to 10 = maximum fear), averaged across four time points (i.e., before the first speech and after each speech). The main secondary outcomes were (1) the average relative dwell time on faces and (2) the overall global external assessment of performance by the evaluation committee in the PST. The average relative dwell time on faces was assessed by eye-tracking and represents the ratio of the participants' dwell time on faces of any of the committee members relative to the participants' total gaze time during the PST (for further information about the eye-tracking analysis see 'Eye-tracking'). The global external assessment of the performance in the PST (VAS ratings; 0: very bad, 100: very good) was assessed three times (i.e., after each speech) by each of the three committee members and averaged across time points and committee members to reach the overall global external assessment of performance.

Primary and secondary outcomes were evaluated before (baseline) and directly after the one-hour VR intervention in study phase 1 and on the second visit, scheduled 3–5 weeks after the last use of the treatment app during study phase 2 (Fig. 1). All further secondary outcome measures are described in the, together with their statistical analysis.

Eye-Tracking

We used eye-tracking to characterize participants' social gaze behavior during the PST. Both eyes were tracked at 120 Hz by a mobile, head-mounted eye-tracking system (Pupil Lab Core, Berlin, Germany; field of view: 100°, world-camera: 30Hz, 720p). The eye-tracking was controlled by the software Pupil Capture (v1.12.17). Gaze detection and mapping were based on a 2D model. At the beginning of the first PST, a built-in 10-point calibration procedure was run and repeated until an accuracy of at least 0.5° and precision of at least 0.2° was achieved. Calibration targets were used, which were projected at a distance of 2.5 m. To keep the distance to the calibration plane constant, we instructed participants to stay within an area marked on the ground during their speeches. After each speech, we run a 5-point accuracy test, allowing for offline recalibration in case of shift/slippage. Speeches, where the predefined minimum accuracy and precision were neither reached by online calibration nor offline recalibration, were excluded from further analyses (<5%). Blinks (filter length: 0.2 s, onset and offset confidence: 0.5) and fixations (maximum dispersion: 3°, minimum duration: 0.3 s, confidence threshold: 0.75) were detected by the default settings of the software.

To quantify pupil size, we used the 2D pupil detection algorithm with default settings (pupil range: 10–100, pupil size and diameter indicated in image pixels as observed in the eye image frame, confidence threshold: 0.6). We controlled for artifacts (e.g., blinks or movement) by excluding pupil diameter outliers. A data point x was defined as an outlier if x<M–3×SD or x>M+3×SD or if it was labeled as a blink and replaced by linear interpolation. Subsequently, pupil recordings were smoothed using a sliding average (83 ms time 90 window, ten samples; see 28). The light settings and exact position of the participants in the room were held constant throughout the entire experiment to control for confounding effects of the lighting conditions. All recorded data were visualized and exported by pupil player (v1.14.9), with the default minimum data confidence of 0.6.

Since we used a wearable eye-tracking headset, the coordinate system of the exported gaze data was fixed to the participant's head, not to the real world. However, automated face detection allowed us to detect the location of the committee's faces. To this end, we implemented a Convolutional Neural Network (CNN) to detect the location of the committee's faces with a face detector available in the dlib library (Python3.7). We implemented a face detection algorithm on the video recorded by the frontal camera (exported mp4-file) on a GPU cluster (NVidia Tesla v100) to extract the committee's faces in each video frame. All detected faces of each video frame were defined as the area of interest (AOI; see Fig. 5).

To quantify fear-related gaze behavior as a main secondary outcome measure, we calculated the duration of all valid fixations that were lying in the AOI (i.e., one of the three faces of the PST committee). We then divided this value by the total duration of all valid fixations to account for different speech durations. The resulting measure is the relative dwell time on faces. As a further secondary outcome measure and another marker of attention towards the committee, we counted the relative number of fixations on faces, irrespective of their length. Finally, we calculated the mean pupil diameter during each speech as a potential marker of general task difficulty and engaged attention. All eye-tracking measures were averaged within participants and PSTs.

Statistical analyses

We performed a per-protocol analysis. All data were analyzed in R version 3.6.1 and validated by a second statistician. We applied linear mixed models (‘nlme’ package) in combination with ANOVA (SS II). The primary and secondary outcomes were analyzed in separate models.

For the primary outcome analysis, a linear mixed model included the SUDS fear rating during the PST as the dependent variable. Group (treatment vs. control group; between-subject factor), time point (t0: baseline, t1: after study phase 1, t2: after study phase 2; within-subject factor) as well as their interaction were included as independent variables. The Participant-ID was included as the random effect. In case of significant interactions, post-hoc tests were applied to characterize the effects of group further. Age and sex were included as covariates. To account for baseline differences, we also included the measurement of the SUDS fear rating at baseline testing in case of post-hoc analyses. Potential significant interactions of covariates with either of the independent variables were further described by post-hoc tests. Non-significant interactions were removed from the statistical model.

Furthermore, to account for loss of participants between study phase 1 and study phase 2, we conducted an intention to treat (ITT) analysis for our primary outcome using the same model specified above. For missing outcomes, we applied the method Last Observation Carried Forward (LOCF) including the last available value of every subject that was reliably assessed. Group assignment was maintained according to randomization for participants entering the ITT.

For the main secondary outcomes’ analysis, we replaced the primary outcome measure as the independent variable with each main secondary outcome in a separate model that was otherwise identical. They were statistically treated the same way as the primary outcome measures but corrected for multiple comparisons ($p < .025$, corresponding to Bonferroni correction for two independent tests). We present results as mean (*SD*) for the intervention and control condition, and associated two-sided p values, as well as adjusted group difference with 95% CIs (‘emmeans’-package). We estimated Cohen’s d as effect size measurement. The estimate of d was based on t values of the linear models. Therefore, d is corrected for the effects of all confounding variables included in the model. By convention, $d = 0.2$ is considered to be a small, $d = 0.5$ to

be an intermediate and $d = 0.8$ to be a large effect. Based on previous studies with brief as well as continuous VR interventions using concepts of exposure therapy and including participants with fear of public speaking and public performance, we expected medium to large effect sizes (45). We estimated a required minimum sample size of $N = 80$ based on a power analysis assuming two-sample t-tests, equivalent to the least complex post-hoc test performed within the linear mixed model analyses (with a power of 80% at $\alpha = .05$, two-tailed; Cohen's $d = 0.65$; software: G*power 3.1).
A clinical trial monitor oversaw data collection and data entry according to a written monitoring plan approved by the IEC before trial conduction.

Acknowledgments
We thank Gabriel Guerra, Galya Iseli and Theresa Müggler for their work as experimenters in the PST. We thank Daria Bühler, Milena Cavegn, Tatjana Fuchs, Christiane Gerhards, Irena Kovacic, Cécile Longoni, Anna Magos, Johanna Otte, Jule Schröder, Nina Waldthaler and Johanna Weibel for helping in trial conduction as well as Amanda Aerni for monitoring the study. This study was funded by the Transfaculty Research Platform of the University of Basel and the Swiss National Science Foundation (SNSF). The views expressed are those of the author(s) and not necessarily those of the SNSF.

Data Sharing
Pseudonymized data will be made available upon reasonable request, which must include an approved ethics protocol and statistical analysis plan. R code for the analyses is freely available from the corresponding author.

References

1. R. Cañigueral, A. F. de C. Hamilton, The Role of Eye Gaze During Natural Social Interactions in Typical and Autistic People. *Front. Psychol.* **10** (2019).

2. R. S. Hessels, How does gaze to faces support face-to-face interaction? A review and perspective. *Psychon. Bull. Rev.* **27**, 856–881 (2020).

3. D. Tang, B. J. Schmeichel, Look Me in the Eye: Manipulated Eye Gaze Affects Dominance Mindsets. *J. Nonverbal Behav.* **39**, 181–194 (2015).

4. M. Weick, C. McCall, J. Blascovich, Power Moves Beyond Complementarity: A Staring Look Elicits Avoidance in Low Power Perceivers and Approach in High Power Perceivers. *Pers. Soc. Psychol. Bull.* **43**, 1188–1201 (2017).

5. N. J. Emery, The eyes have it: the neuroethology, function and evolution of social gaze. *Neurosci. Biobehav. Rev.* **24**, 581–604 (2000).

6. K. M. Dalton, *et al.*, Gaze fixation and the neural circuitry of face processing in autism. *Nat. Neurosci.* **8**, 519–526 (2005).

7. American Psychiatric Association, American Psychiatric Association, Eds., *Diagnostic and statistical manual of mental disorders: DSM-5*, 5th ed (American Psychiatric Association, 2013).

8. J. Chen, E. van den Bos, P. M. Westenberg, A systematic review of visual avoidance of faces in socially anxious individuals: Influence of severity, type of social situation, and development. *J. Anxiety Disord.* **70**, 102193 (2020).

9. F. R. Schneier, T. L. Rodebaugh, C. Blanco, H. Lewin, M. R. Liebowitz, Fear and avoidance of eye contact in social anxiety disorder. *Compr. Psychiatry* **52**, 81–87 (2011).

10. R. Lowe, *et al.*, Avoidance of eye gaze by adults who stutter. *J. Fluen. Disord.* **37**, 263–274 (2012).

11. J. K. Langer, T. L. Rodebaugh, Social Anxiety and Gaze Avoidance: Averting Gaze but not Anxiety. *Cogn. Ther. Res.* **37**, 1110–1120 (2013).

12. K. Horley, L. M. Williams, C. Gonsalvez, E. Gordon, Face to face: visual scanpath evidence for abnormal processing of facial expressions in social phobia. *Psychiatry Res.* **127**, 43–53 (2004).

13. J. W. Weeks, A. N. Howell, A. Srivastav, P. R. Goldin, "Fear guides the eyes of the beholder": Assessing gaze avoidance in social anxiety disorder via covert eye tracking of dynamic social stimuli. *J. Anxiety Disord.* **65**, 56–63 (2019).

14. M. J. Wieser, P. Pauli, G. W. Alpers, A. Mühlberger, Is eye to eye contact really threatening and avoided in social anxiety?—An eye-tracking and psychophysiology study. *J. Anxiety Disord.* **23**, 93–103 (2009).

15. N. T. M. Chen, P. J. F. Clarke, Gaze-Based Assessments of Vigilance and Avoidance in Social Anxiety: a Review. *Curr. Psychiatry Rep.* **19** (2017).

16. H. A. Paul, A Review of: "Hofmann, S. G., & Dibartolo, P. M. (Eds.). (2010). Social Anxiety: Clinical, Developmental, and Social Perspectives. Second Edition." *Child Fam. Behav. Ther.* **33**, 366–371 (2011).

17. J. K. Langer, T. L. Rodebaugh, Social Anxiety and Gaze Avoidance: Averting Gaze but not Anxiety. *Cogn. Ther. Res.* **37**, 1110–1120 (2013).

18. R. M. Rapee, R. G. Heimberg, A cognitive-behavioral model of anxiety in social phobia. *Behav. Res. Ther.* **35**, 741–756 (1997).

19. S. H. Spence, R. M. Rapee, The etiology of social anxiety disorder: An evidence-based model. *Behav. Res. Ther.* **86**, 50–67 (2016).

20. , Cognitive Behavioral Model Of Social Phobia (Clark, Wells, 1995). *Psychol. Tools* (April 12, 2022).

21. J. Reichenberger, M. Pfaller, A. Mühlberger, Gaze Behavior in Social Fear Conditioning: An Eye-Tracking Study in Virtual Reality. *Front. Psychol.* **11** (2020).

22. L. Schulze, B. Renneberg, J. S. Lobmaier, Gaze perception in social anxiety and social anxiety disorder. *Front. Hum. Neurosci.* **7** (2013).

23. N. T. M. Chen, P. J. F. Clarke, C. MacLeod, I. B. Hickie, A. J. Guastella, Aberrant Gaze Patterns in Social Anxiety Disorder: An Eye Movement Assessment during Public Speaking. *J. Exp. Psychopathol.* **7**, 1–17 (2016).

24. N. T. M. Chen, P. J. F. Clarke, Gaze-Based Assessments of Vigilance and Avoidance in Social Anxiety: a Review. *Curr. Psychiatry Rep.* **19** (2017).

25. A. Moukheiber, *et al.*, Gaze avoidance in social phobia: Objective measure and correlates. *Behav. Res. Ther.* **48**, 147–151 (2010).

26. K. Horley, L. M. Williams, C. Gonsalvez, E. Gordon, Social phobics do not see eye to eye: A visual scanpath study of emotional expression processing. *Anxiety Disord.*, 12 (2003).

27. J. Canton, K. M. Scott, P. Glue, Optimal treatment of social phobia: systematic review and meta-analysis. *Neuropsychiatr. Dis. Treat.* **8**, 203–215 (2012).

28. A. N. Howell, D. A. Zibulsky, A. Srivastav, J. W. Weeks, Relations among Social Anxiety, Eye Contact Avoidance, State Anxiety, and Perception of Interaction Performance during a Live Conversation. *Cogn. Behav. Ther.* **45**, 111–122 (2016).

29. I. Konovalova, J. V. Antolin, H. Bolderston, N. J. Gregory, Adults with higher social anxiety show avoidant gaze behaviour in a real-world social setting: A mobile eye tracking study. *PLoS ONE* **16**, e0259007 (2021).

30. N. T. M. Chen, L. M. Thomas, P. J. F. Clarke, I. B. Hickie, A. J. Guastella, Hyperscanning and avoidance in social anxiety disorder: The visual scanpath during public speaking. *Psychiatry Res.* **225**, 667–672 (2015).

31. C. A. Pollard, J. G. Henderson, Four types of social phobia in a community sample. *J. Nerv. Ment. Dis.* **176**, 440–445 (1988).

32. C. B. Pull, Current status of knowledge on public-speaking anxiety: *Curr. Opin. Psychiatry* **25**, 32–38 (2012).

33. A. M. Ruscio, *et al.*, Social fears and social phobia in the USA: results from the National Comorbidity Survey Replication. *Psychol. Med.* **38**, 15–28 (2008).

34. R. W. Seim, C. R. Spates, The Prevalence and Comorbidity of Specific Phobias in College Students and Their Interest in Receiving Treatment. *J. Coll. Stud. Psychother.* **24**, 49–58 (2009).

35. M. B. Powers, P. M. G. Emmelkamp, Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *J. Anxiety Disord.* **22**, 561–569 (2008).

36. S. Stupar-Rutenfrans, L. E. H. Ketelaars, M. S. van Gisbergen, Beat the Fear of Public Speaking: Mobile 360° Video Virtual Reality Exposure Training in Home Environment Reduces Public Speaking Anxiety. *Cyberpsychology Behav. Soc. Netw.* **20**, 624–633 (2017).

37. R. Reeves, A. Elliott, D. Curran, K. Dyer, D. Hanna, 360° Video virtual reality exposure therapy for public speaking anxiety: A randomized controlled trial. *J. Anxiety Disord.* **83**, 102451 (2021).

38. P. Premkumar, *et al.*, The Effectiveness of Self-Guided Virtual-Reality Exposure Therapy for Public-Speaking Anxiety. *Front. Psychiatry* **12** (2021).

39. S. L. Rogers, C. P. Speelman, O. Guidetti, M. Longmuir, Using dual eye tracking to uncover personal gaze patterns during social interaction. *Sci. Rep.* **8** (2018).

40. L. C. Kegel, *et al.*, Dynamic human and avatar facial expressions elicit differential brain responses. *Soc. Cogn. Affect. Neurosci.* **15**, 303–317 (2020).

41. A. P. Association, *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)* (American Psychiatric Pub, 2013).

42. , *Leiner, D. J. (2019). SoSci Survey (Version 3.1.06) [Computer software]. Available at https://www.soscisurvey.de.*

43. R. S. Kennedy, N. E. Lane, K. S. Berbaum, M. G. Lilienthal, Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *Int. J. Aviat. Psychol.* **3**, 203–220 (1993).

44. T. W. Schubert, The sense of presence in virtual environments: *Z. Für Medien.* **15**, 69–71 (2003).

45. E. Carl, *et al.*, Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *J. Anxiety Disord.* **61**, 27–36 (2019).

Figures and Tables

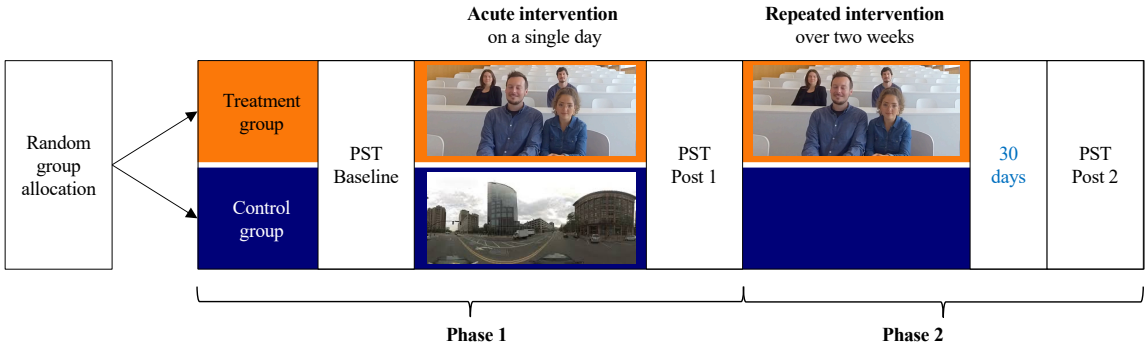

Figure 1: Schematic study procedure.
Left to right: Individuals were randomly allocated to either the treatment or the control group. Afterwards, the baseline public speaking test (PST) was conducted, which comprised a preparation time of 10 min and three speeches with a maximum duration of 3 min each. Before the first and after every speech, participants were asked to indicate their fear level, which was the primary outcome measure. During the speeches, eye tracking allowed us to quantify the relative dwell time on faces of any of the committee members during the PST as a main secondary outcome measure. After the first PST, the participants of the treatment group underwent the gaze exposure treatment in the VR app for 3×20 min, while the participants of the control group explored virtual scenarios without social threat for the same amount of time (acute intervention). The second PST followed a procedure identical to the first one. In study phase 2, the treatment group was requested to complete 9×20 min home training sessions with the gaze exposure app (repeated intervention), while the control group did not receive any task. One month after training completion of the treatment group, a third PST following the same procedure was conducted.

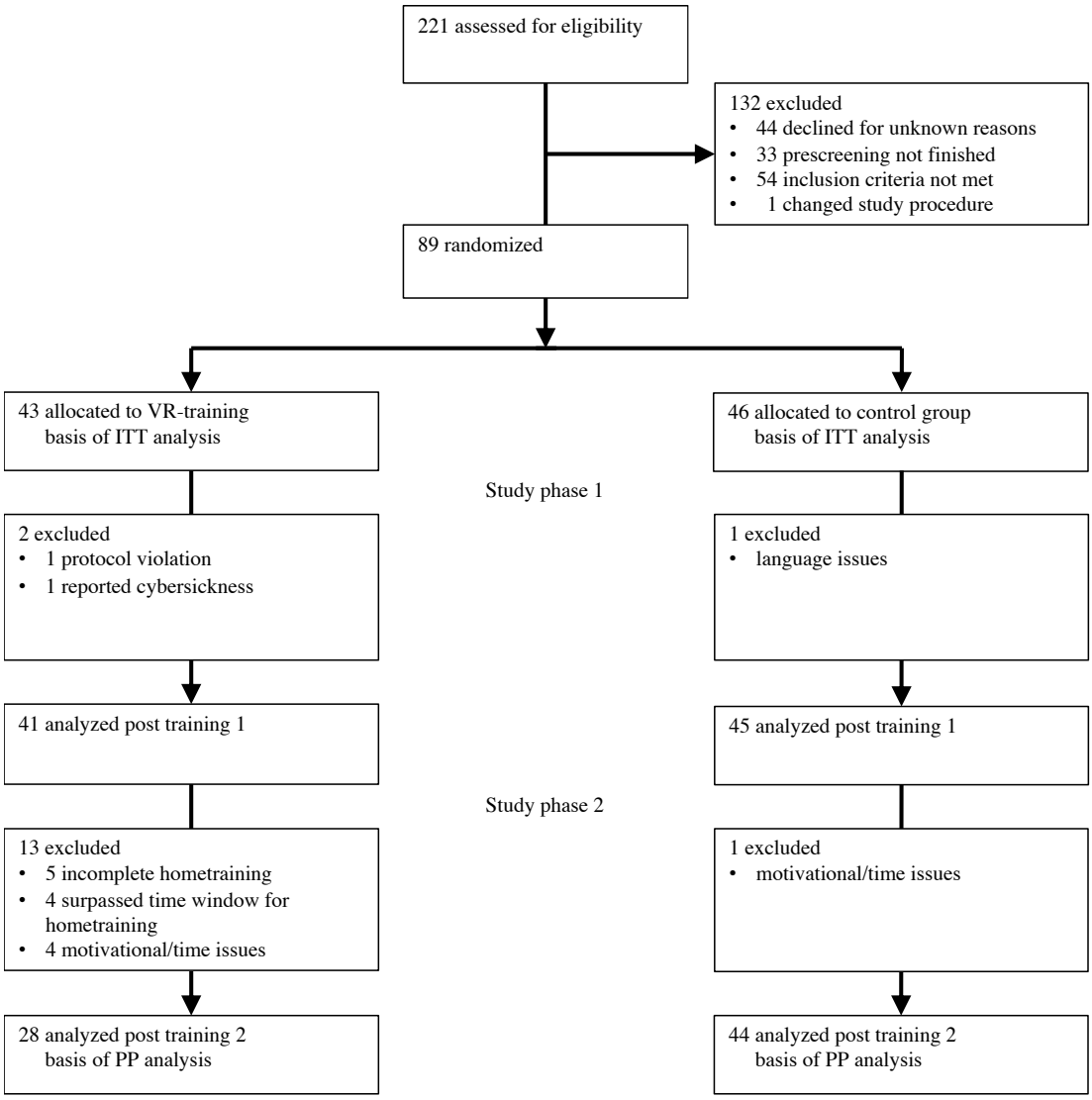

Figure 2. Flowchart of participants. VR, virtual reality; ITT, intention to treat; PP, per protocol. In study phase 2, more participants were excluded from the treatment group compared to the control group. To quantify a potential attrition bias, we performed an intention to treat (Last-Observation-Carried-Forward, LOCF) analysis for the primary outcome measure in addition to the analyses defined in the protocol.

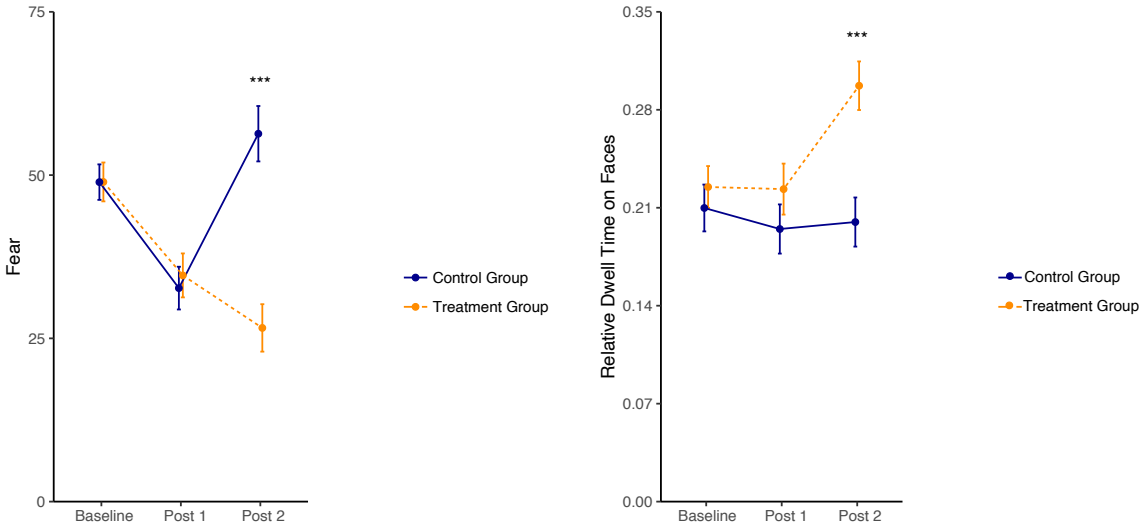

Figure 3. Effects of the gaze exposure treatment on perceived fear and relative dwell time on faces during public speaking. The per-protocol analysis suggested no reduction regarding the subjectively perceived fear and the relative dwell time on faces during the public speaking test in the treatment group immediately after 1-h VR exposure in study phase 1 (Post 1) compared to the control group. However, there was a beneficial effect on subjective fear and the relative dwell time after additional home trainings (9 × 20 min) with the gaze avoidance treatment app, as evident by the difference in fear ratings between the groups 1 month after intervention phase 2 (Post 2).

*** $p < 0.001$. Error bars indicate standard errors of the mean.

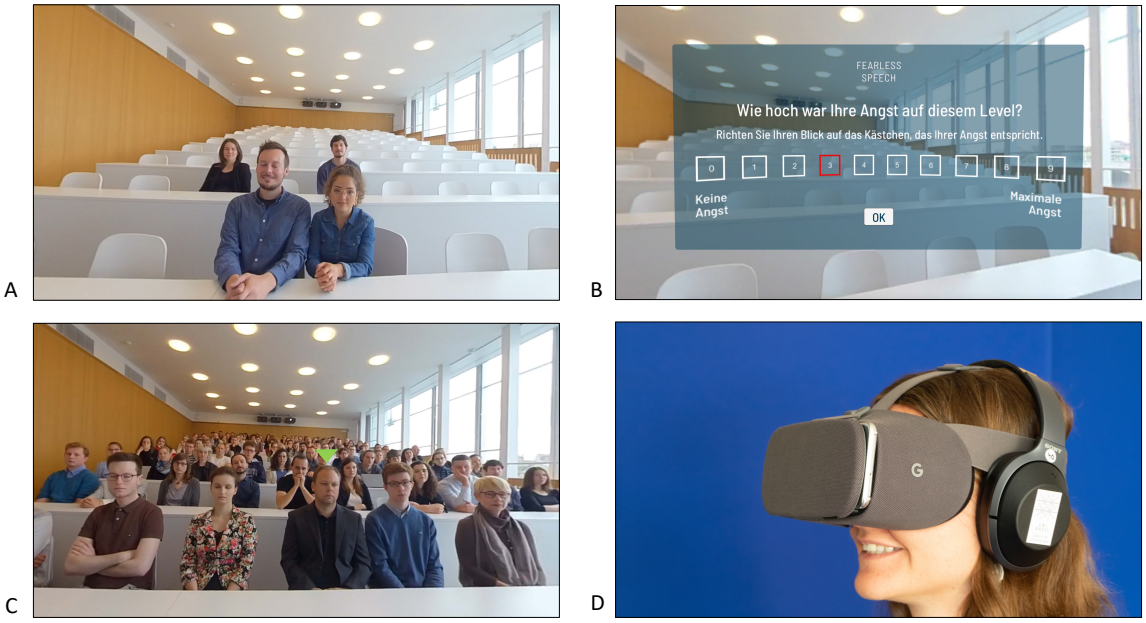

Figure 4. Virtual reality-based gaze exposure application. 4a and **4c** show examples of the lecture hall scenario used in the gaze exposure application. **4a** shows level 1; 4 people with positive emotional expression. **4c** shows level 6; 100 people with negative emotional expression; The green arrow indicated the current target person with whom eye contact should be maintained. As soon as a user gazed towards the eye area of the target person, the green arrow disappeared, and an invisible timer started. After a predefined time, the arrow switched to the next target person to prompt a change of eye contact. Whenever a user exited the target area before the predefined time, the timer paused, and the green arrow reappeared above the current target. **4b** shows an example of the fear ratings ("How high was your subjective fear in this level?", scale of 0 = no fear to 10 = maximum fear) that followed each successful eye contact maintenance. Each level is repeated until the user indicates low subjective fear (<3). **4d** shows the equipment of the study worn by a team member: an android smartphone with a preinstalled application, noise cancelling headphones and a Google Daydream View version 2 VR headset.

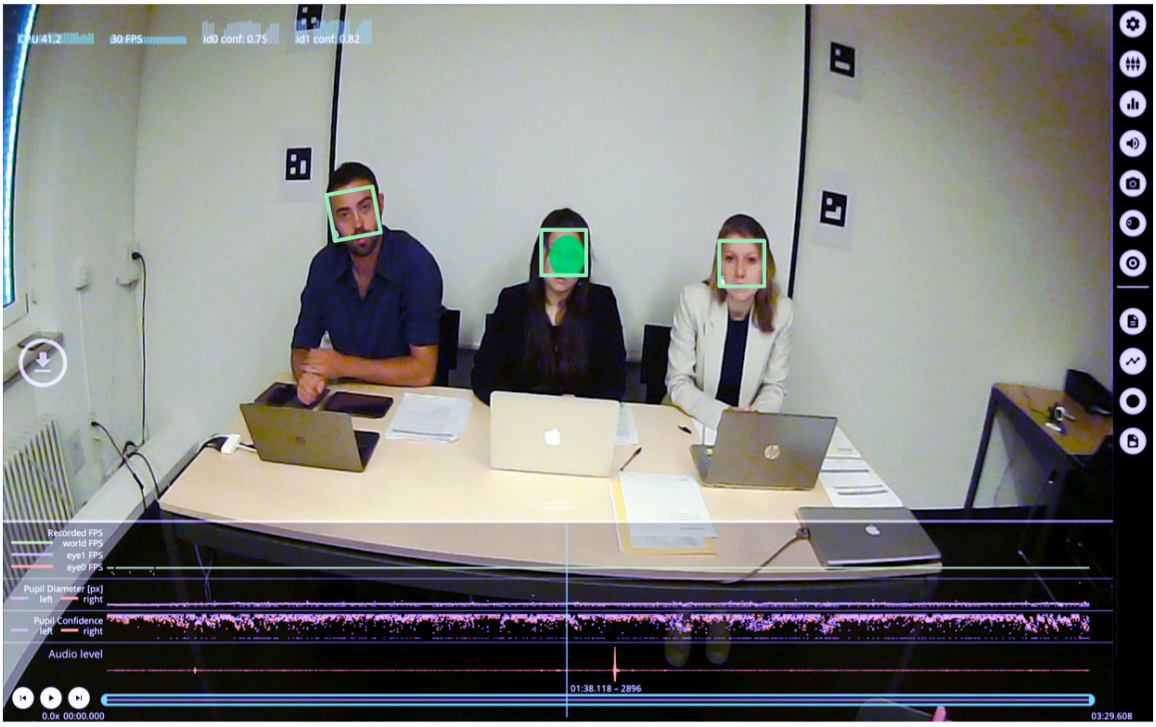

Figure 5. Example of face detection. A video frame is shown, which was recorded from the world camera of the mobile eye-tracking system, thus representing a participant's perspective during the PST. The green boxes mark the bounding area of the faces of the three committee members that were recognized by a deep learning algorithm (Convolutional Neural Network). The total area comprised by the boxes was defined as the area of interest (AOI), and fixations within (as indicated by the green dot) were interpreted to reflect mutual face gaze.

Table 1. Demographic characteristics of participants.

	Treatment group			Control group		
Intervention phase	**Baseline**	**Post 1**	**Post 2**	**Baseline**	**Post 1**	**Post 2**
Participants included	43	41	28	46	45	44
Age (years) (SD)	26.7 (5.5)	26.7 (5.6)	26.2 (5.1)	28.2 (6.2)	28.1 (6.2)	28.1 (6.2)
Men	14	14	9	17	17	17
Women	29	27	19	29	28	27
Education						
Bachelor degree or higher	20	19	14	22	21	20
Vocational education	7	6	3	10	10	10
High school education	13	13	9	10	10	10
Compulsory education	1	1	1	3	3	3
Other	2	2	2	1	1	1

Data are numbers of participants or means (*SDs*).

Table 2. Primary and main secondary outcome measures and differences between conditions. Post-hoc analyses for each time point separately were only conducted in case of significant time point × group interactions. The adjusted group differences are only indicated if significant. Descriptive values are means (*SDs*).

		Treatment group		Control group		p value (F-value)	Adjusted group difference (95% CI)	Cohen's d
		Mean (SD)	n	Mean (SD)	n			
Primary outcome	**PST fear (per protocol)**							
	Time point × group					<0.0001 (F[2, 154]=23.32)		
	Baseline	49.0 (19.5)	43	48.9 (18.4)	46	0.81		
	Post training 1	34.6 (21.5)	41	32.7 (21.9)	45	0.20		
	Post training 2	26.6 (19.2)	28	56.3 (28.1)	44	<0.0001	−29.8 (−41.8 to −17.9)	−1.07
	PST fear (ITT analysis)							
	Time point × group					<0.0001 (F[2,174]=23.15)		
	Baseline	49.0 (19.5)	43	48.9 (18.4)	46	0.15	..	..
	Post training 1	36.0 (22.0)	43	32.4 (21.8)	46	0.55	..	..
	Post training 2	30.2 (22.3)	43	55.2 (28.2)	46	<0.0001	−24.6 (−34.7 to −14.6)	−1.04
Main secondary outcomes	**PST relative dwell time on faces (per protocol)**							
	Time point × group					0.00026 (F[2, 151]=8.71)		
	Baseline	0.22 (0.10)	42	0.21 (0.11)	45	0.56	..	....
	Post training 1	0.22 (0.12)	40	0.19 (0.12)	44	0.20	..	0.97
	Post training 2	0.30 (0.09)	28	0.19 (0.12)	44	<0.0001	0.09 (0.05 to 0.13)	
	PST relative dwell time on faces (ITT analysis)							
	Time point × group					0.0030 (F[2,170]=6.07)		
	Baseline	0.22 (0.10)	43	0.21 (0.11)	46	0.56	..	..
	Post training 1	0.22 (0.12)	43	0.19 (0.12)	46	0.16	..	..
	Post training 2	0.28 (0.11)	43	0.20 (0.12)	46	0.00012	−0.07 (0.03 to 0.11)	0.83
	PST global external assessment of performance (per protocol)							
	Time point × group					0.79 (F[2,154]=0.24)		
	Baseline	47.5 (11.9)	43	45.2 (11.7)	46	..	..	..
	Post training 1	52.3 (11.1)	41	48.9 (10.9)	45	..	..	..
	Post training 2	55.1 (7.5)	28	52.2 (8.7)	44	..	..	..
	PST global external assessment of performance (ITT analysis)							
	Time point × group					0.37 (F[2,174]=1.01)		
	Baseline	47.5 (11.9)	43	45.2 (11.7)	46	..	..	..
	Post training 1	51.2 (12.6)	43	48.1 (11.9)	46	..	..	..
	Post training 2	52.1 (11.3)	43	51.0 (10.5)	46	..	..	..

PST, public speech test; ITT, intention to treat

Supplementary Information

Methods.

Further outcome measures. Further secondary outcomes measures with regards to behavior were (1) fear of eye contact during the PST (SUDS; 0: no fear of eye contact, 100: maximum fear of eye contact), measured analogously to the primary outcome measure, (2) the global self-assessment of performance in the PST, measured analogously to the global external assessment, but with regard to the participants' own rating, (3–5) the global subjectively perceived improvement of fear, eye contact and performance by the VR app (VAS-ratings; −100: much worse than before, 100: much better than before), (6) social anxiety, as measured by the SPIN (0–68, with higher scores corresponding to a higher burden) and (7) fear of negative evaluation, as measured by the FNE-K (12–60, with higher scores corresponding to greater distress).

Further secondary outcomes with regards to eye-tracking were (1) the average relative fixation frequency on faces as well as (2) the average pupil size during the PST (see 'Eye tracking'). The further secondary outcome regarding physiology was (1) the difference between salivary cortisol concentrations before and after the PST. Outcomes of further interest were a total of 16 measures (1–16) with regards to the externally (averaged across committee members) and self-assessed performance in the PST, characterized by 8 VAS ratings of detailed performance aspects (i.e., verbal fluency, verbal expression, vocal modulation, tempo, posture, facial expression, eye contact, nervousness; 0: very bad, 100: very good) and averaged within each PST, (17) the speech duration in seconds, averaged within each PST, (18) the feeling of presence in the VR-session, assessed by the IPQ (–42–42, with higher scores corresponding to a stronger sense of presence), (19) the usability (0–90, with higher scores indicating higher usability) and (20) social immersion (0–500, with higher scores indicating a higher degree of social immersion) of the app, assessed by questions created in-house as well as (21) the minutes of self-exposure to social situations since phase 1. Measures were taken at baseline and after the first and second intervention phase, with the following exceptions: Measures directly addressing improvement (i.e., of fear, eye contact, and performance in the VR app) were assessed after intervention phase 1 and at intervention phase 2. Scores for SPIN, FNE-K, and the cortisol difference were assessed at baseline and intervention phase 2. The IPQ and usability of the app were assessed after intervention phase 1, and the amount of self-exposure only at intervention phase 2. Additionally, only the treatment group provided fear ratings during VR exposure and ratings of the usability and social immersion of the app, rated after the home training at intervention phase 2.

Statistical analysis of further outcome measures. For the analysis of the further secondary outcomes and the outcomes of further interest, we replaced the primary outcome measure as the independent variable by them. This was done for all of them in a separate model that was otherwise identical. For measures directly addressing improvement (i.e., of fear, eye contact, and performance in the VR app), no baseline differences were accounted for in the case of post-hoc analyses. For variables that were only assessed at one time point (i.e., IPQ and usability of the app for both groups only after intervention phase 1, amount of self-exposure to social situations only after intervention phase 2), analyses were reduced to linear models. Variables only assessed for one group (i.e., the app usability and social immersion after intervention phase 2) are reported in a descriptive way. Further secondary outcome measures were grouped according to behavior (7, $p < 0.007$), eye tracking (2, $p < 0.025$), and physiology (1, $p < 0.05$) and corrected for multiple comparisons within those categories. Outcomes of further interest were analyzed in an explorative way. Therefore, statistical significance is reported based on nominal p values.

Gaze training application. The gaze training app was developed using Unity3D (v2018.3.11f1; Unity Technologies, San Francisco, CA, USA) under MacOS Mojave (v10.14.6) and compiled into a standard Android Package file (.apk). All visual material is based on 360° panoramic video clips taken by a 360°-camera (Insta360° Pro; Insta360, Shenzhen, GD, China) and stitched with Insta360Sticher (v3.0.0). Videos are used as skybox textures and all panels were created in the world coordinate system in the Unity3D game engine. All audio material (including logo and effects sound) was produced using Ableton Live 10 Suite (v10.0.1). The VR environment consists of 360° video clips, accompanied by sounds characteristic to each VR scenario and level (e.g., sounds of the audience rustling or coughing).

Table S1. Further outcome measures and differences between conditions. Post-hoc analyses for each time point separately were only conducted in case of significant time point × group interactions. The adjusted group differences are only indicated if significant. Descriptive values are means (*SDs*).

		Treatment group		Control group		p value (F-value)	Adjusted group difference (95% CI)	Cohen's d
		Mean (SD)	n	Mean (SD)	n			
Further secondary outcomes	Behavior	**PST fear of eye contact**						
		Time point × group				<0.0001 (F[2, 154]=18.58)		
		Baseline	37.6 (21.0)	43	37.4 (22.6)	46	0.75	..	..
		Post training 1	25.0 (20.1)	41	26.0 (23.3)	45	0.79	..	..
		Post training 2	18.2 (14.5)	28	49.1 (31.1)	44	<0.0001	–29.6 (–42.0 to –17.3)	–1.03
		PST global self–assessment of performance							
		Time point × group					<0.0001 (F[2, 154]=25.10)		
		Baseline	35.3 (18.3)	43	33.4 (15.4)	46	0.48	..	..
		Post training 1	47.0 (17.5)	41	43.8 (19.7)	45	0.68	..	..
		Post training 2	61.9 (21.1)	28	34.6 (20.0)	44	<0.0001	27.6 (18.1 to 37.1)	1.24
		Perceived improvement of fear							
		Time point × group					<0.0001 (F[1, 68]=21.95)		
		Baseline	..	..	..	..	..	..	..
		Post training 1	66.3 (15.5)	40	66.9 (16.9)	44	0.67	..	..
		Post training 2	76.3 (14.0)	28	49.2 (23.8)	44	<0.0001	26.1 (16.1 to 36.1)	1.27
		Perceived improvement of eye contact							
		Time point × group					<0.0001 (F[1, 68]=27.44)		
		Baseline	..	..	..	..	..	..	..
		Post training 1	63.2 (14.0)	40	60.9 (15.6)	44	0.60	..	..
		Post training 2	75.8 (16.3)	28	45.3 (21.5)	44	<0.0001	30.1 (20.4 to 39.8)	1.51
		Perceived improvement of performance							
		Time point × group					<0.0001 (F[1, 68]=24.94)		
		Baseline	..	..	..	..	..	..	..
		Post training 1	60.8 (15.2)	40	62.5 (16.6)	44	0.51	..	..
		Post training 2	73.3 (17.3)	28	46.6 (22.4)	44	<0.0001	25.9 (15.8 to 36.0)	1.25
		SPIN score							
		Time point × group					0.62 (F[1, 70]=0.24)		
		Baseline	21.7 (10.9)	43	22.7 (8.0)	46	..	..	..
		Post training 1	..	..	..	..	..	..	..
		Post training 2	17.6 (9.9)	28	17.8 (8.9)	44	..	..	..
		FNE-K score							
		Time point × group					0.78 (F[1, 70]=0.07)		
		Baseline	40.2 (10.7)	43	38.5 (9.2)	46	..	..	..
		Post training 1	..	..	..	..	..	..	..
		Post training 2	37.5 (10.2)	28	34.7 (10.6)	44	..	..	..

Eye-tracking	**PST relative fixation frequency on faces**							
	Time point × group					<0.0001 (F[2, 151]=8.26)		
	Baseline	0.21 (0.09)	42	0.20 (0.11)	45	0.57	..	..
	Post training 1	0.22 (0.11)	40	0.19 (0.11)	44	0.24	..	..
	Post training 2	0.29 (0.09)	28	0.19 (0.11)	44	<0.0001	0.09 (0.05 to 0.13)	0.94
	PST Pupil size							
	Time point × group					0.34 (F[2, 151]=1.09)		
	Baseline	48.6 (10.3)	42	47.7 (9.7)	45	..	..	..
	Post training 1	49.2 (10.9)	40	49.6 (10.7)	44	..	..	..
	Post training 2	53.1 (10.0)	28	49.3 (7.6)	44	..	..	..
Physiology	**PST Cortisol (Δ)**							
	Time point × group					0.74 (F[1, 70]=0.11)		
	Baseline	1.3 (5.7)	43	2.0 (6.8)	46	..	..	..
	Post training 1	..	..	..	..	..	..	..
	Post training 2	0.4 (4.7)	28	0.7 (4.7)	44	..	..	..

PST, public speech test; ITT, intention to treat; Δ=Cortisol levels after the PST−Cortisol levels before the PST.

Table S2. Level settings of the gaze exposure application. The levels were gradually increasing in difficulty, as operationalized by the emotional valence of the audience's facial expressions, the size of the audience as well as the time required to maintain face gaze. While the absolute number of faces required to gaze at remained constant, the dwell time required on each face was increased in higher levels, resulting in a longer total dwell time on faces needed for successful level completion.

Scenario	Level	Emotional valence of audience	Size of audience	Number of target faces	Number of faces required to gaze at	Dwell time required per face (s)	Total dwell time required on faces (s)
Proximity	1	Positive	2	2	8	7	56
Proximity	2	Positive	4	4	8	8	64
Proximity	3	Neutral	6	2	8	9	72
Proximity	4	Neutral	6	2	8	10	80
Proximity	5	Negative	9	3	8	11	88
Proximity	6	Negative	12	4	8	12	96
Classroom	1	Positive	2	2	8	7	56
Classroom	2	Positive	4	3	8	8	64
Classroom	3	Neutral	8	4	8	9	72
Classroom	4	Neutral	12	4	8	10	80
Classroom	5	Negative	16	5	8	11	88
Classroom	6	Negative	21	6	8	12	96
Lecture hall	1	Positive	4	2	8	7	56
Lecture hall	2	Positive	8	4	8	8	64
Lecture hall	3	Neutral	16	6	8	9	72
Lecture hall	4	Neutral	32	6	8	10	80
Lecture hall	5	Negative	64	6	8	11	88
Lecture hall	6	Negative	100	7	8	12	96

4.3 Augmented reality exposure therapy for specific phobias

Whereas VR brings the user into a digitally created environment, AR brings a digitally created object into the user's own environment. As a result, AR requires less time and cost to be developed and has the advantage of having to create only one object rather than an entire environment. Since single objects are much simpler than full environments, they can be created with a high degree of realism as well as with personal relevance to each user (Vinci et al., 2020a). With the additional element of interaction with the feared stimuli and seeing one's own body, AR offers promising features to create engaging treatment options for exposure-based treatments for specific phobias (Baus & Bouchard, 2014; Vinci et al., 2020a).

AR has been used in entertainment and gaming, maintenance, architecture, education in general as well as medical training in specific and in cognitive and motor rehabilitation (Chicchi Giglioli et al., 2015). In clinical psychology, AR has mainly been tested for the treatment of small animal phobias, such as cockroaches and spiders, showing clinically relevant effect sizes in reducing fear and avoidance behavior (Botella et al., 2010, 2016; Wrzesien et al., 2013). The AR then was desktop-based, needed a special headset and markers to detect the surface in order to place the virtual objects. The technological developments of smartphones have made it possible to display AR objects without any further devices, increasing the overall practicability and flexibility. Recent studies on substance abuse showed cue reactivity (e.g. cravings, urges) with large effect sizes for cue exposure in AR, indicating its suitability for exposure therapy even beyond small animal phobia (Brandon et al., 2021; Vinci et al., 2020b). A preliminary meta-analysis suggests that AR-based exposure therapies can be as effective as VRET or in vivo exposure (Suso-Ribera et al., 2019).

In our third study **C)**, we utilized the recent developments in AR technology and smartphones to project virtual objects into the real environment without any marker or further devices. We investigated the effectiveness of the developed AR smartphone app *Phobys* to reduce fear and avoidance behavior in individuals with fear of spiders. The intervention, too, is designed as a stand-alone home-training with an automated and gamified AR exposure scheme, navigated by touch with a simple, playful user interface (UI). The stimuli – a 3D spider model – is projected into the user's real environment, allowing a different type of immersion and interaction. In comparison to previous AR exposure studies on small animals, no further guidance is needed. This study sets itself apart by investigating the first AR app of its kind and additionally, testing its effectiveness in a real-life setting in addition to subjective measures.

4.3.1 Effectiveness of a smartphone-based, augmented reality exposure app to reduce fear of spiders in real-life: A randomized controlled trial

Contents lists available at ScienceDirect

Journal of Anxiety Disorders

journal homepage: www.elsevier.com/locate/janxdis

Effectiveness of a smartphone-based, augmented reality exposure app to reduce fear of spiders in real-life: A randomized controlled trial

Anja Zimmer [a,e], Nan Wang [a,e], Merle K. Ibach [a,e], Bernhard Fehlmann [a,e], Nathalie S. Schicktanz [a,e], Dorothée Bentz [a,e], Tanja Michael [f], Andreas Papassotiropoulos [b,c,d,e], Dominique J.F. de Quervain [a,d,e,*]

[a] Division of Cognitive Neuroscience, Faculty of Psychology, University of Basel, CH, Switzerland
[b] Division of Molecular Neuroscience, Faculty of Psychology, University of Basel, CH, Switzerland
[c] Life Sciences Training Facility, Department Biozentrum, University of Basel, CH, Switzerland
[d] University Psychiatric Clinics Basel, CH, Switzerland
[e] Transfaculty Research Platform, University of Basel, CH, Switzerland
[f] Division of Clinical Psychology and Psychotherapy, Department of Psychology, Saarland University, D, Germany

ARTICLE INFO

Keywords:
Fear of spiders
Specific phobia
Smartphone-based AR exposure app
Stand-alone
Behavioural approach test

ABSTRACT

Although in vivo exposure therapy is highly effective in the treatment of specific phobias, only a minority of patients seeks therapy. Exposure to virtual objects has been shown to be better tolerated, equally efficacious, but the technology has not been made widely accessible yet.

We developed an augmented reality (AR) application (app) to reduce fear of spiders and performed a randomized controlled trial comparing the effects of our app (six 30-min sessions at home over a two-week period) with no intervention. Primary outcome was subjective fear, measured by a Subjective Units of Distress Scale (SUDS) in a Behavioural Approach Test (BAT) in a real-life spider situation at six weeks follow-up.

Between Oct 7, 2019, and Dec 6, 2019, 66 individuals were enrolled and randomized. The intervention led to significantly lower subjective fear in the BAT compared to the control group (intervention group, baseline: 7.12 [SD 2.03] follow-up: 5.03 [SD 2.19] vs. control group, baseline: 7.06 [SD 2.34], follow-up 6.24 [SD 2.21]; adjusted group difference -1.24, 95 % CI -2.17 to -0.31; Cohen's d = 0.57, p = 0.010).

The repeated use of the AR app reduces subjective fear in a real-life spider situation, providing a low-threshold and low-cost treatment for fear of spiders.

1. Introduction

Specific phobias are among the most common anxiety disorders, with an estimated lifetime prevalence ranging from 3% to 15 % (Eaton, Bienvenu, & Miloyan, 2018), with fears of animals such as spiders representing one of the most common form (Oosterink, De Jongh, & Hoogstraten, 2009). For those affected, exposure to spiders induces immediate emotional and physiological reactions such as intense fear, panic or disgust, and accelerated heart rate, often resulting in avoidance of the feared stimulus (American Psychiatric Association, 2013; Davey, 2011). These reactions can impede people doing a variety of daily-life activities, resulting in functional impairment for the sufferers, a negative impact on interpersonal interactions and quality of life in general (Bandelow & Michaelis, 2015). Many studies have demonstrated that exposure-based treatments are among the most effective treatments for specific phobias including fear of spiders. The current gold-standard is in vivo exposure therapy, during which therapists expose patients to the feared stimuli in real-life (Choy, Fyer, & Lipsitz, 2007; Wolitzky-Taylor, Horowitz, Powers, & Telch, 2008).

However, many patients with specific phobias do not seek professional help, because they adapted their daily lives to their fear by trying to avoid any contact with the feared stimulus, e.g. a spider (Bandelow & Michaelis, 2015). Among the main reasons for the underuse of in vivo therapy ranks the fear of being exposed to a real phobic stimulus (Chambless & Ollendick, 2001). Further, there is a high drop-out rate of in vivo exposure treatment due to low acceptance (Choy et al., 2007). In vivo exposure therapy can be also challenging for psychotherapists. The intensity and level of perceived threat for the patient can never be fully

* Corresponding author at: Birmannsgasse 8, Basel, 4055, Switzerland.
E-mail address: dominique.dequervain@unibas.ch (D.J.F. de Quervain).

https://doi.org/10.1016/j.janxdis.2021.102442
Received 9 September 2020; Received in revised form 25 May 2021; Accepted 29 June 2021

controlled by the therapists, which might lead to concerns that exposure might be too stressful for their patients (Cook, Biyanova, Elhai, Schnurr, & Coyne, 2010; Öst, 1989). Additionally, for logistic reasons, in vivo exposure treatment can be challenging and time consuming. Consequently, there is a need for novel exposure-based treatment approaches that circumvent the raised limitations of conventional in vivo treatment.

The increasing success of using virtual reality (VR) in therapeutic settings has been documented in several studies. Virtual reality exposure therapy (VRET) has been demonstrated to be an effective and empirically validated alternative to in vivo exposure for several specific phobias (Carl, Stein, & Levihn-Coon, 2018; Wechsler, Kümpers, & Mühlberger, 2019). Furthermore, it has been reported that patients prefer exposure in VR over exposure in vivo (Garcia-Palacios, Botella, Hoffman, & Fabregat, 2007). Despite its good efficacy including transfer to real-life situations (Morina, Ijntema, Meyerbröker, & Emmelkamp, 2015) and high acceptability for both patients and therapists, exposure in VR is still quite restricted to laboratories and experimental studies (Botella, Fernández-Álvarez, Guillén, García-Palacios, & Baños, 2017), and only a minority of clinicians offers VR treatment (Segal, Bhatia, & Drapeau, 2011). Reasons might be the fear of potential technical difficulties, possible side effects of motion sickness, and the continuous monetary expenses for the latest equipment and software.

A variant of VR is augmented reality (AR), which augments the real world with virtual elements in real time (Azuma, 1997). In the treatment context, AR presents the same advantages as VR (e.g. control over the way the exposure is conducted, easier access to the threatening stimuli, no risk of real danger for the patient, a reduction in preparation time, exposure in the comfort of the therapy room or home), and its development only requires a few virtual elements to be designed, which reduces both costs and time of programming. Furthermore, a big advantage is that in AR the patient is able to see his or her own body while interacting with the virtual elements, which can enhance the patient's engagement in the treatment (Baus & Bouchard, 2014).

The use of AR in the treatment of mental disorders is still in its infancy. A few studies have shown that desktop computer-based AR can be successfully used to reduce fear in small animal phobia (cockroaches and spiders) (Botella et al., 2016; Chicchi Giglioli, Pallavicini, Pedroli, Serino, & Riva, 2015). In a preliminary comparison to VRET and in vivo exposure, treatment with exposure through AR has been shown to be equally efficacious to reduce fear in small animal phobia (Suso-Ribera, Fernández-Álvarez, & García-Palacios, 2019).

However, these treatments implementing AR were carried out under laboratory conditions and with continuous surveillance and guidance by an experimenter or clinician, limiting the translation into real-life practice. Additionally, these settings still needed markers in order to detect and specify the area in which the small animal should appear. A recent study showed the success of inducing fear of multiple animal species including spiders with a first markerless AR app (De Witte et al., 2020), paving the way to use mobile AR apps also for the treatment of specific phobias in exposure-based interventions.

In the present study, we developed a stand-alone, smartphone-based AR exposure app – *Phobys* – to reduce the fear of spiders. We also implemented game elements since it has been suggested that, with the appropriate design and use, digital games have the potential to be effective psychotherapeutic tools (Stetina, Felnhofer, Kothgassner, & Lehenbauer, 2012). The AR app was used as a home training with six 30-min sessions over a two-week period. This set-up allowed treatment under real-life and not laboratory conditions. We tested its effectiveness in a randomized controlled trial in subjects with clinical and subclinical fear of spiders. The main outcome measure, i.e. subjective fear in a Behavioural Approach Test (BAT) with a real spider, as well as the secondary outcome measures, such as the performance and subjective disgust in the BAT and the questionnaires to assess fear of spiders, were assessed at baseline and at six weeks follow-up.

2. Material and methods

2.1. Study design and participants

We performed a single-blind, parallel-group, randomized controlled trial to investigate real-life effectiveness of our stand-alone, smartphone-based gamified AR exposure app. We recruited physically healthy participants with fear of spiders from the German speaking general population of Switzerland by online advertisements. We included individuals with subclinical and clinical fear of spiders (DSM-5 (American Psychiatric Association, 2013)), aged 18–40 years. We excluded individuals if they currently received psycho- or pharmacotherapy, had ever been in treatment for fear of spiders or participated simultaneously in another study, showed signs of depression (Beck Depression Inventory II, BDI-II (Beck, Steer, Ball, & Ranieri, 1996) total score ≥ 20) or suicidal ideation (BDI-II item $9 > 0$), had a physical illness or chronic medication intake (except intake of oral contraceptives), were pregnant or had a BAT score over 8 at baseline. Participants were instructed to abstain from alcohol and medication intake for 12 h and from psychoactive substances (including benzodiazepines) for five days before days of testing.

The study protocol (including the definition of primary and secondary outcome measures and the statistical analysis plan) and all procedures were approved by the Ethics Committee of North-West and Central Switzerland (EKNZ) before the start of the study. All participants gave written informed consent for trial participation. Participants received a compensation of CHF 125.- for their participation. The study took place at the Division of Cognitive Neuroscience at the University of Basel, Switzerland. A clinical trial monitor oversaw data collection and entry according to a written monitoring plan approved by the EKNZ before trial conduction. This trial was registered at ClinicalTrials.gov with the identifier: NCT04162509.

2.2. Randomisation and masking

After study inclusion, participants were randomly (matched for the presence of a clinical diagnosis of fear of spiders and sex) allocated to the two groups (intervention group: gamified AR spider exposure app vs. control group: no intervention (all participants gained access to the app after trial participation)). We used two randomization lists for subjects with subclinical fear of spiders (male/female) and two for clinical fear of spiders (male/female). The groups were block-randomized within these randomization lists (in each block of six, three participants were randomly allocated to the intervention group and to the control group, respectively). The experimenter who collected the primary outcome measure in the real-life spider situation was unaware of the group assignment of the participants (single-blind).

2.3. Procedures

After a potential participant contacted the study team, more detailed information about the study was sent by email along with the main inclusion and exclusion criteria. Eligible participants were scheduled for the study. Before study enrolment, we checked all inclusion and exclusion criteria and collected basic demographic data. Fear of spiders was assessed by the section for specific phobia of the diagnostic interview for mental disorders of the DSM-5 (American Psychiatric Association, 2013). Depressive symptomatology and suicidal ideation were assessed by the Beck Depression Inventory (BDI-II (Beck et al., 1996), exclusion criteria BDI-II item $9 > 0$, BDI-II total score ≥ 20). Alcohol consumption and intake of prescribed or illicit drugs were inquired about. Participants filled out questionnaires to collect baseline measures for their fear of and beliefs about spiders (Fear of Spiders Questionnaire (FSQ) (Szymanski & O'Donohue, 1995), German version (Rinck et al., 2002), Spider Phobia Beliefs Questionnaire (SBQ) (Arntz, Lavy, Van den Berg, & Van Rijsoort, 1993), German version (Pössel & Hautzinger, 2003)) and general self-efficacy (General Self-Efficacy Scale (GSE) (Schwarzer

& Jerusalem, 1995), German version (Schwarzer & Jerusalem, 1999)). Finally, we conducted the baseline Behavioural Approach Test (BAT) in real-life that included the assessment of our primary outcome, the Subjective Units of Distress Scale (SUDS) of fear.

The BAT in vivo procedure was similar to the one used by Lass-Hennemann and Michael (2014). Participants were placed in front of a closed room and were asked to open the door and approach a living house spider measuring about 5 cm, which was placed in a sealed transparent plastic container on a table at the far end of the room. Participants were requested to approach the spider and interact with it as far as possible. In detail, the BAT comprised 13 steps: 0 = refuses to enter the test room, 1 = stops 5 m from the container, 2 = stops 4 m from the container, 3 = stops 3 m from the container, 4 = stops 2 m from the container, 5 = stops 1 m from the container, 6 = stops close to the windowsill with the container, 7 = touches the container, 8 = removes the lid, 9 = puts a hand in the container, 10 = touches the spider with one forefinger, 11 = holds the spider less than 20 s, and 12 = holds the spider for at least 20 s. Theses scores ranging from 0 to 12 were given when the BAT was completed (max. 3 min) or the participant indicated not be able to proceed any further during the BAT. To counteract a possible ceiling effect, we excluded participants who were already able to insert their hand into the box (step 9 of the 13 steps) during the baseline BAT.

Subsequently, participants of the intervention group received a short description of the mechanisms underlying exposure therapy and on how to use the AR app during their home training on the smartphones they were given. They further filled out a scale on credibility/expectancy for improvement (Borkovec & Nau, 1972). At the end of the first study day, all participants were assessed for adverse events and sent home if no safety concerns were present.

At follow-up six weeks later, we first asked about intake of alcohol or medications. Afterwards, we conducted the BAT, re-assessed the fear of spiders by the section for specific phobia of the diagnostic interview for mental disorders for DSM-5 (American Psychiatric Association, 2013) and participants again filled out the FSQ, SBQ and GSE and additionally one item each for self-reported reduction of fear and disgust in real-life. The intervention group additionally filled out a usability scale and a questionnaire concerning their feeling of immersion (Georgiou & Kyza, 2017).

2.4. Intervention

The AR app *Phobys* (Fig. 1) consists of eight levels with a pre-defined length of 2 min and a ninth level, which only lasts 30 s. Once the user opens the AR app, all the information on how to use it is given in written form. Task instructions (e.g. looking at the spider, approaching it, putting the hand underneath it) are given via small text pop-up windows. Each level starts with a surface scan of either a table, wall or the floor with a distance of approx. 1 m to the surface, followed by a tap on the display, which places the virtual spider in the scanned area and starts the timer of 2 min for level one to eight and 30 s for level nine, respectively.

The levels comprise different tasks of exposure and interaction with a realistic 3D AR spider model as follows: In level one, after the initial tap on the display to place the virtual spider, the user is instructed to stay at the distance of approx. 1 m to the table and watch the spider, which is not moving, from all sides. In level two, the user is instructed to move the smartphone closer to the spider until a sound indicates the distance to stop and watch the spider, which is not moving, from all sides. In level three, which is otherwise similar to level two, a certain distance (without sound) triggers the spider model to lift its front legs. In level four, the user is again instructed to approach the spider with the

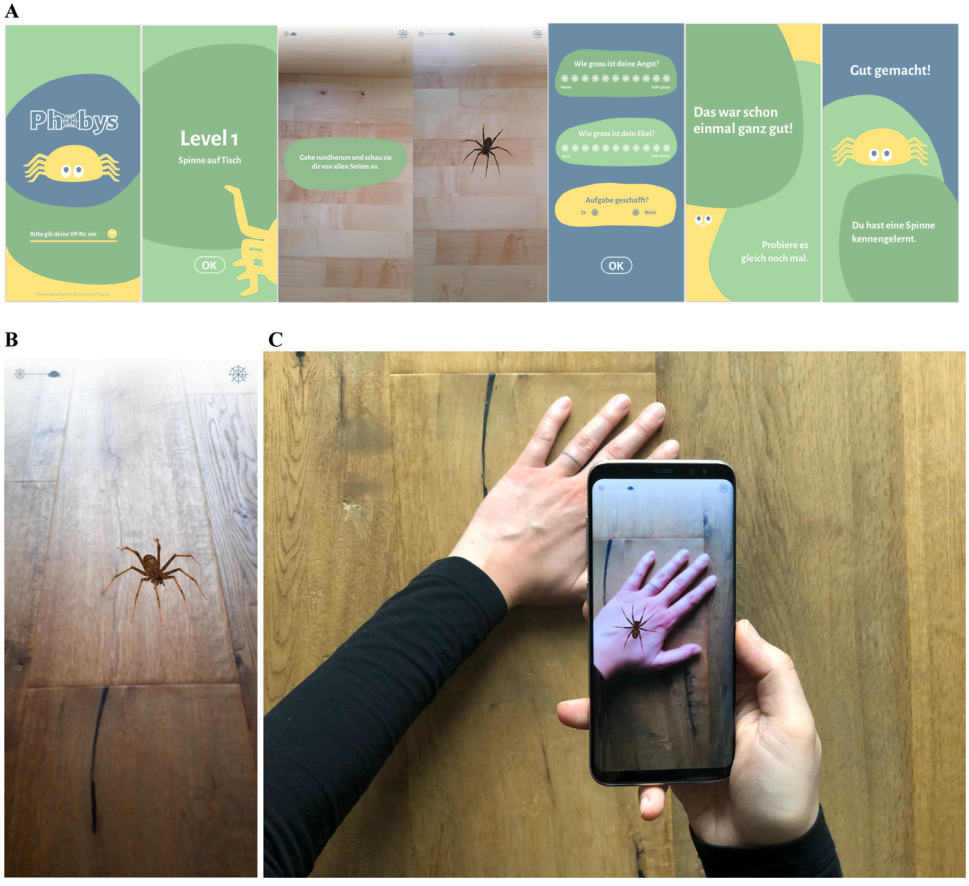

Fig. 1. Phobys. A) User interface, B) 3D AR spider model, and C) third person view of the app of level 5.

smartphone, whereby a certain distance (without sound) triggers the spider to walk away. Moving the smartphone back towards the user's body leads the spider to be walking towards the user, after which the task is repeated. In level five, the user is instructed to put the hand underneath the spider model. A net-icon should be tapped additionally to let the spider walk around in circles. In level six, the initial tap on the display places two spiders in the scene. Again, a net-icon should be tapped to allow the spiders to walk in circles and the user is asked to put the hand underneath them. For level seven, the smartphone is required to face a wall. Tapping on the display places ten spiders in the scene. The task is to collect the spiders by approaching them with the smartphone. Each spider disappears at a certain distance, accompanied by a sound effect. The spiders reappear after 10 s and should be continuously collected until the end of the level. For level eight, the smartphone is required to face the floor. Tapping on the display places many spiders in the scene with a path between them. The task is to follow the path and walk through the group of spiders, if possible several times. Level nine starts exactly as level one, with the smartphone facing the table and the placement of one spider in the scene through tapping on the display. However, this time a timer of only 30 s is started. The user is instructed to move the smartphone closer to the spider. A close distance triggers a kiss sound and hearts floating across the screen.

After each level a rating of fear and disgust on a continuous Subjective Units of Distress Scale (SUDS) from 0 to 10 is given as well as an indication, whether the tasks have been completed. From the first level the user proceeds to further levels according to a pre-defined exposure scheme. The users will repeat each level until their SUDS ratings of fear are 4 or below and have successfully completed the task.

The ratings are then followed by either a unique image (GIF) of an entertaining cartoon spider and rewarding sound effects if the level is completed (e.g. a spider clapping with a sound of cheering voices) or a standard screen informing the user that the current level will be repeated. This stepwise exposure is similar to the procedure developed by Öst for an in vivo intensive one-session exposure treatment for spider phobia (Chambless & Ollendick, 2001; Öst, 2012).

The AR app *Phobys* was developed at the University of Basel, Switzerland using Unity3D (version 2018.3.11f1 [64-bit] Unity Technologies, San Francisco, CA, USA) under MacOS Mojave (version 10.14.6) and compiled into a standard Android Package (.apk) file. The visual material of the user interface (UI) was created with Illustrator CC (version 2019), the GIF-files with a duration per frame of 0.1 s were animated in Photoshop CC (version 2019). The UI itself was designed in Sketch (version 56). The audio material (such as the sound effects) was produced using Ableton Live 10 Suite (version 10.0.1) under MacOS Mojave (version 10.14.6). The 3D spider model has been created by Computer Graphic (CG) designer, M. Gabriel Casamasso (artstation.com/gabrielcasamasso). Blender (version 2.79b) was used to create the spider geometry CG model. Inverse Kinematics (IK) was applied to the articulated spider body to enable the animations (walking, idle and attack animations were manually created in 1 s time slots that contain 30 movement frames). The Blender material was saved as texture and compiled into Filmbox (.fbx) files to make them usable in Unity3D. For the study, the AR app was installed on Samsung smartphones (Galaxy S8, Exynos 8895, 6.20″, 64 GB, resolution: 2960 × 1440 px, memory: 4 GB) running Android 8.0. With the installation of the app also ARCore (version 1.10) was installed to run the augmented reality elements, access to the camera was allowed and the volume for the sound effects was set. After that, no further setting changes or internet connection were necessary to use the app and all data was stored locally.

Participants in the current study were requested to train 6 × 30 min (always starting from the first level, irrespective of achieved level) over a two-weeks span on any chosen day with the single restriction to train only once a day. For each training, date, time and ratings were logged automatically and locally on the smartphones and assigned to the participant number.

2.5. Outcomes

The primary outcome measure, as defined in the protocol, was subjective fear (SUDS, range 0–10) in the BAT with a real-life spider. Our secondary outcomes were performance (range 0–12) as well as the subjective disgust (SUDS, range 0–10) in the BAT, the Fear of Spiders Questionnaire (FSQ Szymanski & O'Donohue, 1995, German version Rinck et al., 2002), the Spider Phobia Beliefs Questionnaire (SBQ Arntz et al., 1993, German version Pössel & Hautzinger, 2003), and one question to assess self-reported reduction of fear of spiders.

The FSQ is a self-report questionnaire and measures avoidance behaviour as well as fear of harm. It consists of 18 items on a 7-point scale (0 = not at all true to 6 = very true, α = 0.96 (Rinck et al., 2002), range 0–108) with higher scores indicating greater severity (Rinck et al., 2002; Szymanski & O'Donohue, 1995). The SBQ is a self-report questionnaire and specifically assesses spider-related catastrophic cognitions. It consists of 48 items (0 % = I do not believe that at all to 100 % = I am absolutely convinced, α = 0.98 (Pössel & Hautzinger, 2003)), separated into spider-related and self-related beliefs, with higher scores indicating greater severity (Arntz et al., 1993; Pössel & Hautzinger, 2003). In the *self-reported reduction of fear* item, participants were asked to self-rate their subjective reduction in fear of spiders in daily-life on a single scale in a range of 0 to 10 (0 = not at all and 10 = a lot).

Other outcomes of interest were *self-reported reduction of disgust* of spiders, the General Self-Efficacy Scale (GSE, German version (Schwarzer & Jerusalem, 1995, 1999)), a credibility/expectancy for improvement scale (Borkovec & Nau, 1972), a scale for usability of the app, and the Augmented Reality Immersion Questionnaire (ARI (Georgiou & Kyza, 2017)).

In the *self-reported reduction of disgust* item, participants were asked to self-rate their subjective reduction in disgust of spiders in daily-life on a single scale in a range of 0 to 10 (0 = not at all and 10 = a lot).

The GSE (Schwarzer & Jerusalem, 1995, 1999) consists of 10 items on a 4-point scale (1 = not at all to 4 = very much, α = 0.76 Schwarzer & Jerusalem, 1999), range 10–40) assessing the general self-efficacy within unknown, difficult situations.

The credibility/expectancy for improvement scale (Borkovec & Nau, 1972) consists of five items on the expectations for treatment improvement on an 11-point scale (0 = not at all to 10 = very much, α = 0.81 (Borkovec & Nau, 1972), range 0–50) and was translated into German and adapted to our app.

The scale for usability was specifically created for the purpose of this study and the app and consisted of eight items on a 11-point scale (0 = not at all to 10 = very much, range 0–80) regarding e.g. its functionality and design, and four open format questions for general feedback (see supplementary materials).

The ARI (Georgiou & Kyza, 2017) is a self-report questionnaire consisting of 21 items on a 7-point scale (0 = not at all to 6 = very much, α = 0.90 (Georgiou & Kyza, 2017), range 0–126) regarding the participants' immersive experience with the AR app, which we translated into German and adapted to our app.

2.6. Statistical analyses

We used linear models in combination with ANOVA (SS II). The analyses were done in R (version 3.6.2, GUI 1.70 (R Development Core Team, 2012)). Dependent variables were our pre-defined primary, secondary and other outcome measures, each investigated in a separate model. The independent variable was the between-subject factor group (intervention or control). As per protocol, the corresponding measures from the baseline were separately included as covariates to account for potential baseline differences. Further covariates were sex, age, and diagnosis (clinical/subclinical). In case of significant interactions of covariates with the factor group, post-hoc tests were applied to describe the interaction. In case of no significant interactions of covariates with

the factor group, the two-fold interactions were removed from the statistical models. Other outcomes of interest were each analysed in an explorative manner and only nominal p values are reported if applicable.

We present results as means (SDs) for the intervention and control group with associated two-sided p values, as well as adjusted group mean differences with 95 % CIs. $P < 0.05$ was considered significant for the primary outcome. For our five secondary outcomes, we set the significance threshold to $p < 0.01$ (Bonferroni correction for five independent tests). Outcomes of further interest were analysed in an explorative way (see appendix).

Usually the steps in the BAT are considered and analysed as a continuous variable (Botella et al., 2016; Lass-Hennemann & Michael, 2014). However, since the continuity of this test is arguable, we additionally performed a Kruskal Wallis test.

We estimated Cohen's *d* as effect size measurement. The estimate of *d* was based on the t value of the linear models. Therefore, *d* is corrected for the effects of all confounding variables included in the linear model. By convention, $d = 0.2$ is considered to be a small, $d = 0.5$ a medium and $d = 0.8$ a large effect (Cohen, 1992).

According to previous AR exposure studies to treat fear of small animals including spiders (Botella et al., 2016), we expected large effect sizes. Based on a power analysis using ANOVA with repeated measurements ($r = 0.5$) and between factors assuming to detect a large effect size ($f = 0.4$) with a power of 95 % and $\alpha = 0.05$ (software: G*Power 3.1) 32 participants in each of the 2 groups are needed resulting in 64 participants.

3. Results

Between October 7, 2019 and December 6, 2019, 71 individuals were screened for trial participation, of whom 5 were excluded after screening. Consequently, 66 individuals were enrolled, of whom 33 were randomly allocated to use the AR app (intervention group) and 33 were allocated to the control group. 66 participants (35 fulfilling DSM-5 criteria for spider phobia) completed the study as planned and were analysed (Fig. 2). Participants' baseline characteristics were balanced across groups (Table 1). Final data was collected on December 6, 2019. No dropouts or adverse events occurred.

Concurrent validity between baseline variables SUDS of fear in BAT, performance in BAT, SUDS disgust in BAT, FSQ, and SBQ were assessed using Spearman's rank correlation coefficient (r_s), which was found to be moderate with a range of $r_s \geq |0.34|$ and $r_s \leq |0.52|$ (all $p < 0.006$; all $n = 66$).

Test-retest reliability was determined by calculating intra-class correlation coefficients (ICCs) separately between Visit 1 and Visit 2 for all outcome measures (SUDS of fear in BAT, performance in BAT, SUDS of disgust in BAT, FSQ, SBQ) within the control group. ICC estimates were calculated using the psych-package (Revelle, 2021) in R based on single-rating, consistency, 2-way mixed models.

Because an exposure in a baseline measurement may affect the scores of the outcomes in a systematic way, we additionally calculated ICC estimated based on single-rating, consistency, 2-way random effect models. Here, we found a good degree of reliability between Visit 1 and Visit 2 SUDS of disgust in BAT (ICC = 0.75, $F(32,32) = 9.1$, $p < 0.0001$), moderate degree of reliability between Visit 1 and Visit 2 in BAT ratings (ICC = 0.72, $F(32,32) = 13.2$, $p < 0.0001$) and SBQ ratings (ICC = 0.51, $F(32,32) = 4.9$, $p < 0.0001$), and poor reliability between Visit 1 and Visit 2 SUDS of fear in BAT (ICC = 0.35, $F(32,32) = 3.2$, $p = 0.0008$), and FSQ (ICC = 0.28, $F(32,32) = 3.7$, $p = 0.0002$).

Uptake of the AR app: 28 (85 %) of 33 participants completed at least one session (30 min), 23 (70 %) at least 1 h, and 11 participants (33 %) at least 2 h of AR exposure. Five (15 %) had an exposure time under 30 min. The mean total time of app use was 91.12 min (SD 44.23).

Since all randomized participants completed the trial and no protocol violations were observed, the per protocol analysis was identical to an intention to treat (ITT) analysis. The analysis revealed that the repeated administration of the AR app led to significantly lower subjective fear in the BAT compared to the control group (intervention group, baseline: 7.12 [SD 2.03], follow-up: 5.03 [SD 2.19] vs. control group, baseline: 7.06 [SD 2.34], follow-up 6.24 [SD 2.21]; adjusted group difference -1.24, 95 % CI -2.17 to -0.31; Cohen's d = 0.57, p = 0.010). Further, the steps reached in the real-life BAT were significantly higher in the intervention group compared to the control group (intervention group, baseline: 5.27 [SD 2.32], follow-up: 6.76 [SD 2.40] vs. control group, baseline: 4.97 [SD 2.52], follow-up: 5.42 [SD 2.67]; adjusted group difference 1.05, 95 % CI 0.46 to 1.64; Cohen's d = 0.41, p = 0.00068). Similar results were found for the subjective ratings of disgust, the questionnaires and participants' perceived reduction of fear. There was a significant reduction in subjective disgust in the BAT of the

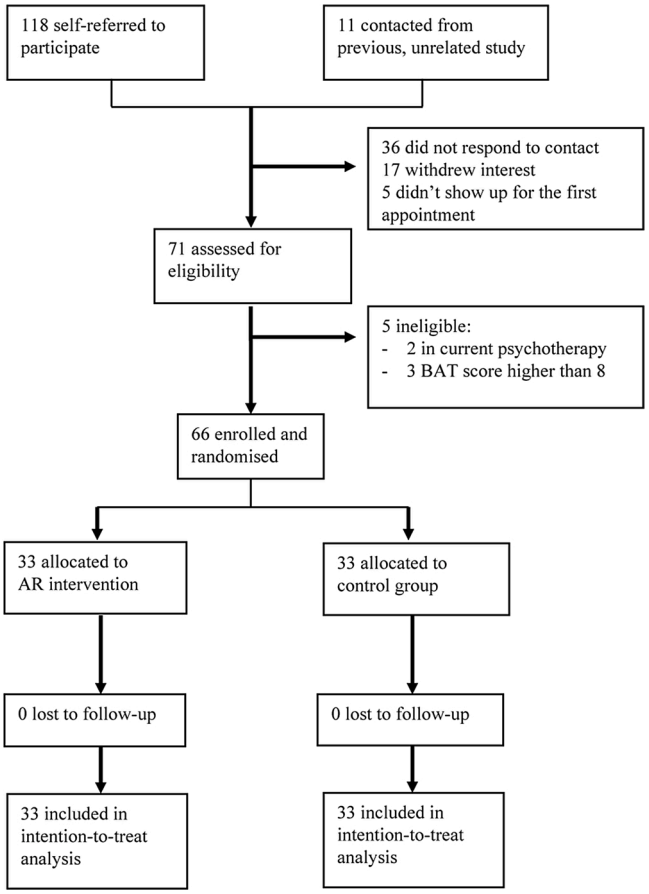

Fig. 2. Trial profile. BAT = Behavioural Approach Test.

Table 1
Baseline characteristics.

	Intervention (n = 33)	Control (n = 33)
Sex		
Male	7 (21 %)	7 (21 %)
Female	26 (79 %)	26 (79 %)
Age (years)	24 (18–39)	24.5 (18–39)
Ethnic origin		
Caucasian	28 (85 %)	29 (88 %)
Asian	3 (9%)	1 (3%)
Other	2 (6%)	3 (9%)
Diagnosis of arachnophobia	18	17
Education		
Highschool/vocational education	7 (21 %)	4 (12 %)
College/university	26 (79 %)	28 (85 %)
Other	0	1 (3%)

intervention group compared to the control group (intervention group, baseline: 7.18 [SD 2.80], follow-up: 6.09 [SD 2.60] vs. control group, baseline: 7.06 [SD 2.76], follow-up 7.03 [SD 2.50]; adjusted group difference -1.03, 95 % CI -1.80 to -0.26; Cohen's d = 0.41, p = 0.0098). The reductions of fear and disgust in in the BAT were significantly correlated in the intervention group (r_s = 0.53, p < 0.0001, spearman correlation of deltas) but not in the control group. A reduction of fear in the intervention group compared to the control group was further shown by corresponding questionnaires (FSQ: intervention group, baseline: 70.24 [SD 19.46], follow-up: 42.18 [SD 19.75] vs. control group, baseline: 68.15 [SD 17.73], follow-up 65.12 [SD 19.15]; adjusted group difference -24.42, 95 % CI -31.60 to -17.24; Cohen's d = 1.30, p < 0.0001; SBQ: intervention group, baseline: 54.53 [SD 20.33], follow-up: 38.81 [SD 19.59] vs. control group, baseline: 55.16 [SD 16.94], follow-up 54.99 [SD 17.98]; adjusted group difference -15.78, 95 % CI -22 to -9.55; Cohen's d = 0.85, p < 0.0001) and participants' subjectively perceived reduction of fear (intervention group: 3.76 [SD 2.51] vs. control group: 1.03 [SD 2.07], adjusted group difference 2.73, 95 % CI 1.60 to 3.85; Cohen's d = 1.20, p < 0.0001) (Table 2).

All intervention effects were independent of the presence of a DSM-5 diagnosis (p ≥ 0.5 for interactions between diagnosis and intervention group) for the primary and secondary outcome measures. No significant interactions were found between the factor group and the covariates age and sex for the primary outcome (p ≥ 0.7) and all secondary outcomes (p ≥ 0.2, corrected for multiple comparison), except for the SBQ, where we found a significant interaction with sex (p = 0.0070). Post-hoc analyses indicated a reduction of the SBQ score of the intervention group compared to the control group for women (intervention group, baseline: 58.18 [SD 16.21], follow-up: 38.48 [SD 17.93] vs. control group, baseline: 55.34 [SD 14.61], follow-up 56.53 [SD 15.79]; adjusted group differences -20.20, 95 % CI -27.28 to -13.11; Cohen's d = 1.24; p < 0.0001), but not for men (p = 0.48).

The analyses additionally showed a significant effect of subjectively perceived reduction of disgust between the intervention and control group (intervention group, follow-up: 2.42 [SD 2.42] vs. control group, follow-up 0.52 [SD 1.37]; adjusted group difference 1.91, 95 % CI 0.93 to 2.90; Cohen's d = 0.98, p = 0.00025). Treatment did not affect general self-efficacy, GSE (p = 0.83).

We looked at the credibility/expectancy for improvement, the usability and the immersion in a descriptive manner, to gain feedback from the intervention group. Analyses show an average credibility/expectancy for improvement of 35.49 (SD 5.97, range 21–46). The overall acceptability of the AR app was very good, and its usability, design and functionality were rated as very appealing with an average value of 51 (SD 16.47, range 7–80). The participants showed an average immersion value of 76.85 (SD 19.07, range 17–90).

In the literature, the BAT is widely considered and analyzed as a continuous variable. In addition to the parametric tests, we also report results from a non-parametric Kruskal Wallis test with eta^2 as effect size showing that the previously reported effects of the treatment for the performance in BAT remained significant (H(1) = 4.56, p = 0.033, eta^2 = 0.056).

Table 2
Outcome measures at both timepoints and differences between groups. Data are mean (SD), unless otherwise indicated. SUDS = Subjective Units of Distress Scale. BAT = Behavioural Approach Test. FSQ = Fear of Spiders Questionnaire. SBQ = Spider Beliefs Questionnaire.

	Intervention (n = 33)	Control (n = 33)	Adjusted group difference (95 % CI)	Effect size (Cohen's d)	p value
Primary outcome					
SUDS of fear in BAT					
Baseline	7.12 (SD 2.03)	7.06 (SD 2.34)	..	..	..
Follow-Up	5.03 (SD 2.19)	6.24 (SD 2.21)	−1.24 (-2.17 to -0.31)	0.57	0.010
Secondary outcomes					
Performance in BAT					
Baseline	5.27 (SD 2.32)	4.97 (SD 2.52)	..	..	..
Follow-Up	6.76 (SD 2.40)	5.42 (SD 2.67)	1.05 (0.46 to 1.64)	0.41	0.00068
SUDS of disgust in BAT					
Baseline	7.18 (SD 2.80)	7.06 (SD 2.76)	..	..	..
Follow-Up	6.09 (SD 2.60)	7.03 (SD 2.50)	−1.03 (-1.80 to -0.26)	0.41	0.0098
FSQ					
Baseline	70.24 (SD 19.46)	68.15 (SD 17.73)	..	..	..
Follow-Up	42.18 (SD 19.75)	65.12 (SD 19.15)	−24.42 (-31.60 to -17.24)	1.30	<0.0001
SBQ					
Baseline	54.53 (SD 20.33)	55.16 (SD 16.94)	..	..	..
Follow-Up	38.81 (SD 19.59)	54.99 (SD 17.98)	−15.78 (-22 to -9.55)	0.85	<0.0001
Reduction of fear					
Baseline	..	..	..	..	..
Follow-Up	3.76 (SD 2.51)	1.03 (SD 2.07)	2.73 (1.60 to 3.85)	1.20	<0.0001

4. Discussion

We report that repeated home-use of the stand-alone, smartphone-based, gamified AR exposure app was effective in the reduction of phobic fear in participants with fear of spiders. Specifically, the app use led to reductions in fear, disgust and avoidance behaviour at medium effect sizes when tested in a real-life situation, and to reductions at large effect sizes in questionnaire-based fear measures.

Studies with in vivo exposure therapy or desktop-based AR exposure treatments typically report large effect sizes. The AR exposure treatments in those studies were carried out under laboratory conditions and continuous surveillance and guidance by an experimenter or clinician (Botella et al., 2016; Chicchi Giglioli et al., 2015; Suso-Ribera et al., 2019). In our study, participants carried out the treatment by themselves in their homes. This unsupervised form of treatment resulted in individual differences in the actual exposure time. The reasons for less exposure may have been compliance issues or technical challenges. The latter were mainly due to issues of a correct surface detection by the app. Despite these challenges and based on the participants' feedbacks in the usability questionnaire, the overall acceptability of the AR app was very good, and its usability, design and functionality were rated as very appealing by the participants. It is noteworthy that even with an average exposure time of approx. 90 min instead of the suggested 180 min, the fear of spiders was reduced at clinically relevant effect sizes, although we can of course not exclude that the placebo effect contributed to the observed effects. Importantly, we showed the benefits of our intervention in a real-life spider situation on subjective fear and disgust as well as on the objectively measurable behavioural level (BAT). Finally, we

found that treatment effects were independent of the presence of a DSM-5 diagnosis of specific phobia, indicating the app to be a beneficial intervention for both subclinical and clinical fear of spiders.

Our study has several limitations. First, it was framed as a smartphone-based intervention to treat fear of spiders. This might have led to a selection bias of participants willing to use modern technologies for treatment purposes, potentially reducing the generalizability of the findings. Second, we only included participants aged 18–40, reducing the generalization of the findings to the older generation. Third, from the current data we do not know whether treatment effects outlast the six weeks we assessed. Fourth, we only tested one intervention regime of 6×30 min over a two-week period. We do not know, if other treatment regimes would have led to other results. Fifth, please note that the test-retest reliability was lower for SUDS fear in BAT than for SUDS disgust in BAT and BAT performance. Sixth, as we did not conduct a full diagnostic interview, we cannot make any statement on the possible influences of other comorbid specific phobias. Last, we did not include a direct comparison to other evidence-based interventions.

Even though our understanding of the underlying mechanisms of exposure has evolved and therapy protocols are constantly being improved through strategies targeting the cognitive aspects of fear (i.e. inhibitory learning, violation of expectancy, dysfunctional beliefs, self-efficacy), there is still a number of individuals experiencing a return of fear (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; Tardif, Therrien, & Bouchard, 2019). We see a great potential in our smartphone-based, stand-alone AR exposure app to act as a complementary tool for psychotherapy and especially as a re-booster or re-fresher of the learned associations and strategies to cope with fear of spiders in real-life.

Recent literature emphasized the role of disgust in small animal phobia, especially in spider phobia (next to blood-injury-injection type and obsessive-compulsive disorder, OCD). It has been discussed, that the emotion of disgust might be much more treatment resistant than fear in the context of exposure-based intervention and might even be increased (Knowles, Jessup, & Olatunji, 2018). Our results show a similar reduction of fear and disgust for the intervention group, but interestingly not for the control group, here we see a slight decrease in fear but almost none in disgust. This observation needs to be further addressed, as well as the potential of the here developed app to target and treat disgust of spiders.

5. Conclusion

Given the current underuse of conventional in vivo exposure therapy for specific phobias, there is definitely a need for alternative evidence-based approaches. Smartphone-based interventions implementing AR technology for exposure purposes have the potential to become a game changer for the current dissemination problem of in vivo treatments. Apps are highly accessible due to the widespread use of smartphones in the general population. Furthermore, digital marketplaces are already in place to enable the dissemination of apps to practitioners or, as self-help tools, directly to patients. Smartphone-based exposure has all the benefits mentioned for stationary AR and, in addition, it can be conducted both in the treatment rooms without cost-expensive gear and as a stand-alone add-on for homework in between sessions (blended treatment). Finally, the here presented beneficial effects of the gamified AR app are likely to encourage people to face their fears in a subtle and fun, yet effective way.

Contributions

AZ and DQ designed the exposure app and the trial, and drafted the paper. NW and MKI programmed and visually designed the app. AZ collected the data. DB contributed clinical advice. AP and TM commented on the design of the exposure app and trial. AZ, BF and NSS analysed the data. All authors commented on the paper.

Funding

The study was funded by the Swiss National Science Foundation (SNSF), from which Anja Zimmer received a Doc.CH grant for the course of four years during her PhD (P0BSP1_181761 / 1), and by the Trans-faculty Research Platform of theUniversity of Basel, Switzerland. The SNSF had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Data availability

The data (de-identified) that support the findings of this study are available on request from the corresponding author [DQ].

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

We thank Amanda Aerni for monitoring the study, Fabian Müller for designing sound effects for the app, Milena Cavegn, Johanna Weibel, Jule Schröder, Daria Bühler, Nina Waldthaler, Johanna Otte, Galya Iseli and Irena Kovacic for helping in trial conduction.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.janxdis.2021.102442.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th edn.). Washington: American Psychiatric Association.

Arntz, A., Lavy, E., Van den Berg, G., & Van Rijsoort, S. (1993). Negative beliefs of spider phobics: A psychometric evaluation of the Spider Phobia Beliefs Questionnaire. *Advances in Behaviour Research and Therapy, 15*, 257–277. https://doi.org/10.1016/0146-6402(93)90012-Q

Azuma, R. T. (1997). A survey of augmented reality. *Presence, 6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Bandelow, B., & Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience, 17*(3), 327.

Baus, O., & Bouchard, S. (2014). Moving from virtual reality exposure-based therapy to augmented reality exposure-based therapy: A review. *Frontiers in Human Neuroscience, 8*, 112. https://doi.org/10.3389/fnhum.2014.00112

Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories–IA and –II in psychiatric outpatients. *Journal of Personality Assessment, 67*, 588–597.

Borkovec, T. D., & Nau, S. D. (1972). Credibility of analogue therapy rationales. *Journal of Behaviour Therapy and Experimental Psychiatry, 3*(4), 257–260. https://doi.org/10.1016/0005-7916(72)90045-6

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent progress in virtual reality exposure therapy for phobias: A systematic review. *Current Psychiatry Reports, 19*, 42. https://doi.org/10.1007/s11920-017-0788-4

Botella, C., Pérez-Ara, MÁ, Breton-Lopez, J., Quero, S., Garcia-Palacios, A., & Baños, R. M. (2016). In vivo versus augmented reality exposure in the treatment of small animal phobia: A randomized controlled trial. *PloS One, 11*(2). https://doi.org/10.1371/journal.pone.0148237. e0148237.

Carl, E., Stein, A., Levihn-Coon, A., et al. (2018). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. *Journal of Anxiety Disorders, 61*, 27–36. https://doi.org/10.1016/j.janxdis.2018.08.003

Chambless, D., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685–716.

Chicchi Giglioli, I. A., Pallavicini, F., Pedroli, E., Serino, S., & Riva, G. (2015). Augmented reality: A brand new challenge for the assessment and treatment of psychological disorders. *Computational and Mathematical Methods in Medicine*. https://doi.org/10.1155/2015/862942

Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical Psychology Review, 27*, 266–286. https://doi.org/10.1016/j.cpr.2006.10.002

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Cook, J., Biyanova, T., Elhai, J., Schnurr, P., & Coyne, J. (2010). What do psychotherapists really do in practice? An Internet study of over 2,000 practitioners.

Psychotherapy Theory Research Practice Training, 47(260). https://doi.org/10.1037/a0019788

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23.

Davey, G. (2011). Disgust: the disease-avoidance emotion and its dysfunctions. *Philosophical Transactions Biological Sciences, 366*(1583), 3453–3465. https://doi.org/10.1098/rstb.2011.0039

De Witte, N. A., Scheveneels, S., Sels, R., Debard, G., Hermans, D., & Van Daele, T. (2020). Augmenting exposure therapy: Mobile augmented reality for specific phobia. *Frontiers in Virtual Reality, 1*(8). https://doi.org/10.3389/frvir.2020.00008

Eaton, W. W., Bienvenu, O. J., & Miloyan, B. (2018). Specific phobias. *The Lancet Psychiatry, 5*(8), 678–686. https://doi.org/10.1016/S2215-0366(18)30169-X

Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. In vivo exposure by patients with specific phobias. *Cyberpsychology & Behavior, 10*(5), 722–724. https://doi.org/10.1089/cpb.2007.9962

Georgiou, Y., & Kyza, E. A. (2017). The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings. *International Journal of Human-Computer Studies, 98*, 24–37. https://doi.org/10.1016/j.ijhcs.2016.09.014

Knowles, K. A., Jessup, S. C., & Olatunji, B. O. (2018). Disgust in anxiety and obsessive-compulsive disorders: Recent findings and future directions. *Current Psychiatry Reports, 20*(9), 1–10.

Lass-Hennemann, J., & Michael, T. (2014). Endogenous cortisol levels influence spider phobia. *Behaviour Research and Therapy, 60*, 39–45. https://doi.org/10.1016/j.brat.2014.06.009

Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour Research and Therapy, 74*, 18–24. https://doi.org/10.1016/j.brat.2015.08.010

Oosterink, F., De Jongh, A., & Hoogstraten, J. (2009). Prevalence of dental fear and phobia relative to other fear and phobia subtypes. *European Journal of Oral Sciences, 117*(2), 135–143. https://doi.org/10.1111/j.1600-0722.2008.00602.x

Öst, L. (1989). One-session treatment for specific phobias. *Behaviour Research and Therapy, 27*, 1–7. https://doi.org/10.1016/0005-7967(89)90113-7

Öst, L. G. (2012). One-session treatment for specific phobias in adult and children (2012). In T. E. Davis, T. H. Ollendick, & L. G. Öst (Eds.), *Intensive one-session treatment of specific phobias* (pp. 58–95). New York: Springer.

Pössel, P., & Hautzinger, M. (2003). Dysfunktionale Überzeugungen bei Spinnenangst. Eine deutsche Version des "Spider Phobia Beliefs Questionnaire". *Zeitschrift für Klinische Psychologie und Psychotherapie, 32*, 24–30. https://doi.org/10.1026/0084-5345.32.1.24

R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research*. R package version 2.1.3. Evanston, Illinois: Northwestern University https://CRAN.R-project.org/package=psych.

Rinck, M., Bundschuh, S., Engler, S., Müller, A., Wissmann, J., & Ellwart, T. (2002). Reliabilität und Validität dreier Instrumente zur Messung von Angst vor Spinnen. *Diagnostica, 48*, 141–149.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. Measures in health psychology: A user's portfolio. *Causal and control beliefs, 1*(1), 35–37.

Schwarzer, R., & Jerusalem, M. (1999). Skalen zur Erfassung von Lehrer-und schülermerkmalen. *Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin, 25*, 2014.

Segal, R., Bhatia, M., & Drapeau, M. (2011). Therapists' perception of benefits and costs of using virtual reality treatments. *Cyberpsychology, Behavior and Social Networking, 14*, 29–34. https://doi.org/10.1089/cyber.2009.0398

Stetina, B. U., Felnhofer, A., Kothgassner, O. D., & Lehenbauer, M. (2012). Games for health: Have fun with virtual reality! Virtual reality in psychological. *Medical and Pedagogical Applications*, 65–80. https://doi.org/10.5772/50677

Suso-Ribera, C., Fernández-Álvarez, J., García-Palacios, A., et al. (2019). Virtual reality, augmented reality, and in vivo exposure therapy: A preliminary comparison of treatment efficacy in small animal phobia. *Cyberpsychology, Behavior and Social Networking, 22*(1), 31–38. https://doi.org/10.1089/cyber.2017.0672

Szymanski, J., & O'Donohue, W. (1995). Fear of spiders questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 26*, 31–34. https://doi.org/10.1016/0005-7916(94)00072-T

Tardif, N., Therrien, C. E., & Bouchard, S. (2019). Re-examining psychological mechanisms underlying virtual reality-based exposure for spider phobia. *Cyberpsychology, Behavior and Social Networking, 22*(1), 39–45.

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 exposure in agoraphobia, specific phobia, and social phobia. *Frontiers in Psychology, 10*, 1758.

Wolitzky-Taylor, K., Horowitz, J., Powers, M., & Telch, M. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical Psychology Review, 28*, 1021–1037. https://doi.org/10.1016/j.cpr.2008.02.007

5 Discussion

The currently available consumer hardware allows us to display immersive technologies with as little as a headset and our smartphones, offering accessible and scalable solutions to address mental health issues. This thesis gave an overview of the current clinical application of smartphone apps enriched through immersive technologies to facilitate the dissemination of exposure-based interventions for specific phobias. In the frame of this thesis, three randomized controlled clinical trials highlight the potential and challenges for smartphone apps with VR and AR as well as specific elements to enhance practicability, usability and engagement such as realistic 360° images and videos, gaze selection for navigation and gamification elements.

In study **A)** we tested a stand-alone, automated and gamified VR exposure training for fear of heights. The training with three realistic scenarios with 360° images followed a predefined exposure scheme, thus no further cognitive elements or guidance were integrated. The results show a significant reduction of fear of heights in self-reported measures already after a single 1-hour session with the app *EasyHeights* compared to the control group. After repeated use in a 2-week home training (6 x 30 minutes), the results indicate a significant reduction of subjective fear and additionally on a behavioral level with reduced avoidance behavior in a real-life situation on a tower. The app *EasyHeights* therefore offers an effective self-help tool to train exposure to heights. As an app-based intervention, it can be easily integrated into the daily life of those affected.

In study **B)** a similarly built stand-alone VR training for fear of public speaking was investigated. *Fearless Speech* was not only designed to reduce fear in a speech situation but additionally improving eye contact behavior. This adds a unique feature to the exposure program and takes advantage of the technology of gaze control in mobile VR. Again, no further guidance or cognitive elements were integrated. The automated, gamified exposure to three virtual scenarios with 360° videos of a virtual audience lead to a significant reduction of fear and an increase in eye contact after repeated use in the 2-weeks home training (9 x 20 minutes), but not after acute use in the single 1-hour session. The effects of the home training were again additionally tested in a real-life speech situation, showing significant results for fear reduction and improved eye contact. The results of the study and the app expands current literature on social anxiety in a sense that the app was specifically designed as an eye contact training which led to a significant reduction of fear. *Fearless Speech* offers a unique exposure program that omits the social interaction during the treatment. This aspect may display a key feature for socially anxious people to not seek professional help, since the contact to a therapist is already a situation leading to overwhelming feelings and anxiety and may amplify avoidance behavior or therapy dropout.

In study **C)** AR was explored to reduce fear and avoidance behavior in fear of spiders. As the first app that implements an automated, gamified stand-alone AR exposure training with no further guidance or cognitive elements, this study especially marks the advancements of technology as compared to previous studies. No markers where needed for the surface detection and only a smartphone was necessary. Results showed a significant reduction of fear, disgust and avoidance behavior both on a subjective level measured through questionnaires and on a behavioral level in a real-life situation after the 2-weeks home

training (6 x 30 minutes) with the app *Phobys*. The app thus offers an effective self-help tool, which does not require further equipment besides the smartphone.

All three apps of these randomized controlled trials are offering innovative and scalable solutions to the dissemination problem not only of exposure-based interventions in general but of immersive technology-based treatment options for specific phobias.

Using 360° images of height situations and 360° videos of a virtual audience for the speech situations ads to the level of realism in VR, which is promising for complex situational fears e.g. fear of flying or certain natural environments such as fear of open water or storms. Using realistic visual material instead of CGI is also especially crucial in social situations to possibly avoid the uncanny valley effect (Mori, 1970). However, this would need further investigation and comparison to other VR scenarios with CGI. For mobile AR, a huge advantage and well-suited element to integrate to the treatment of specific phobias for animals and objects is the interaction with the AR model whilst seeing one's own body. Again, it needs further investigation and comparison to other treatment regimens and in vivo. Nevertheless, interaction with the AR models offers a unique feature, especially for phobias where disgust is evoked alongside fear and avoidance, e.g. with spiders, insects, blood or vomit (Davey, 2011). The research of the past years has strongly focused on creating realistic VR, whereas the future lies at the intersection of technology and real life, creating high hopes for AR (Cipresso et al., 2018; Lindner, 2021; Riva, 2022).

A challenge for mental health apps in general is compliance and adherence (Wilhelm et al., 2020). It has been proposed that gamification elements could promote treatment engagement, leading to more compliance and adherence (Nixon & Howard, 2013). However, they should serve as a motivator to increase task engagement and for specific phobias even overcome avoidance behavior, but not distract the individual from learning that the feared stimuli are not threatening (Hoffman & Chu, 2018). Our three apps all have gamification elements such as levels, points and rewards integrated that could enhance the engagement and adherence. Although stand-alone and automated treatment options such as our investigated apps have their own appeal and benefits, apps that include guidance (i.e., from a therapist or coach) have generally been found to have larger effects than stand-alone interventions (Baumeister et al., 2021; Erbe et al., 2017; Wu et al., 2021). Given that stand-alone mental health apps require fewer resources than guided apps, identifying factors that might increase engagement and effectiveness without therapist involvement could greatly improve dissemination efforts (Wu et al., 2021). Increasing the number of engagement techniques may be beneficial also to our apps, to maximize their impact. Incorporating features to enhance a social community and connect users such as chats, a buddy-system or posts of successes may be the most likely way to engage those affected without the use of therapists or coaches. Furthermore, even though our results indicate that the element of exposure alone is sufficient to reduce phobic symptoms, the scalability of app-based interventions allows to be improved through strategies targeting the cognitive aspects of fear, e.g. inhibitory learning, violation of expectancy, reduction of dysfunctional beliefs and the increase of self-efficacy (Craske et al., 2014; Tardif et al., 2019).

As a number of individuals experience a return of fear even after successful exposure interventions (Eaton et al., 2018; Mystkowski et al., 2002), there is great potential in our smartphone-based, stand-alone VR/AR exposure apps to act as a complementary tool for psychotherapy and

especially as a re-booster or re-fresher of the learned associations and strategies to cope with specific phobias. In addition, they allow the users to adapt the training to their individual needs.

The duration of the VR home training sessions was based on the suggested amount of time to spend in a VR scenario to avoid cybersickness and leaned on previous automated VR exposure studies. The duration of the whole training was based on the maximum amount of time indicated for an intensive one-session exposure (Davis et al., 2012). The duration of the AR home training was leaned on our VR interventions – though the limitations of cybersickness are nonexistent – and the intensive one-session protocols. However, to investigate this time as the right dosage, the training needs to be compared to other treatment regimes. Furthermore, individuals with mental health issues usually need to use app for a long time to realize the full benefit and avoid relapse of mental health disorders. Research has not yet been able to show the long-time effectiveness of mental health apps (So et al., 2013; Torous et al., 2018; Wu et al., 2021). Therefore, there is still a lack of evidence concerning app effectiveness on the long term. Our studies show effects of the app use after 4-6 weeks follow-ups, however, the long-term effectiveness would have to be investigated further.

It is a crucial element to investigate the transfer of the effects in the virtual environment to real-life. This is also where our studies set themselves apart from previous studies on automated VR and AR interventions, usually only conducting self-reported or questionnaire-based measures. In all three studies, the effectiveness of the app-based interventions was shown in the self-reports and additionally in real-life settings, exposing the participants to real heights, a real speech situation and a real spider. Although, we tested the effectiveness in real-life settings the translation to clinical practice needs to be further investigated.

For all studies is has to be mentioned, that participants were likely to be more interested in technology, therefore adding the limitation of a selection bias questioning the generalizability of our results. In the studies for fear of heights and spiders, the population consisted of participants with clinical and subclinical fear and the results were not dependent on the clinical status. For the study on fear of public speaking we had a healthy population with subclinical PSA. Therefore, it needs to be further investigated if the intervention also shows significant effects in a clinical population.

Even though our studies could show significant fear reductions after only a 2-weeks home training with our stand-alone apps, the full clinical effect may come from lowering the threshold for individuals to subsequently encourage to expose themselves in everyday life situations, breaking the vicious circle of avoidance (Hofmann & Hay, 2018) or even considering an in vivo exposure therapy to strengthen and deepen the experience (Lindner et al., 2021a).

It has now been over 25 years since the first mental health applications of VR technology appeared (Riva, 2022) and much has happened since, both in terms of scientific progression and technological advances. The release of consumer VR platforms constitutes a true paradigm shift in the development and dissemination of clinical VR, which continues to inspire new generations of interventions (Lindner, 2021). The rapid and ongoing development of AR systems presents the opportunity to harness the technical advances of VR, while overcoming many of its limitations. Overall, it may be said that the use of ARET has the potential of going well beyond the treatment of small animal phobia (Baus & Bouchard,

2014), providing novel treatment strategies for substance use disorders (Brandon et al., 2021) and OCD (Garcia-Batista et al., 2021).

However, one technology is not simply better than the other because each technology has different strengths and weaknesses. The choice of which technology to apply depends on the scope of the application. Additionally, VR and AR are still developing technologies, posing limitations, challenges and changes for research and clinical practice (Lindner, 2021; Riva, 2022). Although mobile VR and AR interventions offer possible solutions of the dissemination problem and multiple trials and reviews have demonstrated their effectiveness, they have yet to become widespread beyond the research setting. Currently there is little translation into the ordinary clinical setting (Hilty et al., 2020; Mishkind et al., 2017).

Staying current with new technologies in general and mental health apps in specific is a significant challenge for institutions, psychotherapists and psychological societies, requiring a new approach to teaching and clinical supervision to provide new knowledge and establish skills (Hilty et al., 2020). Guidelines and evaluation frameworks are being developed for mental health apps (Lagan et al., 2020, 2021), though their diversity makes it difficult to select an appropriate framework to choose the right app (Chan et al., 2015; Torous et al., 2019). Additionally, they do not necessarily simplify the process of finding a mental health app and still require a significant effort on the part of the user, who must then review the products they have identified (Neary & Schueller, 2018). Although, people value ratings and peer reviews, consumer ratings do not seem to reflect clinical usefulness and show only moderate correlations with objective app quality rating scales (Lagan et al., 2020, 2021). While academics and clinicians aim to develop the most effective treatments, companies often focus more on the business aspects of product development and consumption. The lack of clear definitions and standards is harmful – not only to the field, where progress is impeded, but also to individuals with mental health issues, who may not know which app to trust (Ng et al., 2019). As it is more convenient to focus on those apps appearing early in search results with high ratings, educating patients about mental health apps is an important and evolving role for psychotherapists.

The development of effective mental health apps relies on the interdisciplinary integration of the theoretical knowledge from academia, the technological know-how from industries and the appropriate implementation from clinicians and psychotherapists (Boeldt et al., 2019). Even though the hardware is often available at low cost, the software needs to be customized for each experiment or individual, which requires huge efforts in terms of development as well as time to design multiple environments and functionalities. Additionally, the time required to conduct an RCT and publish results does not align with the rapid development of apps. Trials often have extensive eligibility criteria, which slows recruitment and decreases generalizability (Neary & Schueller, 2018). Research collaboration between the healthcare and software engineering fields could help to further our knowledge of app functionality and effectiveness. This is a real-life challenge associated with any app evaluation effort where no gold standards exist. Therefore, collaborations between clinical research and software engineering as well as user-friendly app design – as we did in our app development and studies – is highly desirable. To facilitate the integration of mental health apps into clinical practice and enhance engagement, it is critical to involve clinicians,

psychotherapists and care takers as well as affected individuals into the development process (Ng et al., 2019; Torous et al., 2019; Wilhelm et al., 2020).

To conclude, mental health apps with immersive technologies hold a particular promise to reduce barriers of dissemination of exposure-based interventions for specific phobias. Their use as effective stand-alone tools for self-help could reduce stigmatization and reach individuals not seeking professional help out of avoidance. Even though their implementation into clinical practice needs to be further investigated, they provide scalable solutions in blended care settings. We find ourselves in interesting yet crucial times to take advantage of the technologies available to make mental health care more accessible, engaging, regulated and evidence-based. This field requires the combined mobilization and collaboration of researchers, clinicians, psychotherapists, engineers, designers and users as well as policy makers and investors to advance the development and implementation of innovative psychotherapies, evolving the way we assess, monitor and treat mental health issues.

6 References

Agras, S., Sylvester, D., & Oliveau, D. (1969). The epidemiology of common fears and phobia. *Comprehensive psychiatry*, *10*(2), 151-156.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th edn.). Washington: American Psychiatric Association.

American Psychological Association, Society of Clinical Psychology. (2017). What is exposure therapy? Clinical practice guideline. https://apa.org/ptsdguideline/patients-and-families/exposure-therapy.pdf

Allan R. (2016, Dec 7). *Pokémon GO usage statistics say it's the most popular mobile game in U.S. history.* Survey Monkey Intelligence. 2016. https://medium.com/@sm_app_intel/pok%C3%A9mon-go-usage-statistics-say-its-the-mostpopular-mobile-game-in-u-s-history-ea09ea2bf6df

Alonso, J., Liu, Z., Evans-Lacko, S., Sadikova, E., Sampson, N., Chatterji, S., Abdulmalik, J., Aguilar-Gaxiola, S., Al-Hamzawi, A., Andrade, L. H., Bruffaerts, R., Cardoso, G., Cia, A., Florescu, S., de Girolamo, G., Gureje, O., Haro, J. M., He, Y., de Jonge, P., … Thornicroft, G. (2018). Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries. *Depression and Anxiety*, *35*(3), 195–208. https://doi.org/10.1002/da.22711

Alqahtani, F., & Orji, R. (2020). Insights from user reviews to improve mental health apps. *Health Informatics Journal*, *26*(3), 2042–2066. https://doi.org/10.1177/1460458219896492

Alsina-Jurnet, I., Gutiérrez-Maldonado, J., & Rangel-Gómez, M. V. (2011). The role of presence in the level of anxiety experienced in clinical virtual environments. *Computers in Human Behavior*, *27*(1), 504–512. https://doi.org/10.1016/j.chb.2010.09.018

Azuma, R. T. (1997). A survey of augmented reality. *Presence, 6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advaces in augmented reality. Computers & Graphics, 25, 1–15.

Bandelow, B., & Michaelis, S. (2015). *Epidemiology of anxiety disorders in the 21st century*. *Dialogues in neuroscience.*

Bandelow, B., Lichte, T., Rudolf, S., Wiltink, J., & Beutel, M. (Eds.). (2014). *S3-Leitlinie Angststörungen*. Springer, Berlin, Heidelberg.

Baños, R. M., Botella, C., Garcia-Palacios, A., Villa, H., Perpiñá, C., & Alcaniz, M. (2000). Presence and reality judgment in virtual environments: a unitary construct? *CyberPsychology & Behavior*, *3*(3), 327-335.

Baños, R. M., Herrero, R., & Vara, M. D. (2022). What is the Current and Future Status of Digital Mental Health Interventions? *Spanish Journal of Psychology*, *25*(3). https://doi.org/10.1017/SJP.2022.2

Baranowski, T., & Lyons, E. J. (2020). Scoping Review of Pokémon Go: Comprehensive Assessment of Augmented Reality for Physical Activity Change. *Games for Health Journal 9*(2), 71–84. https://doi.org/10.1089/g4h.2019.0034

Baumeister, H., Bauereiss, N., Zarski, A. C., Braun, L., Buntrock, C., Hoherz, C., Idrees, A. R., Kraft, R., Meyer, P., Nguyen, T. B. D., Pryss, R., Reichert, M., Sextl, T., Steinhoff, M., Stenzel, L., Steubl, L., Terhorst, Y., Titzler, I., & Ebert, D. D. (2021). Clinical and Cost-Effectiveness of PSYCHOnlineTHERAPY: Study Protocol of a Multicenter Blended Outpatient Psychotherapy Cluster Randomized Controlled Trial for Patients With Depressive and Anxiety Disorders. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.660534

Baumel, A., Muench, F., Edan, S., & Kane, J. M. (2019). Objective user engagement with mental health apps: Systematic search and panel-based usage analysis. *Journal of Medical Internet Research*, *21*(9). https://doi.org/10.2196/14567

Baus, O., & Bouchard, S. (2014). Moving from virtual reality exposure-based therapy to augmented reality exposure-based therapy: a review. *Frontiers in human neuroscience*, *8*, 112. https://doi.org/10.3389/fnhum.2014.00112

Becker, C. B., Zayfert, C., & Anderson, E. (2004). A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behaviour Research and Therapy*, *42*(3), 277–292. https://doi.org/10.1016/S0005-7967(03)00138-4

Becker, E. M., & Jensen-Doss, A. (2013). Computer-assisted therapies: examination of therapist-level barriers to their use. *Behavior therapy*, *44*(4), 614-624.

Becker, E. S., Rinck, M., Türke, V., Kause, P., Goodwin, R., Neumer, S., & Margraf, J. (2007). Epidemiology of specific phobia subtypes: Findings from the Dresden Mental Health Study. *European Psychiatry*, *22*(2), 69–74. https://doi.org/10.1016/j.eurpsy.2006.09.006

Benito, K. G., & Walther, M. (2015). Therapeutic process during exposure: Habituation model. *Journal of Obsessive-Compulsive and Related Disorders*, *6*, 147–157. https://doi.org/10.1016/j.jocrd.2015.01.006

Bentz, D., Wang, N., Ibach, M. K., Schicktanz, N. S., Zimmer, A., Papassotiropoulos, A., & de Quervain, D. J. F. (2021). Effectiveness of a stand-alone, smartphone-based virtual reality exposure app to reduce fear of heights in real-life: a randomized trial. *Npj Digital Medicine*, *4*(1). https://doi.org/10.1038/s41746-021-00387-7

Bhugra, D., Tasman, A., Pathare, S., Priebe, S., Smith, S., Torous, J., Arbuckle, M. R., Langford, A., Alarcón, R. D., Chiu, H. F. K., First, M. B., Kay, J., Sunkel, C., Thapar, A., Udomratn, P., Baingana, F. K., Kestel, D., Ng, R. M. K., Patel, A., … Ventriglio, A. (2017). The WPA-Lancet Psychiatry Commission on the Future of Psychiatry. *The Lancet Psychiatry, 4*(10), 775–818. https://doi.org/10.1016/S2215-0366(17)30333-4

Boeldt, D., McMahon, E., McFaul, M., & Greenleaf, W. (2019). Using virtual reality exposure therapy to enhance treatment of anxiety disorders: Identifying areas of clinical adoption and potential obstacles. *Frontiers in Psychiatry, 10*, 773. https://doi.org/10.3389/fpsyt.2019.00773

Böhm, K., Förstner, U., Külz, A., & Voderholzer, U. (2008). Versorgungsrealität der Zwangsstörungen: Werden Expositionsverfahren eingesetzt? *Verhaltenstherapie*, *18*(1), 18–24.

Botella, C., Bretón-López, J., Quero, S., Baños, R., & García-Palacios, A. (2010). Treating cockroach phobia with augmented reality. *Behavior therapy*, *41*(3), 401-413.

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review. *Current Psychiatry Reports 19*(7). https://doi.org/10.1007/s11920-017-0788-4

Botella, C. M., Juan, M. C., Baños, R. M., Alcañiz, M., Guillén, V., & Rey, B. (2005). Mixing realities? An application of augmented reality for the treatment of cockroach phobia. *Cyberpsychology & behavior*, *8*(2), 162-171.

Botella, C., Pérez-Ara, M. Á., Bretón-López, J., Quero, S., García-Palacios, A., & Baños, R. M. (2016). In Vivo versus augmented reality exposure in the treatment of small animal phobia: A randomized controlled trial. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0148237

Boyd, J. H., D. S. Rae, J. W. Thompson, B. J. Burns, K. Bourdon, B. Z. Locke, and D. A. Regier. (1991). Phobia: prevalence and risk factors. *Social psychiatry and psychiatric epidemiology 25*(6), 314-323.

Brandon, K. O., Vinci, C., Kleinjan, M., Hernandez, L. M., Sawyer, L. E., Sutton, S. K., & Brandon, T. H. (2021). Testing Augmented Reality for Eliciting Cue-Provoked Urges to Smoke: Toward Moving Cue-Exposure into the Real World. *Nicotine and Tobacco Research*, *23*(5), 861–865. https://doi.org/10.1093/ntr/ntaa259

Broicher, T., Gerlach, A. L., & Neudeck, P. (2017). Die Relevanz der Ausbildung für den späteren Einsatz von Expositionsverfahren in der therapeutischen Praxis. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *46*(2), 107-116. https://doi.org/10.1026/1616-3443/a000415

Burleigh, T. J., Schoenherr, J. R., & Lacroix, G. L. (2013). Does the uncanny valley exist? An empirical test of the relationship between eeriness and the human likeness of digitally created faces. *Computers in Human Behavior*, *29*(3), 759–771. https://doi.org/10.1016/j.chb.2012.11.021

Castelvecchi, D. (2016). Low-cost headsets boost virtual reality's lab appeal. *Nature 533*, 153–154. doi: 10.1038/533153a

Chan, S., Torous, J., Hinton, L., & Yellowlees, P. (2015). Towards a Framework for Evaluating Mobile Mental Health Apps. *Telemedicine and E-Health*, *21*(12), 1038–1041. https://doi.org/10.1089/tmj.2015.0002

Chandrashekar, P. (2018). Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *MHealth*, *4*, 6–6. https://doi.org/10.21037/mhealth.2018.03.02

Chen, J., van den Bos, E., & Westenberg, P. M. (2020). A systematic review of visual avoidance of faces in socially anxious individuals: Influence of severity, type of social situation, and development. *Journal of Anxiety Disorders*, *70*, 102193.

Cheng, V. W. S., Davenport, T., Johnson, D., Vella, K., & Hickie, I. B. (2019). Gamification in apps and technologies for improving mental health and well-being: Systematic review. *JMIR Mental Health*, *6*(6). https://doi.org/10.2196/13717

Chicchi Giglioli, I. A., Pallavicini, F., Pedroli, E., Serino, S., & Riva, G. (2015). Augmented reality: a brand new challenge for the assessment and treatment of psychological disorders. *Computational and mathematical methods in medicine*, *2015*. https://doi.org/10.1155/2015/862942

Chou, P. H., Tseng, P. T., Wu, Y. C., Chang, J. P. C., Tu, Y. K., Stubbs, B., Carvalho, A. F., Lin, P. Y., Chen, Y. W., & Su, K. P. (2021). Efficacy and acceptability of different interventions for acrophobia: A network meta-analysis of randomised controlled trials. *Journal of Affective Disorders*, *282*, 786–794. https://doi.org/10.1016/j.jad.2020.12.172

Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical Psychology Review 27*(3), 266-286. https://doi.org/10.1016/j.cpr.2006.10.002

Christensen, H., Leach, L. S., Barney, L., Mackinnon, A. J., & Griffiths, K. M. (2006). The effect of web based depression interventions on self reported help seeking: Randomised controlled trial. *BMC Psychiatry*, *6*. https://doi.org/10.1186/1471-244X-6-13

Cipresso, P., Giglioli, I. A. C., Raya, M. A., & Riva, G. (2018). The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature. *Frontiers in Psychology*, 2086. https://doi.org/10.3389/fpsyg.2018.02086

Clay, R.A. (2021, Jan 1). *Mental health apps are gaining traction: Self-help apps are leading more people to therapy rather than replacing it, psychologists say*. American Psychological Association. www.apa.org/monitor/2021/01/trends-mental-health-apps

Cook, J. M., Biyanova, T., Elhai, J., Schnurr, P. P., & Coyne, J. C. (2010). What do psychotherapists really do in practice? An internet study of over 2,000 practitioners. *Psychotherapy*, *47*(2), 260–267. https://doi.org/10.1037/a0019788

Côté, S., & Bouchard, S. (2009). Cognitive mechanisms underlying virtual reality exposure. *Cyberpsychology and Behavior*, *12*(2), 121–129. https://doi.org/10.1089/cpb.2008.0008

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Curtis, G., Magee, W. J., Eaton, W. W., Wittchen, H. U., & Kessler, R. C. (1998). Specific fears and phobias: Epidemiology and classification. *The British Journal of Psychiatry*, *173*(3), 212-217.

Davey, G. (Ed.). (1997). *Phobias: A handbook of theory, research and treatment*. Wiley-Blackwell.

Davey, G. C. (2011). Disgust: the disease-avoidance emotion and its dysfunctions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1583), 3453-3465.

Davis, T. E., Ollendick, T. H., Reuther, E. T., & Munson, M. S. (2012). One-session treatment: Principles and procedures with children and adolescents. In *Intensive one-session treatment of specific phobias* (pp. 97-125). Springer, New York, NY

de Souza e Silva, A. (2017). Pokémon Go as an HRG: Mobility, sociability, and surveillance in hybrid spaces. *Mobile Media and Communication*, *5*(1), 20–23. https://doi.org/10.1177/2050157916676232

Dellazizzo, L., Potvin, S., Luigi, M., & Dumais, A. (2020). Evidence on virtual reality-based therapies for psychiatric disorders: Meta-review of meta-analyses. *Journal of Medical Internet Research, 22*(8). https://doi.org/10.2196/20889

Depla, M. F. I. A., ten Have, M. L., van Balkom, A. J. L. M., & de Graaf, R. (2008). Specific fears and phobias in the general population: Results from the Netherlands Mental Health Survey and

Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology*, *43*(3), 200–208. https://doi.org/10.1007/s00127-007-0291-z

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

East, M. L., & Havard, B. C. (2015). Mental Health Mobile Apps: From Infusion to Diffusion in the Mental Health Social System. *JMIR Mental Health*, *2*(1), e10. https://doi.org/10.2196/mental.3954

Eaton, W. W., Bienvenu, O. J., & Miloyan, B. (2018). Specific phobias. *The Lancet Psychiatry*, *5*(8), 678-686. https://doi.org/10.1016/S2215-0366(18)30169-X

Ebert, C. (2015). Looking into the future. *IEEE Software 32*(6), 92–97. doi: 10.1109/MS. 2015.142

Eichenberg, C. (2011). Psychotherapie und Internet. *Psychotherapeut*, *56*(6), 468–474. https://doi.org/10.1007/s00278-011-0865-9

Erbe, D., Eichert, H. C., Riper, H., & Ebert, D. D. (2017). Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. *Journal of medical Internet research*, *19*(9), e6588. https://doi.org/10.2196/jmir.6588

Essau, C. A., Conradt, J., & Petermann, F. (2000). Frequency, Comorbidity, and Psychosocial Impairment of Specific Phobia in Adolescents. *Journal of Clinical Child and Adolescent Psychology*, *29*(2), 221–231. https://doi.org/10.1207/S15374424jccp2902_8

Farrer, L., Christensen, H., Griffiths, K. M., & Mackinnon, A. (2012). Web-based cognitive behavior therapy for depression with and without telephone tracking in a national helpline: Secondary outcomes from a randomized controlled trial. *Journal of Medical Internet Research*, *14*(3). https://doi.org/10.2196/jmir.1859

Firth, J., Torous, J., Nicholas, J., Carney, R., Pratap, A., Rosenbaum, S., & Sarris, J. (2017a). The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*, *16*(3), 287–298. https://doi.org/10.1002/wps.20472

Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S., & Sarris, J. (2017b). Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of affective disorders*, *218*, 15-22.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin, 99*(1), 20–35. https://doi.org/10.1037/0033-2909.99.1.20

Fredrikson, M., Annas, P., Fischer, H., & Wik, G. (1996). Gender and age differences in the prevalence of specific fears and phobias. *Behaviour research and therapy*, *34*(1), 33-39.

Freeman, D., Haselton, P., Freeman, J., Spanlang, B., Kishore, S., Albery, E., Denne, M., Brown, P., Slater, M., & Nickless, A. (2018). Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry*, *5*(8), 625–632. https://doi.org/10.1016/S2215-0366(18)30226-8

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

García-Batista, Z. E., Guerra-Peña, K., Alsina-Jurnet, I., Cano-Vindel, A., Cantisano-Guzmán, L. M., Nazir-Ferreiras, A., ... & Garrido, L. E. (2021). Design and Validation of Augmented Reality Stimuli for the Treatment of Cleaning Obsessive-Compulsive Disorder. *Frontiers in Psychology*, *12*, 1417.

Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychology and Behavior*, *10*(5), 722–724. https://doi.org/10.1089/cpb.2007.9962

Garcia-Palacios, A., Hoffman, H. G., Kwong See, S., Tsai, A. M. Y., & Botella, C. (2001). Redefining therapeutic success with virtual reality exposure therapy. *CyberPsychology & Behavior*, *4*(3), 341-348.

Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. Psychological Review, 77(5), 419–450. https://doi.org/10.1037/h0029810

Harvey, A. G., & Gumport, N. B. (2015). Evidence-based psychological treatments for mental disorders: Modifiable barriers to access and possible solutions. *Behaviour Research and Therapy*, *68*, 1–12. https://doi.org/10.1016/j.brat.2015.02.004

Henderson, C., Evans-Lacko, S., & Thornicroft, G. (2013). Mental illness stigma, help seeking, and public health programs. American Journal of Public Health, 103(5), 777–780. https://doi.org/10.2105/AJPH.2012.301056

Hilty, D. M., Randhawa, K., Maheu, M. M., McKean, A. J. S., Pantera, R., Mishkind, M. C., & Rizzo, A. “Skip.” (2020). A Review of Telepresence, Virtual Reality, and Augmented Reality Applied to Clinical Care. *Journal of Technology in Behavioral Science*, *5*(2), 178–205. https://doi.org/10.1007/s41347-020-00126-x

Hoffman, L. J., & Chu, B. C. (2019). When is seeking safety functional? Taking a pragmatic approach to distinguishing coping from safety. *Cognitive and Behavioral Practice*, *26*(1), 176-185.

Hofmann, S. G., & Hay, A. C. (2018). Rethinking avoidance: Toward a balanced approach to avoidance in treating anxiety disorders. *Journal of anxiety disorders*, *55*, 14-21. https://doi.org/10.1016/j.janxdis.2018.03.004

Huckvale, K., Nicholas, J., Torous, J., & Larsen, M. E. (2020). Smartphone apps for the treatment of mental health conditions: status and considerations. *Current opinion in psychology*, *36*, 65-70. https://doi.org/10.1016/j.copsyc.2020.04.008

Huppert, D., Grill, E., & Brandt, T. (2013). Down on heights? One in three has visual height intolerance. *Journal of Neurology*, *260*(2), 597–604. https://doi.org/10.1007/s00415-012-6685-1

Insel, T. R. (2017). Digital phenotyping: technology for a new science of behavior. *Jama*, *318*(13), 1215-1216. https://doi.org/10.1001/jama.2017.11295

Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., Busch, M. A., Hapke, U., Maske, U., Seiffert, I., Gaebel, W., Maier, W., Wagner, M., Zielasek, J., & Wittchen, H. U. (2014). Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: The mental health

module of the german health interview and examination survey for adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, *23*(3), 304–319. https://doi.org/10.1002/mpr.1439

Juan, M. C., Alcaniz, M., Monserrat, C., Botella, C., Baños, R. M., & Guerrero, B. (2005). Using augmented reality to treat phobias. *IEEE computer graphics and applications*, *25*(6), 31-37.

Kertz, S. J., MacLaren Kelly, J., Stevens, K. T., Schrock, M., & Danitz, S. B. (2017). A Review of Free iPhone Applications Designed to Target Anxiety and Worry. *Journal of Technology in Behavioral Science*, *2*(2), 61–70. https://doi.org/10.1007/s41347-016-0006-y

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, *21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Korolov, M. (2014, Oct 4). The real risks of virtual reality. Risk Management. https://www.rmmagazine.com/articles/article/2014/10/01/-the-real-risks-of-virtual-reality-

Kristensen, N., Nymann, C., & Konradsen, H. (2016). Implementing research results in clinical practice-the experiences of healthcare professionals. *BMC Health Services Research*, *16*(1). https://doi.org/10.1186/s12913-016-1292-y

Labrique, A. B., Vasudevan, L., Kochi, E., Fabricant, R., & Mehl, G. (2013). mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. *Global health: science and practice*, *1*(2), 160-171.

Lagan, S., Aquino, P., Emerson, M. R., Fortuna, K., Walker, R., & Torous, J. (2020). Actionable health app evaluation: translating expert frameworks into objective metrics. *Npj Digital Medicine*, *3*(1). https://doi.org/10.1038/s41746-020-00312-4

Lagan, S., Sandler, L., & Torous, J. (2021). Evaluating evaluation frameworks: A scoping review of frameworks for assessing health apps. *BMJ Open, 11*(3). https://doi.org/10.1136/bmjopen-2020-047001

Lambe, S., Knight, I., Kabir, T., West, J., Patel, R., Lister, R., Rosebrock, L., Rovira, A., Garnish, B., Freeman, J., M. Clark, D., Waite, F., & Freeman, D. (2020). Developing an automated VR cognitive treatment for psychosis: gameChange VR therapy. *Journal of Behavioral and Cognitive Therapy*, *30*(1), 33–40. https://doi.org/10.1016/j.jbct.2019.12.001

Larsen, M. E., Huckvale, K., Nicholas, J., Torous, J., Birrell, L., Li, E., & Reda, B. (2019). Using science to sell apps: Evaluation of mental health app store quality claims. *Npj Digital Medicine*, *2*(1). https://doi.org/10.1038/s41746-019-0093-1

LaViola Jr, J. J. (2000). A discussion of cybersickness in virtual environments. *ACM Sigchi Bulletin*, *32*(1), 47-56.

Leigh, S., & Flatt, S. (2015). App-based psychological interventions: Friend or foe? *Evidence-Based Mental Health*, *18*(4), 97–99. https://doi.org/10.1136/eb-2015-102203

Lindner, P. (2021). Better, virtually: the past, present, and future of virtual reality cognitive behavior therapy. *International Journal of Cognitive Therapy*, *14*(1), 23-46. https://doi.org/10.1007/s41811-020-00090-7

Lindner, P., Dafgård, P., Miloff, A., Andersson, G., Reuterskiöld, L., Hamilton, W., & Carlbring, P. (2021a). Is Continued Improvement After Automated Virtual Reality Exposure Therapy for Spider Phobia Explained by Subsequent in-vivo Exposure? A First Test of the Lowered Threshold Hypothesis. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.645273

Lindner, P., Dagöö, J., Hamilton, W., Miloff, A., Andersson, G., Schill, A., & Carlbring, P. (2021b). Virtual Reality exposure therapy for public speaking anxiety in routine care: a single-subject effectiveness trial. *Cognitive Behaviour Therapy*, *50*(1), 67–87. https://doi.org/10.1080/16506073.2020.1795240

Lindner, P., Miloff, A., Zetterlund, E., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2019). Attitudes toward and familiarity with virtual reality therapy among practicing cognitive behavior therapists: a crosssectional survey study in the era of consumer VR platforms. *Frontiers in Psychology, 10*, 1–10. https://doi.org/10.3389/fpsyg.2019.00176

Lindner, P., Rozental, A., Jurell, A., Reuterskiöld, L., Andersson, G., Hamilton, W., Miloff, A., & Carlbring, P. (2020). Experiences of gamified and automated virtual reality exposure therapy for spider phobia: Qualitative study. *JMIR Serious Games*, *8*(2). https://doi.org/10.2196/17807

Luckerson, V. (2014, Mar 24). Facebook Buying Oculus Virtual-Reality Company for $2 Billion. Time. http://time.com/37842/facebook-oculus-rift

Lungu, A. J., Swinkels, W., Claesen, L., Tu, P., Egger, J., & Chen, X. (2021). A review on the applications of virtual reality, augmented reality and mixed reality in surgical simulation: an extension to different kinds of surgery. *Expert review of medical devices*, *18*(1), 47-62. https://doi.org/10.1080/17434440.2021.1860750

Ma, L., Mor, S., Anderson, P. L., Baños, R. M., Botella, C., Bouchard, S., Cárdenas-López, G., Donker, T., Fernández-Álvarez, J., Lindner, P., Mühlberger, A., Powers, M. B., Quero, S., Rothbaum, B., Wiederhold, B. K., & Carlbring, P. (2021). Integrating virtual realities and psychotherapy: SWOT analysis on VR and MR based treatments of anxiety and stress-related disorders. *Cognitive Behaviour Therapy*, *50*(6), 509–526. https://doi.org/10.1080/16506073.2021.1939410

Mack, S., Jacobi, F., Gerschler, A., Strehle, J., Höfler, M., Busch, M. A., Maske, U. E., Hapke, U., Seiffert, I., Gaebel, W., Zielasek, J., Maier, W., & Wittchen, H. U. (2014). Self-reported utilization of mental health services in the adult German population - evidence for unmet needs? Results of the DEGS1-Mental health module (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, *23*(3), 289–303. https://doi.org/10.1002/mpr.1438

Magee, W. J., Eaton, W. W., Wittchen, H. U., McGonagle, K. A., & Kessler, R. C. (1996). Agoraphobia, simple phobia, and social phobia in the National Comorbidity Survey. *Archives of general psychiatry*, *53*(2), 159-168.

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), 103. https://doi.org/10.1097/HRP.0000000000000138

Melo, M., Vasconcelos-Raposo, J., & Bessa, M. (2018). Presence and cybersickness in immersive content: Effects of content type, exposure time and gender. *Computers and Graphics (Pergamon)*, *71*, 159–165. https://doi.org/10.1016/j.cag.2017.11.007

Michael, T., Zetsche, U., & Margraf, J. (2007). Epidemiology of anxiety disorders. *Psychiatry*, *6*(4), 136-142.

Miloff, A., Carlbring, P., Hamilton, W., Andersson, G., Reuterskiöld, L., & Lindner, P. (2020). Measuring alliance toward embodied virtual therapists in the era of automated treatments with the Virtual Therapist Alliance Scale (VTAS): Development and psychometric evaluation. *Journal of Medical Internet Research*, *22*(3). https://doi.org/10.2196/16660

Miloff, A., Lindner, P., Dafgård, P., Deak, S., Garke, M., Hamilton, W., Heinsoo, J., Kristoffersson, G., Rafi, J., Sindemark, K., Sjölund, J., Zenger, M., Reuterskiöld, L., Andersson, G., & Carlbring, P. (2019). Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial. *Behaviour Research and Therapy*, *118*, 130–140. https://doi.org/10.1016/j.brat.2019.04.004

Miloff, A., Marklund, A., & Carlbring, P. (2015). The challenger app for social anxiety disorder: New advances in mobile psychological treatment. *Internet Interventions, 2*(4), 382-391. https://doi.org/10.1016/j.invent.2015.08.001

Mishkind, M. C., Norr, A. M., Katz, A. C., & Reger, G. M. (2017). Review of virtual reality treatment in psychiatry: evidence versus current diffusion and use. *Current psychiatry reports*, *19*(11), 1-8.

Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33–35

Morina, N., Ijntema, H., Meyerbröker, K., & Emmelkamp, P. M. G. (2015). Can virtual reality exposure therapy gains be generalized to real-life? A meta-analysis of studies applying behavioral assessments. *Behaviour Research and Therapy*, *74*, 18–24.

Moukheiber, A., Rautureau, G., Perez-Diaz, F., Soussignan, R., Dubal, S., Jouvent, R., & Pelissolo, A. (2010). Gaze avoidance in social phobia: Objective measure and correlates. *Behaviour Research and Therapy*, *48*(2), 147–151. https://doi.org/10.1016/j.brat.2009.09.012

Mubarak, I. (2021). *Augmented Reality in Psychology: Its Advancement in Therapy for Simple Phobia* (No. 6067). EasyChair.

Mueller, F. D., Fehlmann, B., Wang, N., Ibach, M. K., Schlitt, T., Bentz, D., Zimmer, A., Papassotiropoulos, A. & de Quervain, D. J. (2022). Virtual reality-based gaze exposure treatment reduces fear of public speaking. *PNAS (submitted for publication)*

Musiat, P., & Tarrier, N. (2014). Collateral outcomes in e-mental health: A systematic review of the evidence for added benefits of computerized cognitive behavior therapy interventions for mental health. *Psychological Medicine*, *44*(15), 3137–3150. https://doi.org/10.1017/S0033291714000245

Mystkowski, J. L., Craske, M. G., & Echiverri, A. M. (2002). Treatment context and return of fear in spider phobia. *Behavior Therapy*, *33*(3), 399-416.

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design

principles for ongoing health behavior support. *Annals of Behavioral Medicine*, *52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Naslund, J. A., Aschbrenner, K. A., Araya, R., Marsch, L. A., Unützer, J., Patel, V., & Bartels, S. J. (2017). Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. *The Lancet Psychiatry 4*(6), 486-500. Elsevier Ltd. https://doi.org/10.1016/S2215-0366(17)30096-2

National Institute of Mental Health (2017). Technology and the Future of Mental Health Treatment. Available online: https://www.nimh.nih.gov/health/topics/technology-and-the-future-of-mental-health-treatment/index.shtml

Neary, M., & Schueller, S. M. (2018). State of the Field of Mental Health Apps. *Cognitive and Behavioral Practice*, *25*(4), 531–537. https://doi.org/10.1016/j.cbpra.2018.01.002

Neudeck, P., & Einsle, F. (2012). Dissemination of exposure therapy in clinical practice: How to handle the barriers?. In *Exposure therapy* (pp. 23-34). Springer, New York, NY.

Ng, M. M., Firth, J., Minen, M., & Torous, J. (2019). User engagement in mental health apps: A review of measurement, reporting, and validity. *Psychiatric Services*, *70*(7), 538–544. https://doi.org/10.1176/appi.ps.201800519

Nixon, M. E., & Howard, A. M. (2013). Applying gaming principles to virtual environments for upper extremity therapy games. *Proceedings – 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013*, 3430–3435. https://doi.org/10.1109/SMC.2013.585

Park, M. J., Kim, D. J., Lee, U., Na, E. J., & Jeon, H. J. (2019). A literature overview of virtual reality (VR) in treatment of psychiatric disorders: recent advances and limitations. *Frontiers in psychiatry*, *10*, 505. https://doi.org/10.3389/fpsyt.2019.00505

Pittig, A., & Hoyer, J. (2017). Utilization and barriers of exposure in private practice: The perspective of behavioral psychotherapists. *Zeitschrift für Klinische Psychologie und Psychotherapie*, *46*(4), 223–235. https://doi.org/10.1026/1616-3443/a000441

Pittig, A., Kotter, R., & Hoyer, J. (2019). The Struggle of Behavioral Therapists With Exposure: Self-Reported Practicability, Negative Beliefs, and Therapist Distress About Exposure-Based Interventions. *Behavior Therapy*, *50*(2), 353–366. https://doi.org/10.1016/j.beth.2018.07.003

Polák, J., Sedláčková, K., Landová, E., & Frynta, D. (2020). Faster detection of snake and spider phobia: revisited. *Heliyon*, *6*(5). https://doi.org/10.1016/j.heliyon.2020.e03968

Pollard, C. A., & Henderson, J. G. (1988). Four types of social phobia in a community sample. *Journal of Nervous and Mental Disease*.

Price, M., Yuen, E. K., Goetter, E. M., Herbert, J. D., Forman, E. M., Acierno, R., & Ruggiero, K. J. (2014). mHealth: A mechanism to deliver more accessible, more effective mental health care. *Clinical Psychology and Psychotherapy*, *21*(5), 427–436. https://doi.org/10.1002/cpp.1855

Proudfoot, J. (2013). The future is in our hands: the role of mobile phones in the prevention and management of mental disorders. *The Australian and New Zealand Journal of Psychiatry, 47*(2), 111–113. https://doi.org/10.1177/0004867412471441

Pull, C. B. (2012). Current status of knowledge on public-speaking anxiety. *Current opinion in psychiatry*, *25*(1), 32-38. https://doi.org/10.1097/YCO.0b013e32834e06dc

Riva, G. (2022). Virtual reality in clinical psychology. *Reference Module in Neuroscience and Biobehavioral Psychology*. https://doi.org/10.1016/b978-0-12-818697-8.00006-6

Roth, C. B., Papassotiropoulos, A., Brühl, A. B., Lang, U. E., & Huber, C. G. (2021). Psychiatry in the Digital Age: A Blessing or a Curse?. *International journal of environmental research and public health*, *18*(16), 8302. https://doi.org/10.3390/ijerph18168302

Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the USA: results from the National Comorbidity Survey Replication. *Psychological medicine*, *38*(1), 15-28.

Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience, 6*(4), 332–339. https://doi.org/10.1038/nrn1651

Scarfe, P., & Glennerster, A. (2019). The science behind virtual reality displays. *Annual review of vision science*, *5*, 529-547.

Schienle A. (2021) Spezifische Phobien. In: Schnell T., Schnell K. (eds) Handbuch Klinische Psychologie. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-45995-9_3-1

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). Igroup presence questionnaire. *Teleoperators Virtual Environ.*, *41*, 115-124.

Schueller, S. M., Washburn, J. J., & Price, M. (2016). Exploring mental health providers' interest in using web and mobile-based tools in their practices. *Internet Interventions*, *4*, 145–151. https://doi.org/10.1016/j.invent.2016.06.004

Schulze, L., Renneberg, B., & Lobmaier, J. S. (2013). Gaze perception in social anxiety and social anxiety disorder. *Frontiers in human neuroscience*, *7*, 872. https://doi.org/10.3389/fnhum.2013.00872

Schwartzman, D., Segal, R., & Drapeau, M. (2012). Perceptions of virtual reality among therapists who do not apply this technology in clinical practice. *Psychological Services*, *9*(3), 310–315. https://doi.org/10.1037/a0026801

Segal, R., Bhatia, M., & Drapeau, M. (2011). Therapists' perception of benefits and costs of using virtual reality treatments. *Cyberpsychology, Behavior, and Social Networking*, *14*(1–2), 29–34. https://doi.org/10.1089/cyber.2009.0398

Seim, R. W., & Spates, C. R. (2010). The prevalence and comorbidity of specific phobias in college students and their interest in receiving treatment. *Journal of College Student Psychotherapy*, *24*(1), 49–58. https://doi.org/10.1080/87568220903400302

Slater, M., Khanna, P., Mortensen, J., & Yu, I. (2009). Visual realism enhances realistic response in an immersive virtual environment. *IEEE computer graphics and applications*, *29*(3), 76-84.

So, M., Yamaguchi, S., Hashimoto, S., Sado, M., Furukawa, T. A., & McCrone, P. (2013). Is computerised CBT really helpful for adult depression? A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity. *BMC psychiatry*, *13*(1), 1-14.

Sporinova, B., Manns, B., Tonelli, M., Hemmelgarn, B., Macmaster, F., Mitchell, N., Au, F., Ma, Z., Weaver, R., & Quinn, A. (2019). Association of Mental Health Disorders with Health Care Utilization and Costs among Adults with Chronic Disease. *JAMA Network Open*, *2*(8). https://doi.org/10.1001/jamanetworkopen.2019.9910

Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: A systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology*, *43*(2), 476–493. https://doi.org/10.1093/ije/dyu038

Stinson, F. S., Dawson, D. S., Chou, S. P., Smith, S., Goldtein, R. B., Ruan, W. J., & Grant, B. F. (2007). The epidemiology of DSM-IV specific phobia in the USA: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Psychological Medicine*, *37*(7), 1047–1059. https://doi.org/10.1017/S0033291707000086

Suso-Ribera, C., Fernández-Álvarez, J., García-Palacios, A., Hoffman, H. G., Bretón-López, J., Baños, R. M., Quero, S., & Botella, C. (2019). Virtual Reality, Augmented Reality, and in Vivo Exposure Therapy: A Preliminary Comparison of Treatment Efficacy in Small Animal Phobia. *Cyberpsychology, Behavior, and Social Networking*, *22*(1), 31–38. https://doi.org/10.1089/cyber.2017.0672

Tardif, N., Therrien, C. É., & Bouchard, S. (2019). Re-Examining Psychological Mechanisms Underlying Virtual Reality-Based Exposure for Spider Phobia. *Cyberpsychology, Behavior, and Social Networking*, *22*(1), 39–45. https://doi.org/10.1089/cyber.2017.0711

Tarnawski, M. (2017, November). Treating presence as a noun – insights obtained from comparing a VE and a 360 video. In *Proceedings of SIGRAD 2017, August 17-18, 2017 Norrköping, Sweden* (No. 143, pp. 9-16). Linköping University Electronic Press.

Taylor-Rodgers, E., & Batterham, P. J. (2014). Evaluation of an online psychoeducation intervention to promote mental health help seeking attitudes and intentions among young adults: Randomised controlled trial. *Journal of Affective Disorders*, *168*, 65–71. https://doi.org/10.1016/j.jad.2014.06.047

Thyloth, M., Singh, H., & Subramanian, V. (2016). Increasing burden of mental illnesses across the globe: Current status. *Indian Journal of Social Psychiatry*, *32*(3), 254. https:// doi.org/10.4103/0971-9962.193208

Tønning, M. L., Kessing, L. V., Bardram, J. E., & Faurholt-Jepsen, M. (2019). Methodological challenges in randomized controlled trials on smartphone-based treatment in psychiatry: systematic review. *Journal of medical Internet research*, *21*(10), e15362. https://doi.org/10.2196/15362

Torous, J., Andersson, G., Bertagnoli, A., Christensen, H., Cuijpers, P., Firth, J., ... & Arean, P. A. (2019). Towards a consensus around standards for smartphone apps and digital mental health. *World Psychiatry*, *18*(1), 97. https://doi.org/10.1002/wps.20592

Torous, J., Chan, S. R., Tan, S. Y. M., Behrens, J., Mathew, I., Conrad, E. J., Hinton, L., Yellowlees, P., & Keshavan, M. (2014). Patient smartphone ownership and interest in mobile apps to monitor symptoms of mental health conditions: A survey in four geographically distinct psychiatric clinics. *JMIR Mental Health*, *1*(1). https://doi.org/10.2196/mental.4004

Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evidence-based mental health*, *21*(3), 116-119. https://doi.org/10.1136/eb-2018-102891

Torous, J., & Powell, A. C. (2015). Current research and trends in the use of smartphone applications for mood disorders. *Internet Interventions*, *2*(2), 169-173. https://doi.org/10.1016/j.invent.2015.03.002

Torous, J., & Roberts, L. W. (2017). Needed innovation in digital health and smartphone applications for mental health: transparency and trust. *JAMA psychiatry*, *74*(5), 437-438. https://doi.org/10.1001/jamapsychiatry.2017.0262

Torous, J., Wisniewski, H., Liu, G., & Keshavan, M. (2018). Mental health mobile phone app usage, concerns, and benefits among psychiatric outpatients: comparative survey study. *JMIR* mental health, 5(4), e11715. https://doi.org/10.2196/11715

Vincent, C. J., Niezen, G., O'Kane, A. A., & Stawarz, K. (2015). Can standards and regulations keep up with health technology? *JMIR mHealth and uHealth*, *3*(2), e3918. https://doi.org/10.2196/mhealth.3918

Vinci, C., Brandon, K. O., Kleinjan, M., & Brandon, T. H. (2020a). The clinical potential of augmented reality. *Clinical Psychology: Science and Practice*, *27*(3), e12357. https://doi.org/10.1111/cpsp.12357

Vinci, C., Brandon, K. O., Kleinjan, M., Hernandez, L. M., Sawyer, L. E., Haneke, J., Sutton, S. K., & Brandon, T. H. (2020b). Augmented reality for smoking cessation: Development and usability study. *JMIR MHealth and UHealth*, *8*(12). https://doi.org/10.2196/21643

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, *3*(2), 171-178.

Wang, P. S., Angermeyer, M., Borges, G., Bruffaerts, R., Chiu, W. T., De Girolamo, G., ... & Uestuen, T. B. (2007). Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World psychiatry*, *6*(3), 177.

Wilhelm, S., Weingarden, H., Ladis, I., Braddick, V., Shin, J., & Jacobson, N. C. (2020). Cognitive-behavioral therapy in the digital age: presidential address. *Behavior Therapy*, *51*(1), 1-14.

Wisniewski, H., Liu, G., Henson, P., Vaidyam, A., Hajratalli, N. K., Onnela, J. P., & Torous, J. (2019). Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. *Evidence-Based Mental Health*, *22*(1), 4–9. https://doi.org/10.1136/ebmental-2018-300069

Wolitzky-Taylor, K. B., Horowitz, J. D., Powers, M. B., & Telch, M. J. (2008). Psychological approaches in the treatment of specific phobias: A meta-analysis. *Clinical psychology review*, *28*(6), 1021-1037. https://doi.org/10.1016/j.cpr.2008.02.007

Wrzesien, M., Alcañiz, M., Botella, C., Burkhardt, J. M., Bretón-López, J., Ortega, M., & Brotons, D. B. (2013). The therapeutic lamp: treating small-animal phobias. *IEEE computer graphics and applications*, *33*(1), 80-86.

Wu, A., Scult, M. A., Barnes, E. D., Betancourt, J. A., Falk, A., & Gunning, F. M. (2021). Smartphone apps for depression and anxiety: a systematic review and meta-analysis of techniques to

increase engagement. *NPJ digital medicine*, *4*(1), 1-9. https://doi.org/10.1038/s41746-021-00386-8

Zimmer, A., Wang, N., Ibach, M. K., Fehlmann, B., Schicktanz, N. S., Bentz, D., ... & de Quervain, D. J. (2021). Effectiveness of a smartphone-based, augmented reality exposure app to reduce fear of spiders in real-life: A randomized controlled trial. *Journal of anxiety disorders*, *82*, 102442.