UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

The Effect of a CBT-BASED Mental Health Application (WOEBOT) on the Psychological Distress of South African University Students

By

Keanan Gernandt

Submitted in partial fulfilment of the requirements for the degrees

Master of Arts in Psychology

in the

Department of Psychology

Faculty of Humanities

UNIVERSITY OF PRETORIA

October 2022

Supervisor: Dr. B. Motileng

ABSTRACT

Mental health issues such as psychological distress are widespread among university students throughout the world, including South Africa. Chatbot mental health apps such as Woebot that employ artificial intelligence technology and cognitive behavioural therapy principles may be an alternative mode of mental health support for university students suffering from psychological distress. The aim of this quasi-experimental research design was to evaluate the effect of Woebot on the psychological distress levels of a sample of university students. University students (n=28) with elevated levels of psychological distress (General Health Questionnaire, GHQ-12≥9) were assigned nonrandomly through self-selection to either a mental health app intervention group or a non-intervention comparison group (control group). Self-report data were assessed at pre-treatment and post-treatment (3 weeks). The primary outcome measure was psychological distress change scores from pre-treatment to post-treatment (GHQ-12). A Spearman's rank-order correlation test and an independent samples t-test were employed in this study. The results showed a non-statistically significant negative correlation between the participants' average number of weekly Woebot interaction days and psychological distress. Moreover, the results showed that the Woebot intervention had a reduction effect on the psychological distress of university students. Further research is recommended to establish Woebot's effectiveness on psychological distress in larger samples, over more extended periods and using more rigorous research designs such as randomised controlled trials.

Keywords: mental health apps, psychological distress, university students, South Africa, cognitive behavioural therapy, chatbots.

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere gratitude to my supervisor, Dr. Moteleng, for his unwavering support and patience throughout my master's program. Without his patience and encouragement throughout the past few years, I would not have been able to complete this study. My deepest appreciation also extends to my family, friends, and all those who made this possible by encouraging and assisting me during my studies. Last but not least, I would like to thank everyone who participated in my study for their time and dedication.

DECLARATION

I understand what plagiarism is and am aware of the University's policy. In this regard, I declare that the dissertation hereby submitted to the University of Pretoria for the degree of Master of Arts in Psychology is my original work. Where other people's work has been used (either from a printed source, internet or any other source), this has been properly acknowledged and referenced following the requirements as stated in the University's plagiarism prevention policy. I have not used another student's past written work to hand in as my own. I have not allowed, and will not allow anyone, to copy my work to pass it off as their work.

Initials and Surname: K L Gernandt **Date:** 7 October, 2022

ETHICS STATEMENT

I, **Keanan Gernandt**, student number **15073778**, have obtained ethical approval for the research titled: The Effect of a CBT-Based Mental Health Application on The Psychological Distress of University Students. On 15 November 2021, I received ethical approval (reference number: HUM029/0221) from Prof Karin Harris, Chair of the Research Ethics in the Faculty of Humanities at the University of Pretoria.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

This chapter introduces the study and sets the scene for the chapters to follow. First, the background or context in which the study is embedded is presented. The research problem is briefly outlined and the problem statement specified thereafter. Subsequently, the rationale for the study is presented, along with the research questions and hypotheses. Thereafter, the key concepts and variables of the study are defined. Finally, the structure of the dissertation's chapters is broadly defined.

Background

Psychological distress (PD), which may be defined as a state of emotional suffering marked by symptoms of depression and anxiety, is becoming increasingly prevalent among university students worldwide as well as in South Africa (Arvidsdotter et al., 2016; Auerbach et al., 2018; Bantjes et al., 2016; Drapeau et al., 2012; McGowan & Kagee, 2013; Mirowsky & Ross, 2002; Sharp & Theiler, 2018). Bantjes et al. (2016) found that approximately 12-15 % of South African university students reported moderate to severe symptoms of PD.

The availability of mental healthcare resources and support is essential to assist university students with PD. However, many students have reported feeling disincentivised from seeking mental health support due to stigma, waiting lists and privacy concerns (Bantjes et al., 2020; Kaminer & Shabalala, 2019; Knapstad et al., 2021; Mboya et al., 2020; Sharp & Theiler, 2018). In addition, face-to-face counselling and mental health assistance may not always be available due to campus closures, as was the case during the COVID-19 pandemic (Son et al., 2020). Bantjes et al. (2020) revealed that only 18% of students surveyed at two well-resourced South African universities accessed on-campus mental health facilities during a 12-month period. In

addition, only 28.9% of students sought care for more serious psychological problems.

Mental health applications (MHapps) may be employed in an endevour to reach more students who need mental health assistance. MHapps are a broad category of smartphone applications that serve different purposes and functions to enhance mental health (Bergin & Davies, 2019). In this study, mental health was defined according to The World Health Organization's (WHO) conceptualisation as a "state of wellbeing in which the individual realises his or her abilities to cope with the normal stressors of life, work productively and fruitfully, and contribute to his or her community" (WHO, 2018, p. 1). As a consequence of the widespread availability of smartphones, app-based mental health assistance has recently emerged as a viable and scalable modality (Bakker et al., 2016; Berry et al., 2018; Kretzschmar et al., 2019; Prochaska et al., 2021). Thousands of MHapps are currently available for download in smartphone app stores. For example, there are now apps intended to help users with anxiety, addiction, depression, post-traumatic stress disorder (PTSD), stress, phobias and many other forms of psychological difficulties (Bergin & Davies, 2019; Lagan et al., 2021; Lau et al., 2021; Longyear & Kushlev, 2021; Marshall et al., 2020a ; McCloud et al., 2020; Milne-Ives et al., 2020).

Multiple research studies have recommended the potential use of technology-assisted interventions such as MHapps for PD in university students (Farrer et al., 2019; Fitzpatrick et al., 2017; Kajitani et al., 2020; McCloud et al., 2020). The purpose of this study was to determine the effect of a CBT-based mental health application (Woebot) on the PD levels of South African university students. Woebot is an artificial intelligence-based chatbot MHapp that, in essence, was developed to provide CBT principles, specifically regular interactions using a mood monitoring interface that gauges and provides feedback on the user's current psychological state

(Fitzpatrick et al., 2017; Monnier, 2020).

Research Problem and Question

The recent rise in smartphone availability has resulted in a global effort, spearheaded by the WHO, towards app-based mental healthcare solutions (Marshall et al., 2020b). The need for digital mental health solutions was particularly pertinent during the COVID-19 pandemic when face-to-face mental health support services were inaccessible due to protracted lockdown and social-distancing requirements (Chew et al., 2020). During this time, the demand for psychological support services was higher than usual due to the damaging psychological impact of social isolation, uncertainty and stress on people throughout the world (Chew et al., 2020; Sanderson et al., 2020). Marshall et al. (2021) examined the effectiveness of five different MHapps for anxiety and depression throughout the pandemic. They found that all five apps had positive outcomes for anxiety and depression, which were maintained for six months after the study. Wang et al. (2021) revealed a substantial upsurge in MHapp downloads during the COVID-19 pandemic. These findings indicated the potential of MHapps in crises time and when a rapid form of mental health support is needed.

Although many mental health apps are available for download, researchers have only validated a small percentage of MHapps (Bergin & Davies, 2019; Huckvale et al., 2020; Lagan et al., 2021; Lecomte et al., 2020; Marshall et al., 2020b). Currently, more than half of the South African population own smartphones (Silver & Johnson, 2018). Furthermore, MHapps may be an invaluable resource for those who are unable to access mental health care (Silver & Johnson, 2018). However, minimal research has been conducted on MHapps within the South African context. Moreover, improved understanding is required in relation to the feasibility and effectiveness of freely available MHapps such as Woebot for mental health issues such as

psychological distress (Baumel et al., 2020; Connolly et al., 2021; Marshall et al., 2020a; Rowe & Sauls, 2020; Torous et al., 2018). Furthermore, the effect of the Woebot app on measures of PD among South African university students has not been explicity examined in any studies.

In the interest of exploring the potential of app-based mental health solutions further, the current study aimed to determine the effect of the Woebot app on PD among South African university students. The study aimed to answer the following question: what is the effect of Woebot on the PD levels of a sample of South African university students?

Research Aims and Objectives

This study aimed to determine the effect of a CBT-based mental health application (Woebot) on the PD levels of South African university students so as to shed light on how interacting with Woebot affects university students' PD levels to determine if the app is beneficial for such students.

The first objective was to establish the direction and strength of the relationship between the number of days participants interacted with the Woebot app and their PD change scores, from pre-test to post-test, over three weeks. Statistical methods were employed to identify whether the relationship between Woebot interaction and PD was positive or negative. Accordingly, the following question was examined: What is the correlation between the number of average weekly Woebot interactions and PD reduction from pre-test to post-test?

The second objective was to determine whether there was a statistical difference between the changes in PD between the intervention group and comparison group from pre-test to post-test. Accordingly, the following question was investigated: After three weeks, was there a statistically significant difference between the PD change scores of the treatment group who received the Woebot intervention and the comparison group that did not receive any

intervention?

Research Hypotheses

The following hypotheses were formulated:

1. There is a negative correlation between the number of average weekly Woebot interactions and PD change scores from pre-test to post-test.
2. There is a statistically significant difference after three weeks between the PD change scores of the treatment group that received the Woebot intervention and comparison group that did not receive any intervention.

Dissertation Chapter Outline

The first chapter presented the background context of the study, together with the research problem under investigation. In the second chapter, the literature related to MHapps and the Woebot app is reviewed and integrated to justify the current study. The theoretical framework that informed the Woebot MHapp intervention is discussed in Chapter Three. Subsequently, the methodology employed in the study is examined in Chapter Four. Furthermore, an overview of the research design, sampling, data collection procedures, instruments, data analysis techniques and ethical considerations are provided in this chapter. The findings of the study are presented in Chapter Five. Specifically, the statistical results from the data analysis procedures are provided. Chapter Six comprises a synopsis of the previous chapters and interpretation of the results. The importance of the findings by considering their implications, limitations and recommendations for future research are highlighted in this chapter.

Conclusion

The study was introduced and contextualised in this chapter by providing an overview of

the research problem, namely high PD rates among university students and lack of mental health care seeking and the potential for employing mental health applications (MHapps) to assist a wider range of students with mental health needs. Moreover, to address this problem, it was argued that we could employ mental health applications (MHapps) such as Woebot to assist a wider range of students with mental health needs. The study aimed to address the following question: "what is the effect of Woebot on the PD levels of a sample of South African university students?" In the next chapter, the literature is reviewed to describe and integrate the themes and trends related to the research problem.

CHAPTER 2

Literature Review

Introduction

In the previous chapter, a brief overview of mental health apps, chatbot-app technology, and the Woebot app—the intervention for this study—was provided. In this chapter, a more in-depth literature review on this subject matter, starting with a broad overview of the landscape of mental health applications (MHapps) is presented. Specifically, the general trends and common themes that have emerged in MHapp research are highlighted. Subsequently, MHapp interventions are discussed and the evidence-based framework for selecting the Woebot app as the research intervention is detailed. Finally, artificial intelligence and chatbot MHapp technology research related to the Woebot app are discussed.

The Landscape of Mental Health Apps (MHapps)

"It could be argued that the greatest discoveries for humanity, the next frontier, will not be in quantum technology or space, but will instead be those that help us understand our own minds better" (Delloite, 2021, p. 4). In the last decade, smartphones have become an essential part of the lives of billions of people around the globe (Lundquist et al., 2014; Sarwar & Soomro, 2013). Currently, billions of people around the world and over 20 million South Africans use smartphones and spend many hours per day on their phones (O'Dea, 2020). Smartphone app technology has made it possible for users to have access to information, communication and entertainment 24/7 (Bakker et al., 2016; Malavolta, 2016). Furthermore, smartphone apps have revolutionised how people interact, conduct business, travel and manage productivity (Bakker et al., 2016).

Smartphone app-technology also offers a medium to provide mobile mental health support. For instance, thousands of mental health apps are currently available for download on app stores, ranging from online and mobile therapy apps such as *Talkspace* to stress relief apps such as *Sanvello*, meditation apps like *Calm*, depression apps such as *Depression CBT Self-Help Guide*, bipolar disorder apps like *eMoods*, symptom tracking apps such as *Bearable* and apps for posttraumatic stress disorder, for example, *PTSD Coach* (Bergin & Davies, 2019; Lagan et al., 2021; Lau et al., 2021; Longyear & Kushlev, 2021; Marshall et al., 2020a; McCloud et al., 2020; Milne-Ives et al., 2020). In addition, some MHapps may schedule reminders to take medicine or meditate and encourage people on a therapy waiting list to engage in self-help activities or psychoeducation to supplement their therapy protocol (Marshall et al., 2020a, 2020c).

Multiple international studies have revealed the potential for smartphone mental health apps to help university students with psychological distress (PD) (Darcy et al., 2021; Farrer et al., 2019; Fitzpatrick et al., 2017; Kajitani et al., 2020; McCloud et al., 2020). A recent systematic review of mental healthcare technological solutions for PD suggested that mental health apps were equally effective for university students as other digital mental health treatments such as the web, internet and virtual reality-based therapies (Lattie et al., 2019). McCloud et al. (2020) implemented a mental health app in a sample of university students and found a reduction in anxiety and depression symptoms for six weeks. In a pilot study of a smartphone app, Kajitani et al. found reductions in PD among university students throughout the study. Furthermore, MHapps have shown the potential to increase access to mental health care at universities by reducing barriers to seeking care (Kampel & Orman, 2017).

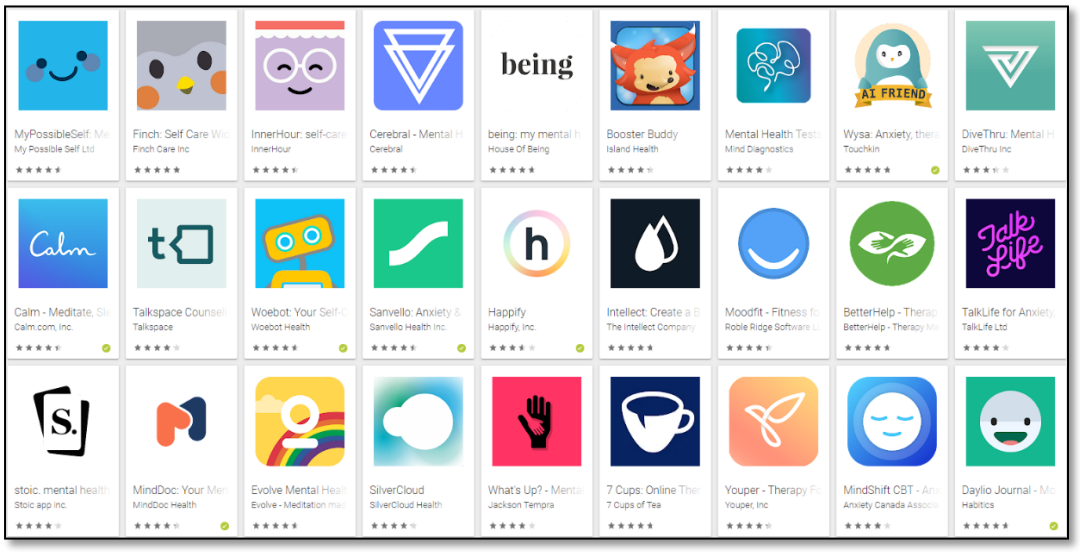
Chew et al. (2020) also found that MHapps have the potential to be used as a primary preventive care measure to help improve users' general mental health. However, with regard to

the potential of MHapps for primary and secondary prevention, it is crucial to consider individuals' capabilities and preferences for using MHapps. Michaelis et al. (2021) revealed that participants in their study were more likely to use MHapps if they had higher levels of technological literacy. Michaelis et al. (2021) also noted that users' opinions regarding using a mental health app are relevant to adherence.

The overabundance of MHapps and determining the most appropriate MHapp for a specific user was another central theme that emerged from the literature. App stores such as Google Play has up to 20 000 MHapps available for download, and MHapps are constantly being added and updated (Clay, 2021). Nevertheless, general users and mental health practitioners are overwhelmed in their quest to decide which app to choose and recommend to patients because of the sheer number of MHapps (Bergin & Davies, 2019).

Figure 1

Examples of Mental Health Apps Available for Download In App Stores



Several researchers have suggested various strategies for determining which MHapps are

appropriate for research and personal use. Bergin and Davies (2019) provided three criteria that can be employed to judge the suitability of MHapps. First, whether the information presented in the app is aligned with current evidence-based practices and theory should be considered. Second, it is crucial to consider the technical features of the app related to the privacy and security of data. Third, the app's usability, which refers to its interface and content's user-friendliness and relevance for a specific target audience, should be examined.

Similarly, Lagan et al. (2021) proposed a more in-depth framework to help clinicians and clients identify suitable MHapps. The Mhealth Index and Navigation Database (MIND) framework poses different questions to assess the appropriateness of various aspects of MHapps, including origin, accessibility, inputs, outputs, privacy, security, clinical foundation, features, engagement and interoperability. The MIND framework may be used to discover app features that are likely to contribute to high user engagement and efficacy. By using the MIND framework, Lagan et al. found that the five most widespread features of MHapps include mood tracking, journaling, mindfulness, psychoeducation and deep breathing. However, they also observed that even MHapps with many downloads and high user et al. (2021) highlighted the need for empirical frameworks, such as the MIND framework, for assessing and evaluating ratings did not meet the metrics for clinical suitability proposed by the MIND framework. Lagan MHapps. They also revealed that popular features, download numbers and user ratings alone should not be the only considerations for deciding whether to use or recommend a particular MHapp.

Another frequent challenge for MHapp developers and researchers is deciding what evidence-based features to incorporate into their app. As noted previously, many MHapps lack rigorous evidence related to features, efficacy and usability (Neary & Schueller, 2018). Bakker et

al. (2016) noted that many MHapps in smartphone app stores lack features that could significantly enhance their functionality. However, it is essential to first recognise some challenges that MHapp developers frequently encounter and how they can be overcome. Chandrashekar (2018) summarised three categorical challenges associated with MHapps: poor regulation of quality and privacy, inconsistent user engagement and a narrow focus on one disorder.

Furthermore, Chandrashekar (2018) recommended how MHapp developers can approach these challenges by including the following highly effective features: features that promote user engagement, such as real-time interactions, app reminders and gamification elements; a simple user interface and experience by including more images instead of text, reducing sentence lengths and employing language that is easy to understand; app features that simultaneously target multiple psychological disorders or symptoms; and self-monitoring features that enable users to record their thoughts, behaviour and actions so as to help users to learn more about themselves. In essence, it is impertaive that MHapp developers ensure that the quality and content of their apps are optimised for the most effective and engaging user experience.

Furthermore, Chandrashekar (2018) recommended that MHapp developers approach these challenges by including features that promote user engagement, such as real-time interactions, app reminders, and gamification elements. Moreover, developers should create a simple user interface and experience by including more images instead of text, reducing sentence lengths, and employing language that is easy to understand. The app should also have features that simultaneously target multiple psychological disorders or symptoms; and self-monitoring features that enable users to record their thoughts, behaviors, and actions so as to help users learn more about themselves.

Framework For Selecting This Study's MHapp Intervention

As discussed in the previous section, various frameworks can assist researchers to select suitable MHapp/s to investigate. Bakker et al. (2016) proposed a framework with 16 evidence-based recommendations for MHapp interventions based on mental health literature, user engagement studies and behavioural change app design research. This framework was chosen because it includes many considerations discussed in the previous section and was deemed to be particularly relevant for guiding the selection of a MHapp for psychological distress. The recommendations from Bakker's MHapp framework for selecting an MHapp intervention are subsequently discussed.

The first recommendation was that the app should be based on cognitive behavioural therapy (CBT). CBT is a psychological treatment that is effective for various problems, including depression, anxiety, substance abuse, relationship problems, eating disorders and severe mental illness (APA, 2017). Numerous studies have shown that CBT improves daily functioning and general well-being significantly (APA, 2017; Dobson & Dobson, 2009; Dozois, 2001; Fenn & Byrne, 2013; Ruggiero et al., 2018).

While CBT is typically applied in the context of a therapy setting, its principles have informed many forms of self-help resources that do not require the support of a therapist (Bakker et al., 2016), thus making it ideal to be incorporated into self-help mediums such as MHapps if its core principles can be adapted. According to DiGiuseppe et al. (2018), the core principles of CBT include the notions that (a) psychological issues are partially based on incorrect or distorted thinking patterns, (b) destructive or harmful behavioural patterns and (c) individuals with psychological difficulties can learn the skills and acquire the understanding to cope better with their psychological problems. Therefore, the first criterion for MHapp selection for this study

was to find an MHapp based on CBT principles.

Bakker et al.'s (2016) second applicable recommendation was that the app should address both anxiety and depression symptoms. In the First Chapter, it was noted the working definition of PD for this study was "a state of emotional suffering marked by symptoms of depression (e.g., lack of interest; sorrow; despair) and anxiety (e.g., agitation; feeling tense) in response to life stressors and difficulties" (Arvidsdotter et al., 2016; Drapeau et al., 2012; McLachlan & Gale, 2018; Mirowsky & Ross, 2002; Wheaton, 2007). Accordingly, the second criterion for MHapp selection for this study was to find an MHapp with the potential to address PD.

The third relevant recommendation was that the MHapp should be designed for nonclinical populations. Bakker et al. (2016) stated that although most smartphone users may not have been diagnosed with a clinical mental disorder, they may suffer from PD periodically. They added that if a MHapp is aimed narrowly at a particular clinical population, it inevitably excludes the bulk of smartphone users from utilising the app. On the contrary, Bakker et al. noted that if a MHapp is intended for nonclinical populations, it may increase the number of eligible and adherent users. Consequently, the third criterion for MHapp selection was to select an MHapp that is suitable for nonclinical populations.

The fourth recommendation was that the MHapp should include automated tailoring. Automated tailoring involves collecting data from the app to identify users' specific requirements and respond accordingly to help address them (Bakker et al., 2016). For example, if users report that they are experiencing anxiety, the app will respond according to that input by providing a tailored response to address anxiety.

The fifth applicable recommendation was that the app should help users monitor and track their thoughts, behaviour and feelings. Lagan et al. (2021) noted that mood tracking is one

of the most widely employed features of MHapps. Mood tracking allows users to monitor their moods, emotions, thoughts and behaviours and helps them learn more about themselves over time (Neary & Schueller, 2018). Therefore, the fourth criterion was that the MHapp should have mood-tracking capabilities.

The sixth MHapp suggestion was that the app should recommend activities aligned with the core principles of the therapeutic framework on which it is based. For example, if the framework is CBT, the app could recommend activities related to mood improvement, coping skills and/or behavioural activation (Bakker et al., 2016).

The seventh recommendation was that the app should provide mental health information, which is also referred to as psychoeducational material, to help users have an enhanced understanding of the dynamics that may lead to mental health problems and ways to prevent them. Lukens and McFarlane (2004) explained that psychoeducation is a treatment modality that integrates psychotherapeutic and educational interventions. Bakker et al. (2016) noted that smartphones' audiovisual and multimedia capabilities could make MHapps an ideal platform for delivering psychoeducation.

The eighth recommendation posited that the app should incorporate real-time engagement with users, which refers to users' ability to obtain assistance whenever they are experiencing psychological difficulties, specifically through immediate text, video and/or image responses from the app (Bakker et al., 2016).

The ninth recommendation was that the MHapp should offer activities that are directly related to the reported problem (Bakker et al., 2016). For example, while calming exercises should be provided for anxiety and stress, self-compassion exercises should be recommended for guilt feelings and self-judgement.

The 10th recommendation was that MHapp should promote nontechnology-based activities. Bakker et al. (2016) maintained that MHapps need to urge users to participate in real-world activities, away from their smartphone devices, which may help foster positive habits.

The 11th recommendation involved including gamification features and elements that promote intrinsic motivation to use the app. Deterding et al. (2011) defined gamification as "the use of game design elements in non-game contexts" (p. 13). Including gamification elements in an MHapp may provide users with a more engaging and rewarding experience. Bakker et al. (2016), for example, encouraged features that promote intrinsic motivation such as goal setting and progression schemes built into the MHapp.

While the 12th recommended feature was the inclusion of a log of past app use or records of previous interactions, the 13th recommendation was the inclusion of notifications to remind users to engage with the app. The 14th and 15th recommendations were for the app to have a simple interface and links or buttons for emergency care services, respectively. Finally, the 16th recommendation was that the MHapp had undergone previous experimental trials to determine its effectiveness.

The 16 recommendation framework was used to choose an MHapp intervention for PD based on CBT principles, with personalised features, gamification and a user-friendly interface. Moreover, it was imperative that the MHapp had been validated by previous studies. In the following section, the Woebot MHapp intervention that was selected in accordance with Bakker et al.'s (2016) recommendations and the considerations discussed earlier in this chapter are provided.

The World of Artificial Intelligence: Chatbots and The Woebot MHapp Intervention

Artificial Intelligence (AI) is an extensive branch of computer science that is involved

with developing smart machines that are able to conduct tasks that usually require human intelligence (Basu et al., 2020). Soroka and Kurkova (2019) stated that AI technology is already incorporated into many aspects of society, ranging from music recommendation algorithms to cars with automated steering. Lovejoy (2019) added that developments in computational power, data gathering and machine learning had influenced the growing interest in AI, as indicated by the recent increase in funding and research on these technologies.

AI has been successfully utilised in the medical field to help with diagnosis, medication development, doctor-patient communication, medical document transcription and remote care delivery to patients (Basu et al., 2020). However, only recently has the potential of AI been recognised in the field of mental health. Graham et al. (2019) posited that AI technology holds great potential for transforming mental healthcare. For instance, AI-based technologies such as electronic health records (EHRs), mood assessment measures, brain imaging records, innovative monitoring systems (e.g., smartphone, video) and social media formats have recently been employed in mental health settings to predict and classify various mental health disorders such as depression, schizophrenia and suicide ideation (Graham et al., 2019). Moreover, AI technology may ensure it is possible to provide existing treatment modalities such as CBT through innovative formats such as chatbots, which may improve treatment accessibility and efficacy substantially (Lovejoy, 2019).

Vaidyam et al. (2019) defined chatbots or conversational agents as “digital tools existing either as hardware or software that use machine learning and artificial intelligence methods to mimic human-like behaviour and provide a task-oriented framework with evolving dialogue able to participate in conversation” (p. 457). Woebot, the MHapp selected as the intervention for this study, is an example of this type of technology. Woebot is an artificial intelligence-based chatbot

MHapp, which was developed to provide CBT principles in short, regular interactions, with a mood monitoring interface that gauges and provides feedback on the user's current psychological state (Fitzpatrick et al., 2017; Monnier, 2020). The development of the Woebot app is the product of more than two decades of collaboration between psychologists at Stanford University and experts in AI and immersive technology (Monnier, 2020). The developers of Woebot explained that the app's chatbot interface was designed for non-clinical samples. Furthermore, the bot employs automated feedback on the user's thoughts, feelings and behaviour (Fitzpatrick et al., 2017).

Furthermore, Woebot provides information related to mental health and actively involves users in real-time activities such as thought reframing and visualisation (Fitzpatrick et al., 2017). The Woebot app has received over $123 million in investment funds to help make mental health support more accessible. Furthermore, the app can accessed by for free by downloading it from the Apple and Google Play stores (WoebotHealth, 2022). The Woebot app has previously been investigated for substance use problems and stress, anxiety and depression, and postpartum mood management (Fitzpatrick et al., 2017; Ramachandran et al., 2020; Vogel et al., 2021). However, to date, no research on the effect of the Woebot app on PD among South African university students has been conducted.

Conclusion

In this chapter, a broad overview of the landscape of mental health applications was presented by highlighting common themes that have emerged in the MHapp research literature. Furthermore, the framework for selecting the MHapp intervention was outlined and AI technology related to chatbots and the Woebot MHapp intervention discussed. In the next chapter, the theoretical framework that was employed in the study is outlined.

CHAPTER 3

Theoretical Framework

Cognitive-behavioural therapy (CBT) serves as the theoretical foundation for the current research. The principles and theory of CBT related to the Woebot app are discussed in this chapter. First, an overview of the history and development of CBT is provided. Second, the main principles of CBT are presented. Next, Woebot's CBT features for psychological distress are highlighted. Additionally, the rationale for why CBT theory may serve as a valuable framework for future developers of mental health applications is discussed.

Brief History of Cognitive Behavioural Therapy (CBT)

CBT refers to a collection of psychological therapies that are based on behavioural and cognitive theoretical frameworks and empirical research (Hofmann et al., 2012; Mennin et al., 2013; Ruggiero et al., 2018). CBT as a therapy technique was first developed in the 1960s and 1970s, partly because of the pioneering work of Albert Ellis and Aaron Beck. Its popularity has snowballed over the ensuing decades (Beck, 1997; Beck, 2011). David et al. (2018) posited that CBT is commonly regarded as the gold standard of psychological therapy and of any therapeutic approach has enjoyed the most robust scientific support (Beck, 2011). Modern CBT comprises a spectrum of intervention methods, including behavioural, emotion-focused and cognitive approaches that may be delivered by a therapist, online modality or self-help modality (David et al., 2018; Hofmann et al., 2012).

Cognitive behavioural therapy methods are constantly evolving and changing. Recent advancements in CBT have incorporated novel components and delivery methods such as mindfulness, group-based CBT, internet and app-based CBT, and self-help CBT (Andersson et al., 2014; Carlbring et al., 2018; Mennin et al., 2013;). For instance, different CBT methods, such

as mindfulness-based therapies, may focus on attention change principles. While specific approaches, such as cognitive therapy, may place a greater emphasis on cognitive change principles. Importantly, no CBT approach focuses exclusively on one mechanism or principle (Mennin et al., 2013). The same techniques used in face-to-face therapy, such as active participation in psychoeducational activities, self-monitoring of moods, emotions, behaviours, and thoughts, and the development of skills and habits that facilitate behavioural adaptation, can be adapted and applied to technological modalities (Beck, 2017).

Core Principles of CBT

Mennin et al. (2013) proposed that most CBT approaches share a common underlying framework, incorporating similar goals, change principles, and therapeutic methods. Notably, the main goal of all CBT approaches is behavioural adaptation. Mennin et al. (2013) asserted that behavioural adaptation might be accomplished by adhering to three essential principles. The first is context engagement, which facilitates adaptive envisioning and acting out novel experiences. The second is attentional change to provide adaptive attentional maintaining, shifting, and expanding. Thirdly, cognitive modification develops a flexible perspective on experiences, altering verbal connotations to experiences. Mennin et al. (2013) stated that, although the intricacies of CBT treatments may vary between programmes, they all share the same underlying building blocks that facilitate the therapeutic process and are best understood concerning the fundamental CBT transformation principles.

In addition to the shared objectives and principles of CBT modalities, three core theoretical assumptions underlie cognitive-behavioural therapy methods (Dobson & Dobson, 2008). The first principle is that cognitive functions and associated content are observable and identifiable. Although specific thoughts or beliefs are often not immediately apparent, individuals

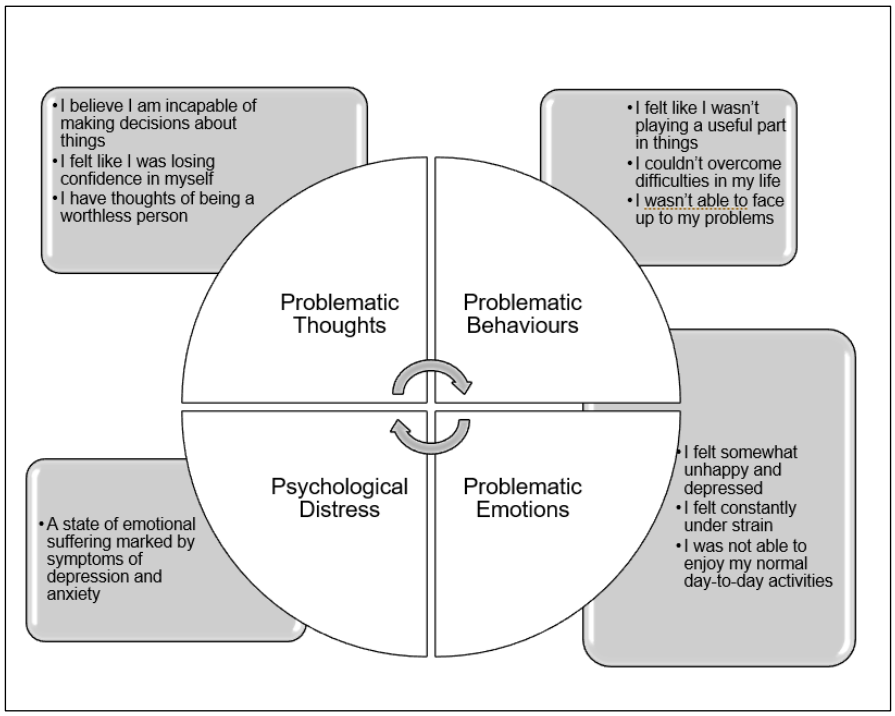
may develop an awareness of them (Dobson & Dobson, 2008). The second critical premise is that our thinking influences how we react to external stimuli. From this perspective, individuals do not just respond emotionally or behaviorally to life events. Instead, CBT maintains that how we think about our world is critical to our response to it (Hofmann et al., 2012). The third fundamental premise of CBT is that such cognitions may be targeted, manipulated, and altered purposefully. As a result, when such cognitions are shifted toward more rational, realistic, and balanced thinking, the individual's symptoms are alleviated, and the individual's adaptability and functioning are enhanced (Hofmann et al., 2012).

CBT, Mental Health Apps and Psychological Distress

According to Melton (2017), thoughts, behaviours, and emotions are interconnected, and any of the three can be a cause of psychological distress. Accordingly, the fundamental theoretical premise of this study was that problematic thoughts, behaviours and emotions result in PD, which can be challenged by employing a modality based on the principles of CBT. The CBT model was adapted and applied to the evaluation of PD by using the General Health Questionnaire 12-Items version in this study, which comprised items related to problematic thoughts, behaviours, and emotions (Figure 2).

Figure 2

CBT Model of Psychological Distress



The Cognitive Model and Woebot

The role of cognition or thought patterns is integral to the CBT model. Beck developed the cognitive model in 1970 to provide a conceptualisation of how thought processes in the CBT model may lead to beneficial or harmful mental health outcomes such as PD (Dobson, 2008). Beck (1976) proposed the cognitive triad, that is, three types of cognition: general, dysfunctional assumptions and negative automatic thoughts. Core beliefs, or schemas, are ingrained views about oneself, others and the world. Core beliefs are deeply held views people hold about the world and their place in it. While core beliefs are usually formed early in life and are shaped by childhood and life events and experiences, dysfunctional assumptions include unrealistic, rigid or conditional standards of life, which result in PD (Fenn & Byrne, 2013). Automatic negative

thoughts comprise thoughts that arise spontaneously in particular contexts. In this study, the assessment of core beliefs included GHQ-12 measure items such as *I had thoughts of being a worthless person* and *I believe I am incapable of making decisions about things*. While dysfunctional assumptions included items such as *I felt capable of making decisions about things* and *I felt like I was losing confidence in myself*, automatic negative thoughts included items such as *I felt like I was not playing a useful part in things (society, school, work, family etc.)* and *I lost sleep over worry*.

Marciniak et al. (2020) found that the most often utilised CBT practices involve cognitive restructuring, which often includes methods such as thought reappraisal, self-monitoring, introspection, and relaxation techniques. Self-monitoring is the systematic observation and documentation of specific target experiences, such as an emotional reactions, distorted thoughts, and detrimental behaviours (Cohen et al., 2013). Certain CBT may emphasise self-monitoring of moods, thoughts, and emotions approaches for patients to identify problematic thoughts and emotions when they arise and recognise patterns between thoughts and emotions (Jarrett & Nelson, 1987 as cited in Cohen et al., 2013). Self-monitoring contributes to this goal by equipping clients with a simple approach for identifying thoughts, emotions, and behaviours through assignments. According to Cohen et al. (2013), self-monitoring, in combination with psychoeducation delivered in session, assists in clarifying and reinforcing the motivation and objectives of CBT. Additionally, techniques such as cognitive restructuring may be used to help patients reframe unpleasant experiences more objectively and realistically (Shurick et al., 2012). Other cognitive techniques used in CBT include assessing the evidence for a belief, distancing tactics, and identifying and questioning cognitive distortions (Beck, 2017). Cognitive distortions are erroneous beliefs that increase the potential of mental health problems such as psychological

distress (Dozois & Beck, 2008).

Cognitive distortions result from a person having a skewed automatic thought in response to a circumstance, which results in emotional and behavioural reactions. Automatic thoughts are often consistent with core beliefs about oneself, other people, and the world. When negative core beliefs and automatic thoughts emerge, a negative, neutral or positive situation may affect the emergence of negative emotions and maladaptive behaviours. Over time, this habitual pattern may contribute to developing and maintaining psychological distress symptoms (Rnic et al., 2016). Identifying and correcting cognitive distortions is a critical component of CBT (Fenn & Byrne, 2013; González-Prendes & Resko, 2012). Beck et al. (1976) defined cognitive distortions as dysfunctional thinking processes that result in problematic behaviour and emotions. Burns (1980) recognised further cognitive distortions, known as the ten cognitive distortions. These include mind reading, catastrophising, all-or-nothing thinking, emotional reasoning, minimising the positive, labelling, overgeneralisations, personalisation, should statements, and mental filtering (Rnic et al., 2016). The Woebot app has features to assist users in identifying cognitive distortions by asking what difficulty they are now experiencing, allowing them to pick the cognitive distortion that is most relevant at the moment, and challenging and rewriting the cognitive distortion more objectively (Darcy et al., 2021; Demirci, 2018; Fitzpatrick et al., 2017).

Conclusion

This chapter discussed the principles and theory of cognitive behavioural therapy as the study's theoretical framework. Firstly, a brief overview of the history of CBT was provided. Secondly, the main principles and methods of CBT practices were outlined. Thirdly, CBT based mental health apps and how CBT theory informs the study was addressed. Finally, Woebot's

CBT-based features were explored concerning how they might assist with psychological distress. The following chapter discusses the study's methodology.

CHAPTER 4

Methodology

Introduction

The research methodology that was employed in the study is presented in this chapter. The research design, including the reasoning behind selecting such, is first outlined. This is followed by an explanation of the participant sampling methods and data collection instruments that were employed. Subsequently, the data collection procedures and the Woebot intervention are described. Finally, the data analysis methods and the ethical considerations applicable to the study are discussed.

Research Paradigm of CBT Methodologies

Wagner et al. (2012) noted that the research process is guided by philosophical beliefs about the nature of reality, knowledge and values whereas the theoretical framework informs comprehension, interpretation, choice of literature and research practice on a given topic of study. Bannink (2017) highlighted that CBT research methodologies are typically based on a positivist paradigm in which human behaviour can be measured objectively through scientific inquiry. The foundation of positivism encompasses the notion that science is the sole source of true knowledge. It asserts that the methodologies, approaches and procedures employed in the natural sciences provide the strongest foundation for studying the social world (Wagner et al., 2012). Similarly, Chow (1992) emphasised much of what is known about cognitive processes from psychological research is derived from positivist-based methods such as experimental designs, quasi-experimental designs and evaluation studies. Furthermore, Beidas et al. (2015) emphasised the use of objective metrics and experimentation to guide CBT-based techniques.

Accordingly, positivism, which served as the researcher's epistemological paradigm and informed the study's research methodology is discussed in the next section.

Research Paradigm: Positivism

Positivism, which is commonly referred to as logical positivism, posits that the scientific method is the sole means of determining truth and objective reality (Park et al., 2020). According to Bogdan and Biklen (2003), Auguste Compte coined the concept positivism to describe a stringent empirical approach in which statements about knowledge are founded on direct experience; it prioritises facts and the causes of behaviour. Compte attempted to differentiate between empirical knowledge and knowledge derived from metaphysics or theology. He believed that scientific knowledge was more indicative of the truth than philosophical speculation. In general, positivism applies scientific approaches to examine human behaviour (Wagner et al., 2012). In modern times, positivism is viewed as being objectivist, which holds that the things around people exist and have meaning regardless of their awareness of them (Crotty, 1998, as cited in Wagner et al., 2012).

In relation to the nature of reality, positivists argue that there is a single tangible reality that is largely consistent throughout space and time (Wagner et al., 2012). This view of reality is referred to as naïve realism, which is part of the researcher's responsibility to uncover. Positivists hold that reality is independent of researchers' interest in it and is essentially objective. They also believe that reality is measurable and can be divided into variables (Wagner et al., 2012). Accordingly, the researcher's view was that the nature of reality is objective and can be measured and operationalised into variables.

According to positivism, the paradigm of natural science is fundamental to the nature of knowledge (Wagner et al., 2012). Positivists define knowledge as assertions, beliefs and truths

that can be scientifically evaluated, confirmed and verified or disproved, and are stable and generalisable. The objective nature of knowledge makes it independent of researchers' values, interests and emotions. Positivists hold that to realise absolute truth for a specific inquiry, researchers merely need the appropriate data collection equipment and/or methods (Wagner et al., 2012). This usually consists of quantitative research methods, which may encompass experimental, quasi-experimental, correlational, causal-comparative and survey designs (Jhangiani et al., 2019) . Furthermore, the data collection procedures typically comprise questionnaires, observations, tests and experiments (Jhangiani et al., 2019). When employing the positivist framework, the objective of a study is to find the governing rules and principles of the universe and predict behaviours and conditions (Wagner et al., 2012). According to positivists, all research must be value-neutral. Consequently, to attain objectivity and neutrality throughout the investigation process, researchers should employ scientific methods for collecting data (Wagner et al., 2012). Accordingly, the researcher aimed to be as value-neutral as possible in this study by employing scientific methods such as questionnaires and scales for data collection.

Positivist Research Methodology

Whereas research methodology encompasses the combination of assumptions about the nature of reality and knowledge, values, theory and practice on a given topic (Wagner et al., 2012), methods are the means utilised for gathering data and are an important part of the methodology. The objectives of a positivist-based research methodology may include predicting outcomes, testing a hypothesis, determining the strength of relationships between variables and establishing a cause-and-effect relationship between variables (Wagner et al., 2012). Employing this paradigm, the aims of this study included testing a hypothesis, determining the strength of

the relationship between the Woebot intervention and PD and endeavouring to demonstrate a causal connection between the intervention and outcome variable.

Positivism is primarily concerned with objectivity and quantifiability; hence, quantitative research designs are emphasised (Park et al., 2020). Quantitative research starts with ideas, hypotheses and/or concepts that are defined as they are applied in the study and employed to identify the variables of importance. It is imperative that the problem statement specify the variables to be investigated and their relationship. In addition, variables are operationally defined so that other researchers may reproduce, evaluate and validate the findings. Operationally defining a variable involves specifying its characteristics based on how it is utilised, measured and observed in the research (Wagner et al., 2012). Accordingly, the emphasis of this study was on quantitative research and operationalising important variables to extract and test a hypothesis for the results to be replicated or verified. Typical research approaches in the positivist paradigm include experimental, quasi-experimental, correlational, causal-comparative, quantitative and randomised controlled trial designs. Moreover, instruments for data collection typically include questionnaires, observations, experiments and tests. In this study, a quasi-experimental approach was employed and data collection methods such as questionnaires utilised (Appendix G).

Research Approach, Strategy and Design

A quantitative research approach was adopted. Wilson (2018) explained that the primary factor that determines the choice of research methodology is the research question. Hence, the decision to use a quantitative research approach was influenced by the nature of the research question: What is the effect of Woebot on the psychological distress (PD) levels of second- and third-year university students at a South African university? When selecting a research design,

accessibility, availability, resource limitations, academic practices and ethics also play a role in selecting a research methodology (Wilson, 2018).

Payne and Williams (2011) stated, "Quantitative methods of social research involve, on the one hand, counting and measuring that human behaviour which is quantifiable, and on the other hand, applying these data as evidence in the interpretation and analysis of the issues addressed by the various social sciences" (p.3). In other words, quantitative research involves measuring and analysing variables using statistical methods to obtain relevant data that inform the research question (Apuke, 2017). In this study, the variables related to the research question were quantified to determine whether a cause-and-effect relationship existed between them. Wilson (2018) noted that cause and effect can be established by analysing the relationship between an independent variable (cause) and a dependent variable (effect) as well as the mechanisms through which the independent variable affects the dependent variable. In addition, predictions can be made about the relationship between variables (research hypotheses) by employing deductive reasoning and theoretical inference (Wilson, 2018). Therefore, a quasi-experimental strategy was deemed fit to determine whether there is a cause-and-effect relationship between the Woebot app intervention and PD rather than just a correlation (Wilson, 2018).

The purpose of quasi-experimental approaches is to establish whether a cause-and-effect relationship exists between variables, with the caveat that participants are not randomly assigned to groups (White & Sabarwal, 2014; Zwarun, 2018). Random allocation was impractical for this study because some of the participants were not willing to complete the three-week Woebot app intervention but were willing to be in the comparison group that received no treatment. Consequently, participants were not assigned randomly to groups. Rather, based on their level of

commitment and availability, they were assigned to the treatment or comparison group.

Another reason for selecting a quasi-experimental strategy was related to the nature of the MHapp intervention. Baumel et al. (2020) noted that quasi-experimental designs are more suitable for technology-related research because they are less time-consuming and expensive than randomised controlled trials (RCTs). In addition, the practicality of quasi-experimental designs is more apt for technologies that evolve and change rapidly such as MHapps (Kumar et al., 2013). However, it is noteworthy that the major drawback of quasi-experimental designs is the lack of randomisation, which limits the extent to which causal relationships can be determined (Baumel et al., 2020; Schweizer et al., 2016). This shortcoming is discussed in further detail in the limitations section of the final chapter.

For this study, a nonrandomised comparison group pretest-posttest design was chosen. This quasi-experimental design comprises a treatment and a comparison group to which participants have not been randomly allocated (Ormrod & Leedy, 2015). The participants were allowed to choose whether they wanted to be in the treatment or comparison group in the pre-study survey (Appendix G). However, non-equivalent groups were established based on shared characteristics such as age range, year of study and baseline PD scores to compromise for the lack of randomisation.

In the following section, the procedures followed during the participant recruitment process are outlined.

Participant Recruitment Procedures

Before initiating the recruitment process, institutional authorisation and permission to conduct the study and recruit participants were obtained (Appendix A). Participant recruitment

occurred from the beginning of March 2022 until April 2022. The sample comprised second-and third-year students from a South African university. This sample was selected in accordance with previous research that revealed that second-and third-year student samples had similar general psychological distress levels across multiple study samples (Chernomas & Shapiro, 2013; Hafen et al., 2006; Tchounwou, 2004). This demographic was also chosen because of similar age ranges and familiarity with smartphone app technology (Fitzpatrick et al., 2017; Kajitani et al., 2020; Kretzschmar et al., 2019; McCloud et al., 2020).

The initial phase of participant recruitment occurred via social media platforms, namely, Facebook and LinkedIn. First, a Facebook group was created to advertise the study for second- and third-year students at the researcher's university. Second, a LinkedIn page was created to advertise the study group page to those in the researcher's network. The research flyer and informational descriptions about the study were posted on the LinkedIn page, with a link to the pre-study survey. Third, WhatsApp groups were were also used to advertise the study. Electronic flyers of the study (Appendix C) and informational posts were published in the groups as well as groups affiliated to the researcher's university. The informational posts contained a brief description of the study and a link to the pre-study survey. The pre-study survey contained the participant information sheet, demographic questions and the GHQ-12 psychological distress measure (Appendix G).

The second recruitment phase was conducted on campus. It involved advertising the study to second- and third-year students present on campus during the recruitment period. Because there were still COVID-19 measures in place during March 2022, there was a limited number of students on the campus. The on-campus recruitment process involved displaying the study poster to students and informing them briefly about the nature of the study. The posters

contained a QR code, which could be scanned with the individual's smartphone camera and would direct them to the pre-study survey on Qualtrics if they were interested in participating. In addition, respondents were informed that they would be emailed participation instructions if they qualified.

The intention of the initial recruitment plan was to recruit participants through social media solely. However, despite several posts and communications on multiple social media platforms, the target sample remained largely unresponsive. This problem has been highlighted in previous studies that recruited participants similarly by using online-only formats. Linardon and Fuller-Tyszkiewicz (2020) noted that in studies where participants signed up online without making personal or phone contact with the researcher, the attrition rates were much higher in comparison to cases where the researcher made some form of interpersonal contact with the respondents. They further explained that this may be because participants who sign up online may be less motivated to continue participating in the study in the ensuing days and weeks due to the effort. In contrast, when researchers contact respondents personally, it may afford an enhanced opportunity to explain the study and motivate participants to continue participating during more extended periods (Eysenbach, 2005; Linardon & Fuller-Tyszkiewicz, 2020). Linardon and Fuller-Tyszkiewicz (2020) thus suggested that it could be helpful for experimental studies on MHapps to incorporate more personalised study enrollment methods.

Sampling Procedures

Purposive sampling, which is a type of nonprobability sampling method, was employed to recruit students through the previously mentioned mediums of Facebook, Linkedin, Whatsapp and face-to-face. Nonprobability sampling is a method for recruiting participants without random

selection (Bigsby, 2018). While random selection implies that every member of a population has an equal chance of being recruited, non-probability sampling omits some population segments from possibly being recruited (Bigsby, 2018). Nonprobability sampling is often employed when the researcher does not have complete access to the whole population. Rather, the researcher uses a range of alternative sampling procedures to limit the list of prospective participants for the study (Bigsby, 2018).

Purposive sampling, also known as judgement sampling, is regarded as the most useful form of non-probability sampling. Researchers rely on their own experience, previous research and/or ingenuity to find participants who are considered representative of the population. Furthermore, they typically employ specific selection criteria to identify the most suitable individuals (Etikan, 2016; Wagner et al., 2012).

Although purposive sampling is typically associated with qualitative research, Palinkas et al. (2015) noted that quantitative data can be generated from a purposive sampling strategy. It is imperative that the data is suited to the specific research objective and assumptions and requirements associated with the research methodology (Palinkas et al., 2015). A primary benefit of purposive sampling is that it assists researchers to justify generalisations from the sample they are studying (Sharma, 2017). One major limitation of purposive sampling is that it may be highly prone to researcher bias (Sharma, 2017).

The inclusion criteria of this study were as follows:

- Second- or third-year student;
- Age 18 to 24 years;
- PD cut-off score of 9 or above on the GHQ-12;
- Currently enrolled at the researcher's university;

- Owned a smartphone with Android or IOS operating systems; and
- Daily internet access.

Sample Size Estimate

The preliminary sample size estimate was 64 participants. The following equation was used to calculate the sample size for two independent groups: 16∕(estimated effect size)*squared* (Allen, 2011). The estimated effect size was set at 0.50 (medium effect), the level of significance at 80% and the level of significance at 5%.

Treatment and Comparison Group Assignments

As noted previously, the sample comprised a treatment and a comparison group to which participants were not randomly allocated. Rather, based on their level of availability and commitment, they were allowed to choose whether they wanted to be in the treatment or comparison group in the pre-study survey (Appendix G). The treatment group received the Woebot app intervention. Participants were instructed to interact with the Woebot app daily or for as many days as possible during a three-week period. The interaction entailed checking in and completing the activity or learning material prescribed for that day, which usually lasted between 5 and 15 minutes. Checking in involved opening the app and selecting the current mood represented by the list of emojis. Three options were generally available for interaction: (a) chatting about a problem, which directly assisted anxiety, depression, loneliness, procrastination, etc.; (b) learning something new in that Woebot provides engaging and practical short lessons on various topics related to mental wellbeing; and (c) tracking and journaling to allow users to record their current mood and track their patterns over time.

In addition, the participants in the treatment group were asked not to use any other mental health apps besides the Woebot app throughout the study. These restrictions were imposed to avoid confounding study findings related to mental health app usage. At the end of each week, the participants in the treatment group were sent a short survey in which they were required to indicate how many days they had used the Woebot app during the previous week. They could choose to record this manually with an attached usage log that accompanied the instructional document or they could use the mood tracker in the Woebot app, which indicated how many days users checked in with the app. Moreover, no personal information or exact app usage details were requested.

The participants in the comparison group received no assigned MHapp intervention and were asked to refrain from using any mental health apps throughout the study. Furthermore, those in this group were emailed instructions that indicated that they would receive a follow-up survey at the end of the three-week intervention period (Appendix F).

The General Health Questionnaire-12 (GHQ-12)

Goldberg developed the General Health Questionnaire (GHQ) in 1972 in an endeavour to develop a screening measure for detecting psychological distress (Cuéllar-Flores et al., 2014). The GHQ is accepted globally as the standard measure for psychological distress assessment. It has been employed as the primary measure of PD (Furukawa et al., 2003, as cited in Drapeau et al., 2012). The GHQ-12 (Appendix G) is commonly used for non-clinical and epidemiological assessments of psychological distress (Drapeau et al., 2012). The GHQ-12 has high construct validity and reliability for assessing PD in university populations (Yaghubi et al., 2012).

Furthermore, the GHQ-12 has been employed as an outcome measure for PD in mental health app studies (Kajitani et al., 2020).

With regard to the psychometric statistics of the GHQ-12 among young adults in Africa, Gelaye et al. (2015) reported the GHQ-12 had an internal consistency of 0.83 when employed to screen psychological distress symptoms in young Ethiopian adults. They also revealed a comparative fit index (CFI) of 0.951 and root mean square error of approximation (RMSEA) of 0.050, which are in the acceptable range. De Kock et al. (2014) applied the extended version of the GHQ-12, the GHQ-28, to a black South African sample and found internal consistency reliability estimates (Cronbach's alpha) ranged from .70 to .83. Furthermore, Goldberg et al. (1997) found the validity coefficient of the GHQ-12 was 0.88, extrapolated across 15 nations worldwide.

The GHQ-12 includes the following items: ability to concentrate; lost sleep over worry; playing a useful part in society; capable of making decisions; constantly under strain; cannot overcome difficulties; enjoying normal activities; addressing problems; unhappy and depressed; losing confidence in yourself; thinking of yourself as worthless; and feeling reasonably happy. The items use a 4-point severity/frequency scale (0–3) to rate the extent to which respondents have experienced each symptom during the previous two weeks. The expressions *recently* and *during the last few weeks* are used occasionally instead of the two-week reference period. In addition, the item scores can be added to create a total psychological distress score. The GHQ scales have been validated in clinical (Segopolo et al., 2009) and non-clinical samples (Nerdrum et al., 2006). The GHQ-12 measures generate a total score, which ranges from 0 to 36. The positively phrased items (items 1, 3 ,4 ,7 ,8 ,12) are scored from 0 (better/more than usual) to 3 (much less than usual) and the negatively phrased items (2, 5, 6, 9, 10, 11) from 3 (much more

than usual) to 0 (not at all). Various cut-off score ranges have been proposed for the GHQ-12, with cut-off scores of 15 or higher usually reserved for screening for the presence of a psychiatric disorder (Centofanti et al., 2019; Endsley et al., 2017; Kim et al., 2013). However, in accordance with an exploratory factor analysis study of the GHQ-12, a score of 9 was set on the GHQ-12 Likert scale to screen for psychological distress. A score of 9 was determined to be optimal to screen for psychological distress among university student populations (Yaghubi et al., 2012).

Demographics Questionnaire

The Qualtrics demographic questionnaire was used to collect information on the respondents' year of study, faculty, age, smartphone ownership, internet access, email address, gender and racial category (Appendix G).

Data Collection Procedure Flow

Once the respondents had completed the pre-study survey that comprised the demographic questionnaire and GHQ-12, their responses from Qualtrics were collected and organised in Excel. The data were used to determine those eligible for participation based on the inclusion criteria. Those who met the inclusion criteria were divided into those that chose to be part of the treatment group and those that chose to be part of the comparison group.

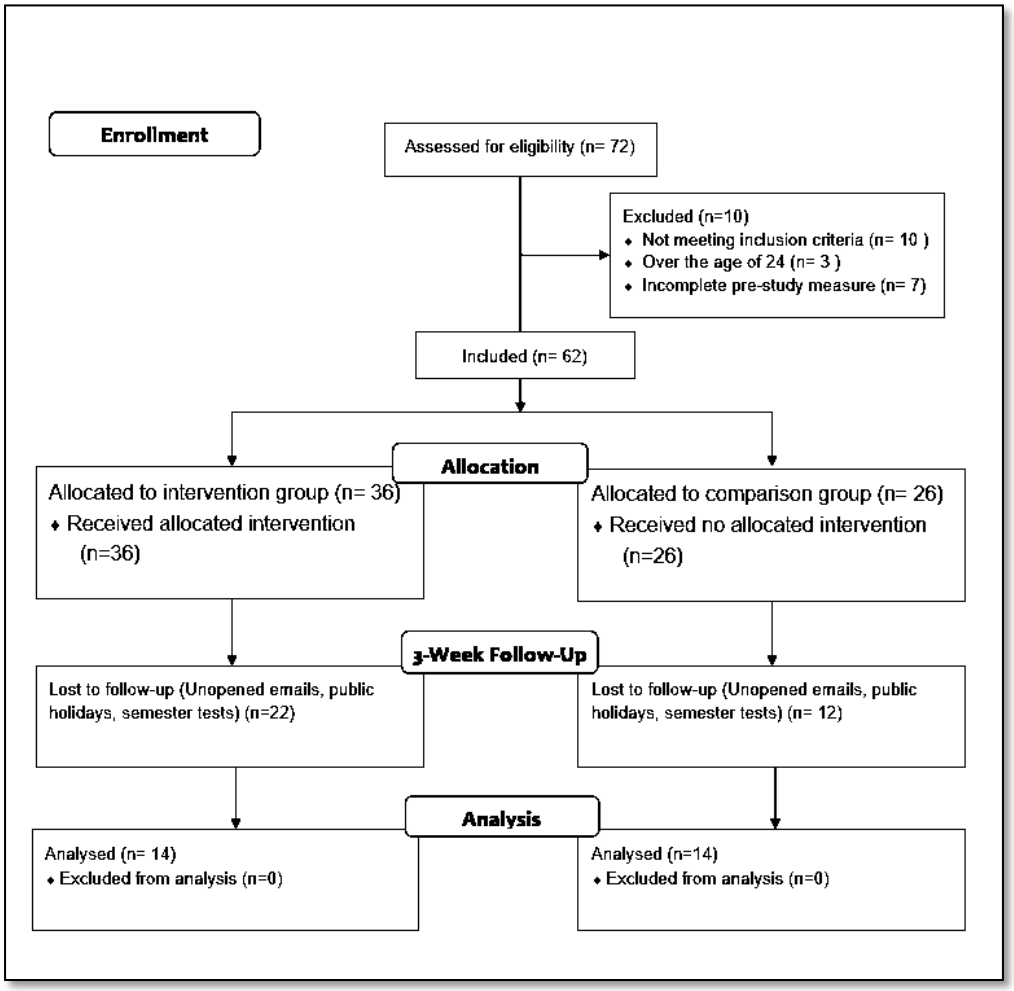
The participants in the treatment group were sent an email a few days after completing the survey, which contained instructions for downloading and using the Woebot app (Appendix F). The starting date of the intervention, data collection of the app and a link to Woebot's privacy policy were also outlined in the instructions. Moreover, the email contained an attached informed consent form linked to a Qualtrics survey (Appendix I).

After the first week of the study, each participant was emailed a Qualtrics survey to record the previous week's app usage data and indicate whether they had used other mental health apps besides Woebot (Appendix J). Their weekly usage data were recorded in an Excel spreadsheet. This weekly data capture process was followed for the first two weeks of the study. After the last week of the study, the participants in the treatment and comparison groups were emailed a link to the post-study surveys.

The post-study survey for the treatment group was used to collect information about the previous week's Woebot usage and the post-study GHQ-12 PD measure (Appendix H). Similarly, the post-study measure for the comparison group was employed to collect data about their PD during the previous few weeks. They were also asked if they had used any mental health or wellbeing apps during the previous three weeks.

Figure 3

Study Consort Diagram



Data Analysis

All data were analysed using International Business Machines (IBM) SPSS version 28. Participants were excluded from the analysis if they: (a) omitted information from their baseline measurements; (b) did not complete the post-study survey; and/or (c) had not used Woebot at all based on their weekly interaction report (Bakker et al., 2018).

The statistical analysis procedures included descriptive and inferential statistical methods. The descriptive statistics were computed by using SPSS 28. Descriptive statistics are employed to organise or summarise a collection of data. Percentages, measurements of central tendency (mean, median, mode), measures of dispersion (range, standard deviation, variance) and

correlation coefficients are some examples of descriptive statistics (Jhangiani et al., 2019, p. 49). The norm, average and centre of a set of scores are described by employing measures of central tendency. In relation to distribution, while the mode is the most frequent score, the median is the middle point of the score distribution and the mean is average set of scores (Jhangiani et al., 2019; Ormrod & Leedy, 2015).

Dispersion measures are also classified as descriptive statistics. They describe how widely a set of scores varies (Jhangiani et al., 2019). While the range is a measure of spread that determines the distance between a distribution's maximum and minimum scores, the standard deviation is a more advanced measure of the spread that calculates the average distance between scores and the mean (Jhangiani et al., 2019). The standard deviation squared is referred to as the variance. The variance is used to calculate the distance between scores and the mean but in a different unit of measurement (Jhangiani et al., 2019).

On the contrary, inferential statistics are employed to draw inferences about larger populations by using data from smaller samples (Ormrod & Leedy, 2015). Inferential statistics were utilised to determine whether Woebot's effect on PD could be extrapolated to the general population of university students. However, before performing inferential statistics, the data distribution was checked for normality and the 3 sigma rule using a t-test. A histogram showing the data distribution curve is displayed in the following chapter. In addition, the psychometric properties of the data were evaluated using Cronbach's alpha. In the following section, the inferential statistical measures used for the study are detailed: Spearman's rank order correlation and the independent samples t-test.

Spearman's Rank-Order Correlation

To evaluate the relationship between weekly average Woebot interaction days over three weeks and psychological distress change scores from the pretest to posttest, a Spearman's rank-order correlation test was performed. According to Laerd Statistics (2018), Spearman's rank-order correlation or Spearman's correlation calculates a coefficient to measure the strength and direction of the relationship between two continuous or ordinal variables. Ordinal variables have at least two categories organised in a particular order or rank. These categories can be ranked, but they cannot be assigned values in a manner that establishes a consistent difference between them as would be possible with a continuous variable (Laerd Statistic, 2018). Continuous variables can be measured by employing either intervals or ratios (Ormrod & Leedy, 2015). Interval variables are quantifiable and can be measured on a continuum, for example, the temperature measured in Celsius or Fahrenheit degrees. Ratio variables are interval variables with the additional condition that a measurement of 0 (zero) indicates the absence of such a variable (Laerd Statistic, 2018).

Spearman's correlation was selected instead of Pearson's correlation coefficient, which evaluates the strength and direction of a linear relationship between two continuous variables (Laerd Statistics, 2018). Since the exact number of interactions within a particular day and interaction durations were not recorded for this study, the average weekly Woebot interactions were measured as an ordinal variable, ranging from 0 to 7 days per week. The number of weekly Woebot interactions for each participant was obtained by averaging the number of interaction days over three weeks. To complete an interaction day, one of three options was given: (a) chat about a problem, (b) learn something new, or (c) track and journal (Appendix F). At the end of each week, the participants were asked to report the total number of interaction days indicated by the mood tracker on the app. Changes in the psychological distress pre-and posttest scores were

measured as a continuous variable, specifically an interval scale. The dependent variable change score was calculated by computing the difference between the pre-test and post-test GHQ-12 measures in SPSS.

Statistical assumptions of Spearman's correlation. To perform a Spearman's correlation, a few assumptions about the study design and the data must be met (Laerd Statistics, 2018). The first assumption is that there must be two variables measured on a continuous and/or ordinal scale; that is, there can be (a) two continuous variables, (b) two ordinal variables or (c) one continuous and one ordinal variable. The second assumption is that the two variables have to correspond to pairs of observations. A single paired observation reflects a single participant's score on each variable. As 28 participants were included in the final statistical analysis, this equated to 28 paired observations. The third assumption is that the two variables must have a monotonic relationship. This assumption can be tested by creating a scatterplot and inspecting the graph visually (Laerd Statistics, 2018).

Null Hypothesis Significance Testing for Spearman's Correlation. The null and alternative hypotheses for Spearman's correlation may be expressed in various ways. First, the methods by which these hypotheses are articulated depend on how broad or specific the researcher wants the hypotheses to be (Conover, 1999). Because Spearman's correlation is only specific for a monotonic relationship, more general assertions might lead to some typical misconceptions of Spearman's correlation. Therefore, Conover (1999) recommended that a hypothesis should be specific rather than broad.

Sprent and Smeeton (2007) stated that the most precise null and alternative hypothesis formulation is when both hypotheses are stated related to Spearman's population correlation

coefficient (ρs). As per Laerd Statistics (2018), describing the hypothesis in this manner minimises the risk that Spearman's correlation results will be misinterpreted.

As such, the null hypothesis of this test is generally stated as follows:

H0: $\rho s = 0$

This means that in the null hypothesis, Spearman's correlation coefficient is equal to 0 in the population. In other words, there is no correlation between Woebot interaction and psychological distress in the real population. The absence of any relationship between the ranks is shown by a correlation coefficient of zero (0).

The alternative hypothesis is

HA: $\rho s \neq 0$

This means that in the alternative hypothesis Spearman's correlation coefficient is not equal to 0 in the population. In other words, there is a correlation between Woebot interaction and psychological distress in the real population. The presence of a relationship between the ranks is shown by a correlation coefficient greater than zero (0).

Following the hypothesis statement, a p-value is computed and compared with a pre-defined cut-off value in order to reject the null hypothesis and accept the alternative hypothesis. The predefined cut-off value was .05 for this study, which means that if the p-value was less than .05, the alternative hypothesis, namely, that a correlation existed between Woebot interaction and PD, would be accepted and the null hypothesis, namely, that no correlation existed between Woebot interaction and PD, would be rejected.

In practice, SPSS Statistics automatically computes a p-value for Spearman's correlation, which is the default when using SPSS Statistics.

Spearman's Correlation Study Design Implementation. According to Laerd Statistics (2018), Spearman's correlation can be employed in conjunction with other statistical tests, as was the case in this study, specifically calculating a change score between two-time points and then using Spearman's correlation to determine whether there is a relationship between the change score and another variable. Consequently, Spearman's correlation was utilised to determine whether the PD change score, that is, the average change in GHQ-12 scores from pretest to posttest, and the participants' average number of Woebot interaction days were correlated.

Interpreting Results for Spearman's Correlation. Spearman's rank-order correlation coefficient value (rs or ρ) represents the strength and direction of the relationship between two variables (Laerd Statistics, 2018). The correlation coefficient can have any value between +1 and -1, where +1 represents a perfect positive relationship of ranks and -1 represents a perfect negative relationship of ranks. The absence of any relationship between the ranks is shown by a correlation coefficient of zero (0).

Unlike Pearson's correlation, there are no guidelines for assessing the strength of the relationships between different variables when employing Spearman's correlation test. However, the stronger the relationship between the variables, the closer the correlation coefficient is to +1 or -1. Furthermore, the closer the correlation coefficient is to 0, the weaker the relationship between the variables (Laerd Statistics, 2018).

The second stage in interpreting the results, which are presented in the following chapter, involved detecting the statistical significance of the Spearman rank-order correlation coefficient. This statistical procedure determined whether the null hypothesis could be accepted. Given that α was set at 0.05, a statistically significant Spearman rank-order correlation would indicate that, if

the null hypothesis was supported, there would be less than a 5% chance that the strength of the relationship occurred by chance (Laerd Statistics, 2018).

Independent Samples T-Test

The independent samples t-test determines whether the means of two independent groups differ on a continuous dependent variable (Ormrod & Leedy, 2015). This test is beneficial to determine whether the difference between two groups is statistically significant (Laerd Statistics, 2015). In this study, the independent-samples t-test was used to determine if the mean change in psychological distress scores from the pretest to posttest differed between the treatment and comparison groups.

Assumptions of the Independent-Samples T-Test. According to Laerd Statistics (2015), an independent-samples t-test requires six assumptions. While the first three assumptions involve the study design and measures used, the second three relate to the data obtained. Specifically, the first assumption is that the dependent variable must be quantifiable on a continuous scale. The PD pretest to posttest change score was measured at a continuous level by employing the GHQ-12. The second assumption is that an independent variable comprises two categorical, independent groups. The categorical independent variable in this study was the two treatment groups, namely, the treatment group that received the Woebot app and the comparison group that did not receive the Woebot app.

The third assumption is that observations should be independent, which implies that there should be no association between the observations in each group of the independent variable or between the groups themselves. In an independent samples t-test, independent groups are those with no association between the participants, which is achieved by having different participants

in each group. Because no participant was in the treatment group and the comparison group simultaneously, an independent group was established. In addition, given that the study sample included second-and third-year students from diverse faculties, it was reasonable to assume that most of the participants were not directly related. Moreover, the study was conducted online and there were no group interactions or briefings between the researcher and participants.

Independent-samples t-tests must also meet three data-related assumptions to be valid. First, there should be no outliers in the two independent variable groups for the dependent variable. Second, the dependent variable should be approximately normally distributed for each independent variable group. Third, the variance of the dependent variable should be equal in each independent variable group, that is, variance homogeneity.

Significance of Studying the Differences between Two Independent Groups. According to Ormrod and Leedy, (2015), the importance of a t-test for independent samples is to determine whether the difference between two samples means is due to sampling variance or whether there is an actual difference in the population. In other words, the independent sample t-test determines whether it is probable that the population means of the groups are distinct, not simply the sample means.

Null and Alternative Hypotheses. The null hypothesis for an independent-samples t-test is generally stated as:

H_0: the population means of the two groups are equal (i.e. $\mu_1 = \mu_2$)

The alternative hypothesis is

H_A: the population means of the two groups are not equal (i.e. $\mu_1 \ \mu_2$)

An independent-samples t-test determines the probability that the sample group means are at least as divergent as the results indicate, provided the null hypothesis has been supported. If the probability ($p < .05$) is small enough, the alternative hypothesis may be accepted and the null hypothesis rejected. Consequently, it is unlikely that the population means of the two groups are equal. However, if the probability is greater ($p > .05$), the null hypothesis is retained and the alternative hypothesis is rejected. This result shows that the two groups' population means are likely to be equal (Laerd Statistics, 2015).

Effect Sizes. As a null hypothesis significance test, although the independent-samples t-test indicates if the differences between group means are *real* in the population, it does not indicate the size of this difference (Laerd Statistics, 2015). Therefore, it is suitable to calculate the effect size to circumvent this limitation. In this study, the effect size measure (*d)* was used, which is one type of effect size that determines the significance of the independent variable by describing the difference between the group means as a ratio of the standard error of the mean difference (Laerd Statistics, 2015).

Sample Size for Independent Samples T-tests. For independent sample t-tests, the study should include six or more participants in each group so the results can be generalised to a broader population (Laerd Statistics, 2015). A balanced design is one in which each group has the same number of participants. Alternatively, if sample sizes are not the same in all groups, the design is unbalanced (Laerd Statistics, 2015). Generally, the more imbalanced the design, the higher the negative impact a violation of an assumption has on the test's validity (Laerd Statistics, 2015). Accordingly, a balanced design is preferable. In this study, 14 participants in each group were included in the data analysis.

According to Lakens (2021), resource and time constraints often have a direct impact on the amount of data that can be collected. Although larger sample sizes of 30 or more participants per group are typically viewed as optimal for data analysis, resource and time constraints greatly limited the number of participants that could be recruited during COVID-19 related campus closures within three weeks to compensate for the high attrition rates. Furthermore, previous MHapp research reported similar concerns with sample sizes that were smaller than expected due to time and resource limitations. For instance, Ly et al. (2017) had 14 participants in each group, and Aziz et al. (2022) had 30 total participants in their guidance-based apps study.

Study Design Implementation. An independent-samples t-test is commonly used to analyse the results of three types of study designs: (a) to determine whether there are differences between two independent groups, (b) to determine whether there are differences between interventions, and (c) to determine whether there are differences in change scores (Laerd Statistics, 2015). In each group, the same dependent variable is measured at the beginning (pre-study measure) and end (post-study measure) of the intervention and a change score is established (the post-values minus the pre-values). An independent-sample t-test was used to compare the change scores of the two groups to determine the difference in the change for each group. The primary question was whether there was a more significant change in PD levels in the treatment group than in comparison group. Any change between the two groups would, in theory, be the result of the intervention (Laerd Statistics, 2015).

Null and Alternative Hypotheses

The null hypothesis was thus formulated:

H_0: $\mu_{treatment} = \mu_{comparison}$

Treatment and comparison mean PD change scores are equal in the population.

The alternative hypothesis was formulated as follows:

H_A: $\mu_{treatment} \neq \mu_{comparison}$

Treatment and comparison mean PD change scores are not equal in the population.

Shapiro-Wilk Test for Normality. The Shapiro-Wilk test is recommended for small sample sizes of approximately 50 participants and when the visual interpretation of Normal Q-Q plots or other graphical representations used to test for normality are unsuitable (Laerd Statistics, 2015). The Shapiro-Wilk test determines whether data for each category of the independent variable is normally distributed. Accordingly, there are as many Shapiro-Wilk tests as there are independent variable categories (Laerd Statistics, 2015). Two tests were performed for this analysis, one for each category of the independent variable, the treatment and comparison groups.

Independent-Samples T-test Procedure

The SPSS Statistics procedure for performing an independent-samples t-test also computes and displays Levene's test for the equality of variances, which is used to test for the assumption of variance homogeneity. When the assumption of homogeneity of variances is met and violated, SPSS Statistics outputs two t-test results.

Assumption of Homogeneity of Variances. An important assumption of the independent-samples t-test is that the two groups' variances are equal in the population. Failure to adhere to this assumption generally increases the chance of making a Type I error. The equality of variances is often referred to as the homogeneity of variances (Laerd Statistics, 2015).

A crucial assumption of the t-test for independent samples is that the variances of the two groups are equivalent in the population. If this assumption is violated and variances are uneven, the likelihood of committing a Type I error increases (Laerd Statistics, 2015).

SPSS Statistics employs Levene's Test of Equality of Variances to determine whether variances differ in the population. In other words, Levene's test assesses whether the two samples were derived from populations with equal variance. If the population variances of both groups are equivalent, this test will provide a p-value greater than 0.05 (i.e., p >.05), demonstrating that the assumption of variance homogeneity has been met. However, if the test yields a p-value less than 0.05 (p .05), the population variances are uneven and the assumption of homogeneity of variances is violated (Laerd Statistics, 2015).

Ethical Considerations

Before commencing with a research study, the researcher must determine who must be informed about the study and obtain their authorisation to conduct research. Persons who grant researchers access to an organisation or community to do research are referred to as gatekeepers (Wagner et al., 2012). Therefore, before conducting this research study, the researcher acquired ethics approval from the Faculty of Humanities Ethics Board of the researcher's university on November 15, 2021 (Appendix A).

Throughout the research process, the researcher attempted to ensure that the benefits of the research project outweighed any risks to the participants or university. Moreover, the researcher endeavoured to act responsibly and with integrity and to conduct the research in a just manner in order to respect the participants' human rights and dignity (Jhangiani et al., 2019).

Protecting participants from harm was a key consideration throughout the research process. In addition, the study was designed in such a way as to avoid exposing participants to any needless physical and/or psychological risks by using a relatively unintrusive intervention in the form of an app, and allowing participants to leave the study if they found the Woebot app to be harmful in any (Ormrod & Leedy, 2015). For instance, the participants were provided access to a psychologist who was on standby in case any problems emerged throughout the intervention period (Appendix E). Moreover, the researcher did made sure not to embarrass, frighten, offend or harm participants. Further precautions were taken to avoid causing harm to the participants, including limiting time commitments, resolving their worries about confidentiality and avoiding humiliation and loss of pride or dignity by not disclosing anyone's identity (Wagner et al., 2012).

Traditional research approaches have often ignored power and justice concerns in the researcher-participant relationship (Wagner et al., 2012). In the context of research, power and justice refer to researchers treating participants fairly and not abusing their position to pressurise individuals into participating or responding in a specific manner. Consequently, the researcher did not exploit his position as a postgraduate students to pressurise undergraduate students to participate. Moreover, no participants were chosen in accordance with their socioeconomic status, race and/or minority status.

The second key ethical consideration that was adopted throughout the study was voluntary and informed participation. Participants were informed that if they provided consent to participate in the study, they still had the right to withdraw at any time during the study. Moreover, they were informed that any involvement in the study was entirely voluntary (Ormrod & Leedy, 2015). Accordingly, they were provided with voluntary and informed participation

documents (Appendix I). In addition, they were informed about the purpose of the research through the participation information sheet (Appendix D).

The right to privacy requires that any study that involves human participants should consider their right to privacy (Ormrod & Leedy, 2015). Throughout the research process, the researcher ensured that each participant's right to privacy was respected. In addition, how each participant reacted or behaved was not reported in the findings. Furthermore, the participants' actual identities were not revealed.

Deception was another ethical issue that was considered. Wagner et al. (2012) noted that deception may take several forms, including covert research conducted without the participants' awareness and attitudinal biases towards participants and/or findings. The researcher endeavoured to maintain neutrality in the data collection and analysis by not maintaining any predetermined biases and/or assumptions regarding the participants and research findings. Moreover, the researcher informed all the participants that he was conducting research to collect data for his dissertation.

Confidentiality and anonymity are two concepts that are inextricably connected in studies that involve human participants. Confidentiality refers to not disclosing a participant's information to others and presenting results in such a manner that individuals cannot be identified, mainly through anonymisation (Wiles et al., 2008). In other words, confidentiality is keeping participants' information and data during a research project completely secret and private. Anonymity ensures that a participant's name is not directly linked with their reported data (Gravetter & Forzano, 2012). Confidentiality also entails not revealing any information obtained from participants in ways that might be used to identify them, whether intentionally or unintentionally (Wiles et al., 2008). In this study, confidentiality of information was guaranteed

by confirming that only the supervisor and researcher had access to relevant data related to the participants' identities. Any results provided in the dissertation findings, conference proceedings and academic publications ensured that the participants remained anonymous by utilising code tags rather than names (Christensen et al., 2014; Gravetter & Forzano, 2012). Furthermore, the university's identity was not disclosed in the research findings. After the dissertation is published, the electronic data will be kept in a file with a password at the researcher's university at the Department of Psychology for at least 5 years. If applicable, future use of the stored data is subject to further research ethics review and approval.

Qualtrics and Woebot Security and Privacy

Although high-end firewall systems and scans safeguarded the confidentiality of all collected survey data on Qualtrics, checks were regularly done to guarantee that any vulnerabilities in the system are discovered and addressed swiftly (Qualtrics, 2022). Qualtrics also utilises Transport Layer Security (TLS) encryption (also known as HTTPS) for all transmitted data and surveys that can be protected with passwords. Moreover, Qualtric services are hosted by trusted data centres that are individually inspected using the industry-standard SSAE-18 method (Qualtrics, 2022). The anonymity of participant data was further reinforced through the enhanced anonymity and anonymity threshold functions on Qualtrics(Qualtrics, 2022). Furthermore, the participants' identities were not linked to the gathered data. Rather, code tags were prescribed on SPSS.

Although not required, Woebot typically asks users to provide their email addresses. Woebot also ensures appropriate security measures to protect personal data from any accidental or unlawful destruction, loss, alteration, unauthorised disclosure and access (WoebotHealth,

2022). Access to information stored on the Woebot app is provided only on a need-to-know basis to those whose roles require them to process personal data (WoebotHealth, 2022). The researcher ensured that Woebat's privacy policy was included in the intervention instructions document (Appendix F). Furthermore, participants were not required to provide their actual identities when registering for the Woebot app.

Conclusion

In this chapter, the methodology employed in the study was presented. The research design was explained as a quasi-experimental pre-test, post-test comparison group design. The participant sampling and data collection processes were outlined. Moreover, the data collection instruments and treatment conditions were described. Finally, the data analysis methods were discussed and the ethical considerations of the study presented. A summary of the results of the intervention are presented in the following chapter.

CHAPTER 5

Results

Introduction

In this chapter, the results of the statistical analysis methods overviewed in the previous chapter are presented. First, the descriptive statistics for the sample are revealed. Second, the various statistical reports related to the Spearman's rank-order correlation test are provided. Finally, the statistical results of the independent samples t-test are presented.

Research Questions and Objectives

The following research question was posed: What is the effect of a CBT-based mental health application (Woebot) on the psychological distress (PD) levels of university students in South Africa?

To answer the research question, the objectives of this study were as follows:

(1) To establish the direction, form and strength of the relationship between each participant's average number of weekly Woebot interactions and pretest posttest PD change scores by conducting a Spearman's correlation test.

(2) To determine whether there is a statistically significant difference between the PD change score means of the treatment group who used the Woebot app and comparison group who did use the Woebot app after three weeks by conducting an independent samples t-test.

Statistical Hypotheses

The research hypotheses of the study were:

Hypothesis 1

The null hypothesis (H0):

H0: $\rho s = 0$

Spearman's correlation coefficient between weekly Woebot interactions and PD change scores is equal to 0 in the population. This indicates that no correlation exists between the Woebot interaction and PD changes extrapolated to the population of university students.

The alternative hypothesis (H1):

HA: $\rho s \neq 0$

Spearman's correlation coefficient between weekly Woebot interactions and PD change scores is not equal to 0 in the population. This indicates that a correlation exists between the Woebot interaction and PD changes extrapolated to the population of university students.

In other words, the presence of any relationship between the Woebot interaction and PD extrapolated to the real population is demonstrated by a correlation coefficient above or below zero (0).

This would indicate that a correlation exists between the Woebot interaction and PD changes in the real population.

Hypothesis 2

The null hypothesis (H0) for the independent-samples t-test was:

H_0: the population means of the two groups are equal (i.e. $\mu_1 = \mu_2$)

In other words, there is no statistical difference between the means of the PD change scores of participants who received the Woebot intervention (treatment group) and those who did not receive any intervention (comparison group).

The alternative hypothesis (H1) was:

H_A: the population means of the two groups are not equal (i.e. $\mu_1 \neq \mu_2$)

In other words, there is a statistically significant difference between the means of the PD change scores of participants who received the Woebot intervention (treatment group) and those who did not receive any intervention (comparison group).

Sample Description

The final sample included 28 second- and third-year university students who were assigned non-randomly through self-selection to either the treatment or comparison condition. The average number of weekly Woebot interactions was four days in the treatment group. The average age of the participants in both groups was 20 years. The frequency tables of the descriptive characteristics of the sample are presented in the following section.

Age

Table 1

Frequency Table Showing Age Distribution of the Sample

Age		
	N	%
18	1	3.6%
19	5	17.9%
20	9	32.1%
21	5	17.9%
22	6	21.4%
23	2	7.1%

Note. The left column indicates the sample age range, the central column the number of participants and the right column the percentage of participants in a certain age bracket.

Gender

Table 2

Frequency Table Showing Gender Distribution of the Sample

Gender		
	N	%
Female	19	67.9%
Male	9	32.1%

Note. The left column represents the number of participants and the right column the percentage of the sample that fell into the particular gender bracket.

Faculty

Table 3

Frequency Table Showing the Faculty Distribution of Participants in the Sample

Faculty		
	N	%
EBI	7	25.0%
EDU	1	3.6%
EMS	2	7.1%
HUM	14	50.0%
NAS	4	14.3%

Notes. The left column denotes the faculties of the participants: EBI=Engineering, Built Environment and Information Technology; EDU=Education; EMS=Economic and Management Sciences; HUM=Humanities; and NAS=Natural and Agricultural Sciences. The central and right columns indicates the numbers and percentages of the sample in each faculty, respectively.

Racial Category

Table 4

Frequency Table Showing Racial Category Distribution of The Sample

Racial Category		
	N	%
African/Black	7	25.0%
White/Caucasian	21	75.0%

Note. While the left column denotes the racial categories of the sample, the central and right columns represent the number and percentage of participants in each racial category, respectively.

Year of Study

Table 5

Frequency Table Showing the Year of Study of the Sample

Year of study

	N	%
2nd	11	39.3%
3rd	17	60.7%

Note. While the left column denotes the year of study of the sample, the central and right columns represent the number and percentage of participants in each year of study, respectively.

Psychological Distress Scores

Table 6

Descriptive Table of the Population Mean Age, GHQ-12 Pre- and Post-Study Scores

Population Descriptive Statistics

	N	Mean	Std. Deviation	Variance
Age	28	20.571	1.294	1.673
GHQ Pre	28	16.786	4.601	21.168
GHQ Post	28	12.321	5.701	32.504
Valid N (listwise)	28			

Std. Deviation and Variance use N rather than N-1 in denominators.

Note. The left column denotes the sample's age and pre-study and post-study GHQ-12 scores. The (N) column denotes the sample numbers and the Mean column indicates the entire sample's mean age and GHQ pre- and post-study scores. The remaining columns indicate the standard deviation and variance of the ages and GHQ pre- and post-scores of the entire sample.

Table 7

Gender Differences in Psychological Distress Pre- and Post-Study GHQ-12 Scores

	Males vs Females AVG Pre and Post PD Scores	
	Avg PD Pre Scores	**Avg PD Post Scores**
Males	16	12,54
Females	17	9,4

Note. The left column denotes the gender categories of the sample. Avg PD Pre Scores refer to the mean pre-study GHQ-12 scores and Avg PD Post Scores refer to the mean post-study GHQ-12 scores of each gender group.

Weekly Average Woebot Interaction Days

Table 8

Frequency Table Showing Average Weekly Woebot Interaction Days for Entire Sample

avgweekly	N	%
0	14	50.0%
1	2	7.1%
2	1	3.6%
3	3	10.7%
4	2	7.1%
5	5	17.9%
7	1	3.6%

Notes. The left column denotes the average number of days per week of Woebot interaction over three weeks, ranging from 0 days for the comparison group to 7 days for the treatment group. The central and right column denotes the number and percentage of the participants in the sample, respectively in each average weekly usage bracket.

GHQ-12 Psychometric Properties Report

Cronbach's alpha was used to calculate the internal consistency of the GHQ-12. Lee Cronbach created alpha, which is stated as a number between 0 and 1, to offer a measurement of the internal consistency of a test or scale (Tavakol & Dennick, 2011). Internal consistency refers to how closely all the test items assess the same concept or construct. It is therefore linked to how

closely the test items are related to one another. Before a test is used for research or examination purposes, internal consistency should be assessed to ensure the validity thereof.

On the other hand, reliability estimates display the degree of measurement inaccuracy in a test (Tavakol & Dennick, 2011). This definition of reliability relates to the correlation of a test with itself. Cronbach alpha values of 0.7 or higher demonstrate acceptable internal consistency (Taber, 2018). The Cronbach alpha value calculated for the GHQ-12 was .716, indicating an acceptable internal consistency. The results of the Cronbach calculations are displayed in Table 9.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.716	.703	12

Table 9

The GHQ-12 scale Cronbach calculations

Item-Total Statistics					
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
1-Able to concentrate	37.5926	128.481	.101	.311	.720
2-Lost sleep over worry	37.9630	123.037	.333	.555	.707
3-Playing a useful part	37.7407	127.123	.190	.538	.717
4-Capable of making decisions	36.9630	116.806	.199	.489	.719
5-Constantly under strain	35.1481	102.208	.544	.635	.668
6-Couldn't overcome difficulties	36.7407	108.969	.344	.695	.699
7-Able to enjoy daily activities	36.2593	105.969	.396	.736	.691
8-Able to face up to problems	37.3333	119.077	.181	.566	.720
9-Feeling unhappy and depressed	36.1481	99.516	.494	.623	.674
10-Losing confidence	36.0000	103.846	.443	.406	.683
11-Thinking of yourself as worthless	37.0741	109.456	.310	.547	.705
12- Feeling reasonably happy	36.6667	99.154	.644	.684	.652

Convergent Validity and Discriminant Validity of the GHQ-12

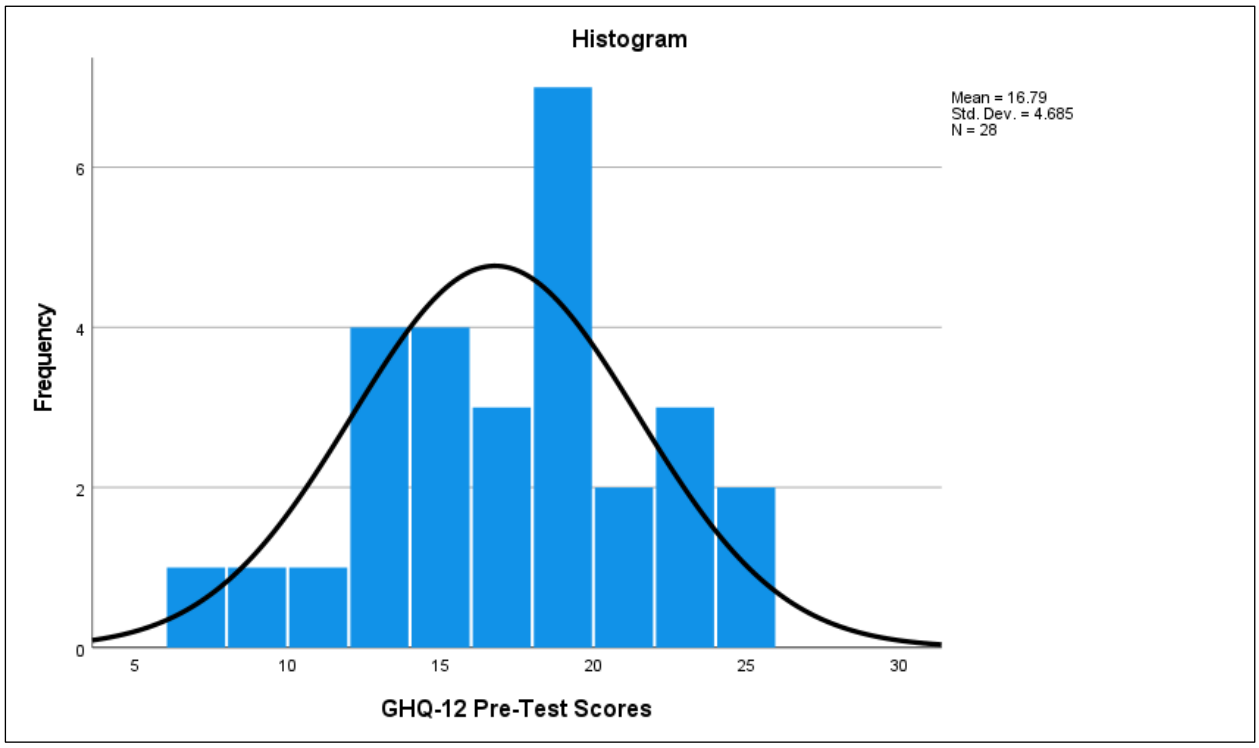
While convergent validity is the degree to which the same trait is measured by different methods, discriminant validity is the degree to which traits are different (Hill & Hughes, 2007). Convergent validity is usually demonstrated by an average factor loading of 0.7 (DeCoster, 1998). By employing principal component factor analysis, the convergent validity for the GHQ-12 was established with average factor loadings of 0.73. To establish discriminant validity, average variance extracted (AVE) analysis was conducted to determine if the square root of every AVE value belonging to each latent construct was larger than the correlations among any pair of latent constructs. Discriminant validity was established, with the square root of the AVE values larger than the correlations of the latent constructs (.56 > 0,01) (Hill & Hughes, 2007).

GHQ-12 Scale Data Distribution Properties

The GHQ-12 scale data were examined to determine whether it met the features of a normal distribution and was thus a bell shape, with scores clustered in the centre rather than the tails.

Figure 4

A Histogram Showing The Data Distribution of GHQ-12 Baseline Scores of the Entire Sample



Notes. GHQ-12 pretest scores of the sample score had a normal distribution, as demonstrated by a normal curve.

Spearman's Rank-Order Correlation Results

The Spearman's correlation test results revealed a weak negative correlation between the average number of days per week the participants used the Woebot app and PD change scores over three weeks, r_s = -.364. In other words, PD levels decreased as the number of days spent using the Woebot app increased. The strength of the correlation was weak, as indicated by the value of the Spearman correlation coefficient.

Determining Statistical Significance

Spearman's rank-order correlation was performed to assess the relationship between the weekly average number of Woebot usage days and PD change scores over three weeks. The final analysis included 28 participants. Preliminary analysis showed the relationship to be monotonic, as assessed by visual inspection of a scatterplot. In addition, there was a statistically non-significant correlation between the weekly average number of Woebot interaction days and PD change scores, $r_s(26) = -.364$, $p > .005$.

Null and Alternative Hypothesis Report

The results of Spearman's correlation test showed that there was a relationship between weekly Woebot interaction days and PD change scores although it was not statistically significant, rs(26) = -.364, p = .057. Since p > .05, the null hypothesis was retained.

Spearman's Rank-Order Correlation Test Main Findings Report

A Spearman's rank-order correlation test was performed to assess the relationship between the participants' average number of weekly Woebot interaction days and PD change scores over three weeks. The participants included 28 second- and third-year students. Preliminary analysis showed the relationship was monotonic, as assessed by a visual inspection of a scatterplot. The Spearman's correlation test revealed that although there was a relationship between weekly Woebot interaction days and PD change scores, it was not statistically significant, rs(26) = -.364, p = .057.

Independent Samples T-Test Results

Data Outlier Report

Figure 5

A Boxplot Showing The Mean Distribution Of Psychological Distress Change Scores Of The Treatment And Comparison Groups

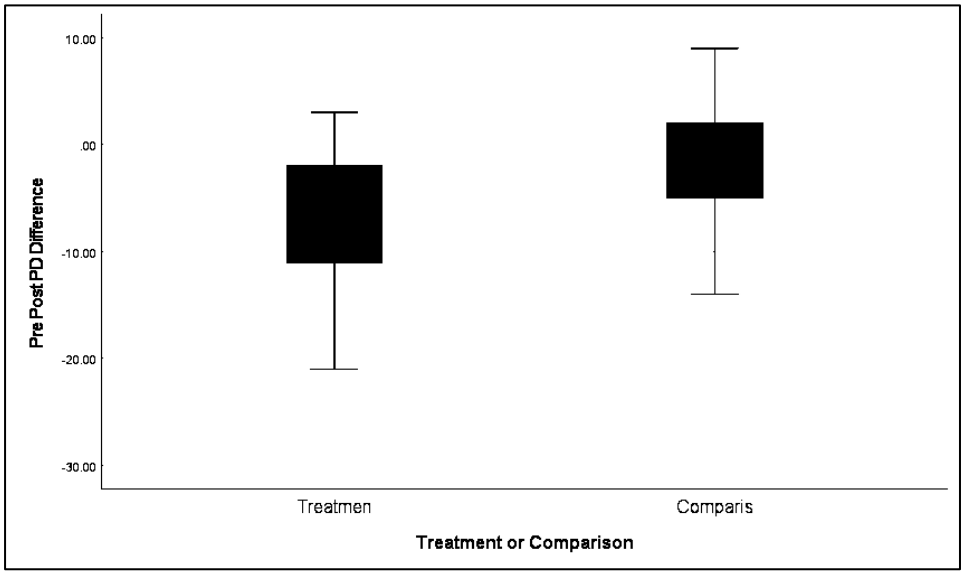


Figure 5. The data had no outliers, as assessed by an inspection of the boxplot.

Table 10

Table Showing Data Distribution Using the Shapiro-Wilk Test for Normality

Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Treatment or Comparison	Statistic	df	Sig.	Statistic	df	Sig.
Pre Post PD Difference	Treatment	.156	14	.200*	.953	14	.608
	Comparison	.175	14	.200*	.954	14	.632

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Notes. Psychological distress (GHQ-12 score) was normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$).

Descriptive Group Statistics Report

There were 14 treatment group participants and 14 comparison group participants. The treatment group's PD (GHQ-12 score) changes from pretest to posttest (M = -7.00, SD = 6.78) were higher than the comparison group's PD score changes from pretest to posttest (M = -1.92, SD = 6.47).

Table 11

Table Showing Psychological Distress (GHQ-12) Scores Difference Between the Treatment and Comparison Group pre and post study

	Woebot vs Comparison AVG Pre and Post PD Means	
	Pre PD Scores	**Post PD Scores**
Woebot Group	18	8,8
Comparison Group	16	17,33

Note. The left column denotes the group categories of the sample. Pre PD Scores refer to the mean pre-study GHQ-12 scores and Post PD Post Scores refer to the mean post-study GHQ-12 scores of the treatment and comparison groups.

Table 12

Table Showing Assumption Of Homogeneity Of Variances Report Assessed Using Levene's Test For The Equality Of Variances

		Levene's Test for Equality of Variances	
		F	Sig.
Pre Post PD Difference	**Equal variances assumed**	.001	.971
	Equal variances not assumed		

Notes. Variances were homogeneous, as assessed by Levene's test for the equality of variances ($p = .971$).

Mean Difference Between Groups Report

The treatment group's mean psychological distress pre-test post-test change score was -5.07, 95% CI [-10.22 to 0.07], higher than the comparison group's mean psychological distress pre-test post-test change score. Furthermore, a difference of t(26) = -2.024, p = .053. was established between the psychological distress pretest and posttest change scores of the treatment group and comparison group, respectively.

Null and Alternative Hypotheses Report

Since the reported p-value (.53) < .05, the null hypothesis was retained.

Effect Size Report

The calculated effect size (*d*) was 0.76, which is considered a large effect size.

Independent-Samples T-Test Main Findings Report

There were 14 participants in the treatment group and 14 participants in the comparison group. An independent-samples t-test was performed to determine whether there were differences in the pretest and posttest changes in the PD scores between the participants in the treatment group and the participants in the comparison group. There were no outliers in the data, as assessed by an inspection of a boxplot. The PD change scores for each group were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$). Furthermore, the variances were homogeneous, as assessed by Levene's test for equality of variances ($p = .971$). The treatment group had a greater reduction in PD scores from the pretest to posttest ($M = -7.00$, $SD = 6.78$)

than the comparison group (*M* = -1.92, *SD* = 6.47), *M*= -5.07, 95% CI [-10.22 to 0.07], *t*(26) = -2.024, *p* = .053, *d* = .76.

Figure 6

Bar Graph Showing the Mean Difference Between the Treatment and Comparison Group's PD Change Scores

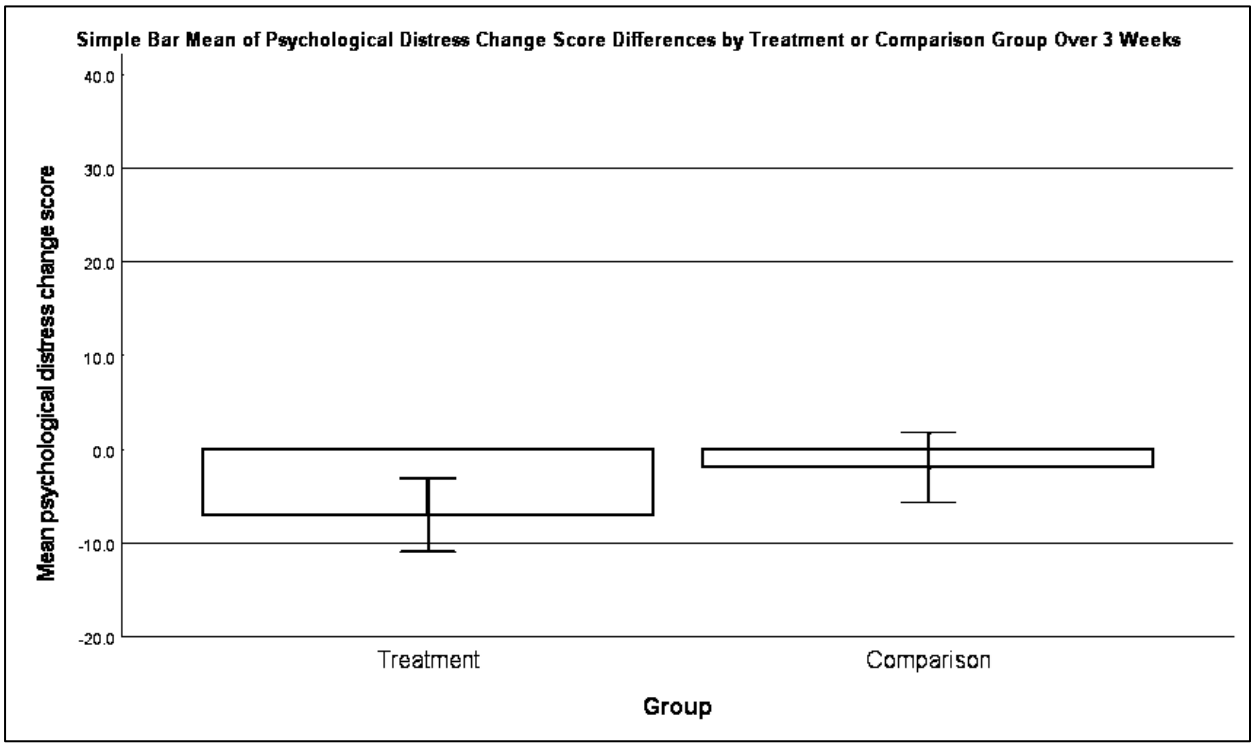


Figure 6. This figure shows that the treatment group had a greater mean reduction in PD than the comparison group after three weeks.

Conclusion

In this chapter, the statistical findings of the study are presented. First, the statistical hypotheses were presented together with their respective testing procedures. Second, the results of the Spearman correlation test and the independent samples t-test were presented in detail. In

the following chapter, the findings of the study are interpreted. The next chapter discusses the study's findings, implications, limitations and recommendations.

CHAPTER 6

Discussion, Implications, Limitations, Recommendations and Conclusion

Introduction

In this study, the effect of Woebot on the psychological distress (PD) levels of a sample of second- and third-year students at a South African university was examined. In Chapter One, the background and context of the research were outlined. The emergence of smartphone app technology for mental health, namely, mental health apps or MHapps, was introduced. A brief overview of PD among university students throughout the world, including in South Africa, was provided as well as mental health care barriers that prevent many students from seeking care. It was highlighted that very little research on MHapps exists within the South African context, specifically among populations that frequently use smartphones such as university students. The Woebot MHapp was introduced as the selected intervention for this study. Accordingly, the research endeavoured to determine whether the Woebot MHapp would effectively reduce the PD of a sample of South African university students over three weeks. The operationalised statements of the research questions, objectives and hypotheses posed for the study concluded the chapter.

A review of the literature constituted the Second Chapter. This included a general overview of the field of mental health apps, frameworks for guiding researchers and developers to select MHapps and MHapp features, and the incorporation of artificial intelligence (AI) technology into MHapps. The literature review had two primary purposes. The first was to provide a broader context and justification for conducting research on MHapps by highlighting the potential of MHapps for PD in university students, the role of MHapps for primary prevention and the importance of digital literacy and attitudes when selecting a target group for

MHapp research. Moreover, the overabundance of MHapps was highlighted. Consequently, multiple criteria and frameworks were presented to assist users, researchers and developers to select suitable MHapps. Specifically, features related to evidence-based content, privacy and user-friendliness emerged as key considerations when selecting a MHapp for research purposes. In the second section of Chapter Two, Bakker et al.'s (2016) evidence-based MHapp recommendations that informed the selection of the Woebot MHapp intervention for this study was provided. Each recommended MHapp feature was outlined briefly and its suitability for this study discussed. In the final section of this chapter, the research on AI integration into MHapps, specifically regarding chatbot app technology, was introduced. Within this context, the development and features of the Woebot app were presented.

In the Third Chapter, Cognitive Behavioural Therapy (CBT) as the theoretical basis for the study was presented by providing a brief history of CBT, an overview of its basic concepts and a discussion of how CBT and positivism were applied to the methodology and research approach of this study.

In Chapter Four, the research methodology that was employed to implement the Woebot MHapp intervention was explained. First, consideratons of the research design were discussed. A quasi-experimental pretest posttest non-equivalent group design was selected. Purposive sampling was used to recruit the sample by using social media and on-campus recruitment methods. The treatment and comparison group conditions were described and the psychometric properties of the General Health Questionnaire 12-Item version were presented. In addition, the demographic questionnaire was described and a CONSORT diagram was provided to summarise the data collection and data analysis process. Thereafter, an overview of the data analysis procedures was provided. A Spearman's rank-order correlation test was selected to analyse the

relationship between the number of Woebot interaction days and PD change scores over three weeks. In addition, an independent samples t-test was chosen to analyse whether there was a statistically significant difference between the treatment group that used the Woebot app and the comparison group that did not use any MHapps over three weeks. Subsequently, the statistical assumptions and descriptions of the two statistical tests were provided and the rationale for selecting the tests were outlined. Subsequently, the statistical assumptions and descriptions of the two statistical tests were provided, and the rationale for selecting the tests was outlined.

In the following section, the ethical considerations for the study were specified by outlining the principles of non-harm, voluntary participation, ethics approval, confidentiality, anonymity and privacy. Moreover, the privacy and security properties of the Woebot app and Qualtrics were detailed.

The statistical results of the study were presented in Chapter Five. First, the statistical hypotheses were stated for Spearman's correlation and the independent samples T-test. Subsequently, the findings of each test's statistical assumptions and main statistical findings were displayed.

The previous five chapters are integrated into a discussion of the findings in this chapter. First, the main findings are discussed in realtion to the supporting and contradictory findings from previous research on MHapps and the Woebot app. Second, the practical and theoretical implications of the results as well as the strengths and limitations of the study are considered. Finally, recommendations for future research are outlined.

Discussion of findings

The first research hypothesis was that there is a negative correlation between the number of average weekly Woebot interactions and PD reduction from pretest to posttest. The second

research hypothesis was that there is a statistically significant difference between the PD change scores of the treatment group that received the Woebot intervention and comparison group that did not receive any intervention after three weeks.

The first finding revealed that the number of days per week that the participants interacted with the Woebot app was negatively correlated with psychological distress. In other words, the participants that interacted with the app more frequently experienced a greater decrease in psychological distress over three weeks. Although the exact number of daily interactions and nature of those interactions were not recorded, the average number of weekly usage days served as a valuable metric for consistency of use and adherence to the Woebot app. Previous studies on the Woebot app and similar MHapps established weekly average usage counts to determine the adherence rates and correlations between app usage counts and mental health outcome measures (Fitzpatrick et al., 2017; Ly et al., 2017; Morris et al., 2015).

The second finding revealed that the treatment group had a greater reduction in PD scores from the pretest to posttest than the comparison group. However, this difference was statistically non-significant ($p > 0.5$). This finding of a statistically non-significant difference has also been reported in previous MHapp studies. For instance, Ly et al. (2017) found that the intervention group did not differ significantly from the comparison group on measures of psychological well-being after the two-week intervention on a chatbot MHapp. Moreover, in a meta-analysis, Weisel et al. (2019) revealed that multiple MHapp interventions for various mental health conditions failed to produce statistically significant results. However, it is essential to emphasise that many factors may have influenced the finding of non-statistical differences between the groups in this study. Norton and Strube (2001) noted that four primary factors may influence the power of a statistical test: level, the difference between group means, variability among subjects and sample

size. In this study, the sample size may have had the most substantial influence on the statistical results. Although the sample included 14 participants in each group in the final analysis, Norton and Strube (2001) emphasised that the least complicated way to increase the statistical power of a test is to have a large sample size. Therefore, the relatively small sample size was singled out as one of the potential explanations for the statistically non-significant difference between the groups. The variability among the participants may have also affected the statistical significance. While there were 19 female participants, there were nine male participants. In addition, the participants were enrolled in a variety of faculties: there 14 in the humanities faculty, seven in engineering, four in natural sciences, two in economic and management sciences and one in education. Furthermore, the racial distribution of the sample was relatively unbalanced and under-representative, with 21 White/Caucasian and seven African/Black students. In addition, while 11 were second-year students, 17 were third-year students.

Despite the statistically non-significant difference ($p > 0,5$) between the PD change scores after three weeks, the Woebot treatment group still reported a higher average decrease of five points on the GHQ-12 measure than the comparison group. Considering that the average number of weekly Woebot interaction days for the treatment group was four days per week, it is plausible that the Woebot app may have influenced the more significant decrease in PD observed in the treatment group. This concurs with previous research on the Woebot app that revealed more significant improvements in mental wellbeing in groups that used the Woebot app for most days of the week compared to control groups that did not use the app at all (Demirci, 2018; Fitzpatrick et al., 2017).

The high attrition rates observed in the study were another significant finding. The attrition rate was 54.8% from the pretest to the posttest, which is much higher than in previous

MHapp studies on chatbot-based apps. For instance, Ly et al. (2017) revealed a high adherence rate of 78.6% for a MHapp intervention and Fitzpatrick et al. (2017) reported an attrition rate of 17% for Woebat. However, several MHapp intervention studies have reported similarly high attrition rates of between 50% and 60% (Bakker et al., 2016; Hochheimer et al., 2016; Marshall et al., 2020a; McCloud et al., 2020). Linardon and Fuller-Tyszkiewicz (2020) noted that a large percentage of participants in MHapp intervention studies drop out and do not use the app as instructed.

Furthermore, Linardon and Fuller-Tyszkiewicz (2020) emphasised that the development of retention strategies may be essential for decreasing attrition and allowing a more effective long-term use of mental health apps. Many factors may have contributed to the high attrition rates in this study. First, the study was conducted online. Research has demonstrated that internet-based studies report higher attrition rates and incompleted questionnaires (Ormrod & Leedy, 2015; Price et al., 2012). Second, no financial incentives were offered to the participants in this study. Although this eliminated the risk of enrollment bias and exploitation, previous digital intervention studies have provided some form of incentive to participants, which may have reduced attrition rates (Resnik, 2015). Third, there was a relatively long intervention period of three weeks in comparison to previous chatbot studies that lasted for approximately two weeks (e.g., Fitzpatrick et al., 2017; Ly et al., 2017; Prochaska et al., 2021). Fourth, higher attrition rates have been observed in digital health intervention studies on young adults (Achilles et al., 2020). As this study's sample comprised young adults between 18 and 24 years, the high attrition rates were consistent with studies with similar sample age groups. Achilles et al. (2020) provided several intervention suggestions to improve young adults' adherence to digital health interventions, including (a) outlining the expectations of programme usage at the outset, (b)

embedding engagement checks and providing feedback on responses, (c) including the user in the design procedure to support an understanding of user needs and (d) implementing persuasive design features. More detailed expectations of intervention usage from the outset and more regular engagement checks and feedback with participants could have been included in this study.

Achilles et al. (2020) also recommended research design features to improve the digital intervention adherence of young adults. These include (a) operationalising adherence from the outset, (b) offering justifications for the adherence guidelines, (c) measuring users' expectations and preferences and (d) implementing study designs other than randomised controlled trials to study adherence factors. Adherence could have been operationalised from the outset in this study by justifiying adherence to a set level of engagement with the intervention.

Another notable finding was the strong effect size of d = .76 between the treatment and comparison group and the psychological distress change scores over three weeks. This is relatively high in comparison to previous MHapp studies. Ly et al. (2017) found a between-group effect size of d = 0.14 between their MHapp intervention group and control group. Moreover, in a chatbot MHapp intervention, Inkster et al. (2018) observed a similar effect size of d=0.63 between the groups. A meta-analysis conducted by Weiss et al. (2016) demonstrated that average between-group effect sizes of d = 0.44 were typical for behavioural interventions for mental wellbeing. Similarly, a systematic review and meta-analysis of smartphone app studies for mental health revealed effect sizes of 0.33 for depression outcomes and 0.43 for anxiety outcomes (Weisel et al., 2019). The practical and theoretical implications of these findings are discussed in the following section.

Implications

This study has several implications for MHapp research and development. First, this is a pioneering study for chatbot MHapps in the South African context. Chatbot technology is a growing area of research and various researchers have found that using chatbots may be especially beneficial for overcoming mental health barriers (Darcy et al., 2021; Skjuve & Brandtz, 2018). Furthermore, the Woebot app among South African university student samples was also a first. Skjuve and Brandtz (2018) noted that chatbot apps have shown considerable potential for use among young adults and adolescents due to the confidentiality and anonymity they provide, thus allowing users to share information with the bot without concerns about judgements and stigma. Although the findings did not reveal that the Woebat app had a substantial effect on the PD levels of the sample, its capacity to provide an open and interactive space for users to share their mental health concerns has been documented in previous studies. Darcy et al. (2021) found that the Woebot app was able to establish a therapeutic bond with participants comparable to human-delivered therapy interventions. Darcy et al. (2021) noted further that if digital interventions can mimic factors traditionally believed to be uniquely human, such as establishing rapport in therapy, they may have much greater potential for enhancing mental well-being.

The second implication involves the importance of employing MHapp frameworks and criteria for selecting MHapp features as the basis for research and/or app development. As noted previously, Bakker et al.'s (2016) evidence-based framework for MHapps was employed. Furthermore, various frameworks such as the MIND framework utilised by Lagan et al. (2021), the criteria of usability, evidence basis and technical features outlined by Bergin and Davies (2019) and the MHapp challenges and suggestions proposed by Chandrashekar (2018), which

were detailed in Chapter Two, can serve as the basis for selecting or revising MHapp features for research and/or MHapp development for particular target groups.

Strengths of the Study

This study has several strengths. The relatively longitudinal intervention length of three weeks may be regarded as the study's first strength. Several MHapp studies on chatbot app interventions have lasted only two weeks (e.g., Fitzpatrick et al., 2017; Ly et al., 2017; Prochaska et al., 2021). The addition of a third week provided a more comprehensive indication of the participants' adherence and habitual Woebot usage patterns over a more extended period.

Second, the Woebot app intervention was conducted in a real-world setting and not in a lab-controlled environment. Therefore, the implementation of a MHapp intervention in such a way so as to simulate real-world application is crucial to establish generalisability to broader populations of users (Hwang et al., 2021; Nahum-Shani et al., 2015, 2018).

Third, the study's between-subjects research design may be regarded as a strength. Bhandari (2021) noted that the first benefit of using a between-subjects design includes a reduction of carryover effects, which are the learning or practice effects when one group is exposed to multiple treatment conditions. Bhandari added that the between-subjects study design allows for a shorter study duration because all participants are assigned to only one condition.

Fourth, the inclusion of a comparison group is a strength. According to Lewis-Beck et al. (2004), the inclusion of a comparison group that does not receive the treatment is beneficial for testing the effects of an intervention. Furthermore, by implementing a comparison group in a between-subjects design, the risk of intervention fatigue effects, practice effects and carryover effects associated with within-subject designs are reduced (Lewis-Beck et al., 2004).

Fifth, employing a pretest posttest non-equivalent comparison group design may be

regarded as a strength. Ormrod and Leedy (2015) posited that the inclusion of a pretest establishes that the treatment and comparison groups are at least similar to the dependent variables measured at baseline. Furthermore, a pretest posttest non-equivalent design allows the magnitude and difference between the pre- and posttest dependent variable measure of treatment and comparison group to be compared (Ormrod & Leedy, 2015). Assuming statistically significant differences in the dependent variable are observed between the two groups after the intervention, one may conclude that the posttest differences are due to the intervention (Ormrod & Leedy, 2015).

A sixth strength encompassed how potential extraneous variables were controlled to account for the effects of maturation, historical bias, regression artefact, instrumentation bias, the Hawthorne effect and attrition. Maturation refers to changes in a participant's internal conditions that occur as a function of time (Christensen et al., 2014). The changes involve biological and psychological processes such as age, learning, fatigue, boredom and hunger that are not related to specific external events but reside within the individual. Christensen et al. (2014) stated that including a comparison group that does not receive the intervention, as in this study, is a suitable way to mitigate the influence of maturation.

Historical bias is any event that occurs during the intervention that may have an influence on the outcome of the intervention other than the treatment itself (Christensen et al., 2014). Christensen et al. (2014) suggested that including two independent groups, comprising one treatment group and one comparison group, helps to reduce the impact of historical bias. Accordingly, this research study included two independent groups, namely, a treatment and comparison group to account for the influence of historical bias.

Regression artefact occurs when the selection of participants is based on extremely high

or low scores. The natural tendency that results from this is that while the scores of participants with very high initial scores tend to decrease posttest, those with very low initial scores tend to show an increase in posttest scores even without participating in any treatment (Christensen et al., 2014). The problem of regression artefacts was addressed by including two independent groups, namely, a treatment and comparison group.

Instrumentation bias refers to inconsistencies in the measurement of the dependent variable throughout the study (Christensen et al., 2014). This form of bias can be addressed by applying the same measurement scale in the same manner from pretest to posttest (Ormrod & Leedy, 2015). Accordingly, to reduce the chances of instrumentation bias, the GHQ-12 measure was applied in the same Qualtrics format to participants at the pretest and posttest. The Hawthorne effect refers to a phenomenon in which participants in a research study change their behaviour only because they are aware of being part of an experimental intervention (Ormrod & Leedy, 2015). Christensen et al. (2014) noted that the Hawthorne effect can be mitigated by adding a comparison group, as in this study. Attrition refers to losing participants because they did not show up or dropped out of the study (Christensen et al., 2014). Ormrod and Leedy (2015) stated that researchers who include human participants in their study must contend with participants who do not show up for the experiment at the designated time and place or do not participate in the assignment conditions required by the study. The effect of attrition on internal validity is particularly problematic for two-group between-subjects designs but not as much for within-group designs (Christensen et al., 2014). Furthermore, attrition affects the internal validity of a study, not only because of the loss of participants but because this might produce differences in the groups that cannot be attributed to the experimental treatment. This is known as differential attrition (Christensen et al., 2014). Attempts were made to reduce attrition in this

study by communicating with participants through email and minimising the length of the weekly surveys.

Limitations of the Study

This study has several limitations. The first involved employing a quasi-experimental design. The lack of randomisation in quasi-experimental designs limits the extent to which causal relationships between independent and dependent variables can be determined (Schweizer et al., 2016). Furthermore, quasi-experimental designs may introduce various forms of bias into the research. For example, selection bias in the treatment and comparison groups may affect the results in non-equivalent group designs (Shadish et al., 2002). The most common types of selection bias in quasi-experimental designs include maturation bias, regression artefact, historical bias, instrumentation bias and the Hawthorne effect. Reporting bias is also common in quasi-experimental designs in which only studies with positive findings and not studies with negative or statistically non-significant results are published (Shadish et al., 2002).

Second, there was only one comparison group that did not receive any treatment. According to Yoon et al. (2011), including more than one comparison group helps improve internal validity when randomisation is unfeasible and provides a way to test successive hypotheses. For instance, previous MHapp studies have included more than one comparison group, with one group not receiving an app and the other group a different app to use during the intervention phase (Bakker et al., 2018; Flett et al., 2020; Lee & Jung, 2018; Ly et al., 2017).

Third, the sample of the study was small, which may have had an impact on the results of the data analyses. Larger sample sizes are beneficial for interpreting statistically significant results, estimating more accurate estimates of treatment effects and establishing greater generalisability from the sample (Biau et al., 2008). Furthermore, there were also limitations

regarding the sampling of the population. The sample included individuals aged between 18 and 24 years who were either second- or third-year students at the researcher's university. Furthermore, the sample only included students with access to smartphones and the Internet. In addition, most participants were sampled based on availability. However, due to COVID-19 pandemic restrictions, many students were not attending lectures on the campus. Therefore, the generalisability of the results of the specific sample may not have been entirely in line with the actual population.

Recommendations for Future Research

Based on the limitations of the study, various recommendations for future research emerged. First, it would be beneficial if a mixed-methods design was employed when studying the effect of an MHapp in future research as this would provide a more comprehensive indication of MHapp users' subjective experiences and the mechanisms that led to changes in their mental health outcomes. For instance, multiple previous studies on the effects of chatbot MHapps have combined quantitative and qualitative data to significant effect (e.g., Fitzpatrick et al., 2017; Inkster et al., 2018; Ly et al., 2017). Although this approach is much more time-consuming and resource-demanding than a quantitative or qualitative design, mixed methods designs have many benefits. First, mixed methods designs afford researchers the opportunity to address a research problem completely by collecting, analysing and interpreting both quantitive and qualitative data (Ormrod & Leedy, 2015). Second, gathering quantitative information may compensate for the limitations of qualitative data and vice versa (Ormrod & Leedy, 2015). The quantitive results of this study could have been reinforced with qualitative information related to participants' experiences, attitudes and perceptions about using the Woebot app. In addition, qualitative information is beneficial for generating a causal relationship hypothesis that can then be tested by

employing an experimental quantitative design (Ormrod & Leedy, 2015). A qualitative approach, such as unstructured interviews with university students, could have first been employed in this study and subsequently a quantitative experimental design based on the hypothesis formulated using the qualitative data. Finally, mixed methods may provide more scientifically sound conclusions if the quantitative and qualitative data align through triangulation (Ormrod & Leedy, 2015). Thus, if the results of the statistically non-significant effect of the Woebot app on the PD of university students concurred with qualitative, subjective reports of the Woebot app, the study's conclusions may have been strengthened.

The second recommendation involves conducting a futher study on the Woebot app for psychological distress among the university population by employing more rigorous experimental designs such as randomised controlled trials. Hariton and Locascio (2018) asserted that randomised controlled trials (RTCs) allow the causal relationship between an intervention and outcome measure to be examined rigorously. Previous MHapp studies, such as an RCT conducted by Bakker et al. (2018) to compare three different MHapps, had significant findings based on the rigour of the design. Furthermore, alternative research designs exploring the app's user experience and engagement features could also be beneficial for future research on the Woebot app. Gaffney et al. (2019) suggested streamlining interventions, conducting comparison studies with other interventions and studying the mechanisms that lead to increased adherence, reach and effectiveness of MHapps as further avenues for MHapp research.

Moreover, it is recommended that future research employ probability sampling methods to recruit a more representative sample. Because this study utilised non-probability purposive sampling, the sample distribution may have been less representative of the characteristics of the actual population than if a probability sampling method had been used (Ormrod & Leedy, 2015).

Fricker (2017) highlighted that for online-based studies, probability sampling strategies such as using a list-based sampling frame, a non-list-based random sampling method and pre-recruited panels could be used to recruit participants.

Third, a post-study follow-up survey weeks or months after the intervention to determine the sustained effect and adherence to Woebot is recommended. Harrer et al. (2018) conducted a three-month follow-up survey to monitor the sustained effect of an internet-based, app-supported stress management intervention for college students. Birney et al. (2016) included a ten-week follow-up assessment and found the Mood Hacker mobile web app had a sustained influence on depression. Similarly, Proudfoot et al. (2013) found that anxiety, stress and depression symptoms assessed post-intervention and at a three-month follow-up remained stable. It is recommended that future studies include follow-up measures such as those noted to strengthen the validity of findings for long-term effects of mental health apps.

Fourth, it is recommended that future studies on the Woebot app or MHapp interventions include a larger sample. As noted previously, larger sample sizes improve the internal validity of the study by estimating the effect of the treatment as well as the external validity of the study more precisely by increasing the representativeness of the sample and generalisability of the results (Biau et al., 2008). Moreover, future studies on the Woebot app for university student populations in South Africa could include samples from multiple universities to determine whether the findings can be generalised to all South African universities. In addition, post-graduate student samples could be compared to undergraduate samples to determine whether the graduate level has an effect on app adherence and efficacy for psychological distress.

Finally, it is recommended that more objective ways to record and monitor user interactions be included. In this study, data were collected by employing self-report measures to

capture the number of weekly interactions with the Woebot app. If possible, it is advised that future researchers could gain access to API Webhooks from the developers of the Woebot app. This would allow data such as log-in counts or daily usage to be recorded more accurately.

Conclusion

The effect of the Woebot MHapp on the PD levels of South African university students was examined in this study. The findings showed a non-statistically significant negative correlation between the participants' average number of weekly Woebot interaction days and psychological distress. The findings demonstrated a large effect size of the Woebot intervention, even though the results were statistically non-significant.The findings also showed that the Woebot treatment group had a higher reduction in psychological distress than the comparison group. Thus, the results showed that the Woebot intervention had a reduction effect on the psychological distress of university students.

Further research is required to evaluate the effect of the Woebot app on psychological distress among South African university student samples by employing more rigorous research designs such as randomised control trials, extended intervention periods and larger sample sizes. In addition, given the prevalence of psychological distress among university students, smartphone mental health apps similar to Woebot could be employed as a potential option in an endeavour to provide alternative mental health support to university students.

REFERENCES

Achilles, M. R., Anderson, M., Li, S. H., Subotic-Kerry, M., Parker, B., & O'Dea, B. (2020). Adherence to e-mental health among youth: Considerations for intervention development and research design. *DIGITAL HEALTH*, *6*, 205520762092606. https://doi.org/10.1177/2055207620926064

Allen, J. C. (2011). Sample size calculation for two independent groups: A useful rule of thumb. *Proceedings of Singapore Healthcare*, *20*(2), 138–140. https://doi.org/10.1177/201010581102000213

Apuke, O. D. (2017). Quantitative research methods: A synopsis approach. *Kuwait Chapter of Arabian Journal of Business and Management Review*, *6*(11), 40–47. https://doi.org/10.12816/0040336

Arvidsdotter, T., Marklund, B., Kylén, S., Taft, C., & Ekman, I. (2016). Understanding persons with psychological distress in primary health care. *Scandinavian Journal of Caring Sciences*, *30*(4), 687–694. https://doi.org/10.1111/scs.12289

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Murray, E., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Stein, D. J., Vilagut, G., Zaslavsky, A. M., Kessler, R. C., & WHO WMH-ICS Collaborators. (2018). WHO World Mental Health surveys international college student project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology*, *127*(7), 623–638. https://doi.org/10.1037/abn0000362

Aziz, M., Erbad, A., Almourad, M. B., Altuwairiqi, M., McAlaney, J., & Ali, R. (2022). Did usage of mental health apps change during COVID-19? A comparative study based on an

objective recording of usage data and demographics. *Life*, *12*(8), 1266. https://doi.org/10.3390/life12081266

Bannink, F. (2017). Positive CBT in practice. In C. Proctor (Ed.), *Positive Psychology Interventions in Practice* (pp. 15–28). Springer International Publishing. https://doi.org/10.1007/978-3-319-51787-2_2

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2016). 2. Mental health smartphone apps: review and evidence-based recommendations for future developments. *JMIR Mental Health*, *3*(1), e7. https://doi.org/10.2196/mental.4984

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2018). 3. A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behaviour Research and Therapy*, *109*, 75–83. https://doi.org/10.1016/j.brat.2018.08.003

Bantjes, J. R., Kagee, A., McGowan, T., & Steel, H. (2016). Symptoms of posttraumatic stress, depression, and anxiety as predictors of suicidal ideation among South African university students. *Journal of American College Health*, *64*(6), 429–437. https://doi.org/10.1080/07448481.2016.1178120

Bantjes, J., Saal, W., Lochner, C., Roos, J., Auerbach, R. P., Mortier, P., Bruffaerts, R., Kessler, R. C., & Stein, D. J. (2020). Inequality and mental healthcare utilisation among first-year university students in South Africa. *International Journal of Mental Health Systems*, *14*(1), 5. https://doi.org/10.1186/s13033-020-0339-y

Basu, K., Sinha, R., Ong, A., & Basu, T. (2020). Artificial Intelligence: How is it changing medical sciences and its future? *Indian Journal of Dermatology*, *65*(5), 365–370. PubMed. https://doi.org/10.4103/ijd.IJD_421_20

Baumel, A., Torous, J., Edan, S., & Kane, J. M. (2020). There is a non-evidence-based app for that: A systematic review and mixed methods analysis of depression- and anxiety-related apps that incorporate unrecognized techniques. *Journal of Affective Disorders, 273*, 410–421. https://doi.org/10.1016/j.jad.2020.05.011

Beck, A. T. (1976). Cognitive therapy and the emotional disorders. *International Universities Press*.

Beck, A. T. (1997). The past and future of cognitive therapy. *The Journal of Psychotherapy Practice and Research, 6*(4), 276–284. PubMed.

Beck, J. S. (2011). *Cognitive behavior therapy: Basics and beyond (2nd ed.). Guilford Press.*

Beck, A. T., & Haigh, E. A. P. (2014). Advances in cognitive theory and therapy: The generic cognitive model. *Annual Review of Clinical Psychology, 10*(1), 1–24. https://doi.org/10.1146/annurev-clinpsy-032813-153734

Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., … & Mandell, D. S. (2015). Free, brief, and validated: standardized instruments for low-resource mental health settings. Cognitive and Behavioral Practice, 22(1), 5-19.

Benjamin, C. L., Puleo, C. M., Settipani, C. A., Brodman, D. M., Edmunds, J. M., Cummings, C. M., & Kendall, P. C. (2011). History of Cognitive-Behavioral Therapy in youth. *Child and Adolescent Psychiatric Clinics of North America*, *20*(2), 179–189. https://doi.org/10.1016/j.chc.2011.01.011

Bergin, A., & Davies, E. B. (2019). Technology Matters: Mental health apps - separating the wheat from the chaff. *Child and adolescent mental health*, *25*(1), 51–53. https://doi.org/10.1111/camh.12363

Berry, K., Salter, A., Morris, R., James, S., & Bucci, S. (2018). Assessing therapeutic alliance in the context of mhealth interventions for mental health problems: Development of the Mobile Agnew Relationship Measure (mARM) Questionnaire. *Journal of Medical Internet Research*, *20*(4), e90. https://doi.org/10.2196/jmir.8252

Bhandari, P. (2021, March 12). *What is a between-subjects design?* Scribbr. https://www.scribbr.com/methodology/between-subjects-design/

Biau, D. J., Kernéis, S., & Porcher, R. (2008). Statistics in brief: the importance of sample size in the planning and interpretation of medical research. *Clinical orthopaedics and related research*, *466*, 2282-2288. https://doi.org/10.1007/s11999-008-0346-9

Bigsby, E. (2018). *Sampling, Nonprobability*. SAGE Publications, Inc. https://doi.org/10.4135/9781483381411

Birney, A. J., Gunn, R., Russell, J. K., & Ary, D. V. (2016). MoodHacker mobile web app with email for adults to self-manage mild-to-moderate depression: Randomized controlled trial. *Journal of Medical Internet Research* (*JMIR) MHealth and UHealth*, *4*(1), e8. https://doi.org/10.2196/mhealth.4231

Bogdan, R. C., & Biklen, S. K. (2003). *Bogdan, R. C., & Biklen, S. K. (2003). Qualitative research of education: An introductive to theories and methods (4th ed.). Boston: Allyn and Bacon.*

Burns, D. D. (1980). *Feeling good: The new mood therapy*.

Centofanti, S., Lushington, K., Wicking, A., Wicking, P., Fuller, A., Janz, P., & Dorrian, J. (2019). Establishing norms for mental well-being in young people (7–19 years) using the General Health Questionnaire-12. *Australian Journal of Psychology*, *71*(2), 117–126. https://doi.org/10.1111/ajpy.12227

Chandrashekar, P. (2018). 21. 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

Chernomas, W. M., & Shapiro, C. (2013). Stress, depression, and anxiety among undergraduate nursing students. *International Journal of Nursing Education Scholarship*, *10*(1), 255–266. https://doi.org/10.1515/ijnes-2012-0032

Chew, A. M. K., Ong, R., Lei, H.-H., Rajendram, M., K V, G., Verma, S. K., Fung, D. S. S., Leong, J. J., & Gunasekeran, D. V. (2020). Digital health solutions for mental health disorders during COVID-19. *Frontiers in Psychiatry*, *11*, 582007. https://doi.org/10.3389/fpsyt.2020.582007

Chow, S. L. (1992). Positivism and cognitive psychology: A second look. *Positivism in psychology: Historical and contemporary problems*, 119-144. https://doi.org/10.1007/978-1-4612-4402-8_10https://doi.org/10.1007/978-1-4612-4402-8_10

Christensen, L. B., Johnson, B., & Turner, L. A. (2014). *Research methods, design, and analysis* (Twelfth Edition). Pearson.

Clay. (2021). *Mental health apps are gaining traction*. https://www.apa.org/monitor/2021/01/trends-mental-health-apps

Cohen, J. S., Edmunds, J. M., Brodman, D. M., Benjamin, C. L., & Kendall, P. C. (2013). Using self-monitoring: implementation of collaborative empiricism in cognitive-behavioral therapy. *Cognitive and Behavioral Practice*, *20*(4), 419–428. https://doi.org/10.1016/j.cbpra.2012.06.002

Connolly, S. L., Hogan, T. P., Shimada, S. L., & Miller, C. J. (2021). Leveraging implementation science to understand factors influencing sustained use of mental health apps: A narrative review. *Journal of Technology in Behavioral Science*, *6*(2), 184–196. https://doi.org/10.1007/s41347-020-00165-4

Conover, W.J. (1999) Practical Nonparametric Statistical. 3rd Edition, John Wiley & Sons Inc., New York, 428-433.

Cuéllar-Flores, I., Sánchez-López, M. P., Limiñana-Gras, R. M., & Colodro-Conde, L. (2014). The GHQ-12 for the assessment of psychological distress of family caregivers. *Behavioral Medicine*, *40*(2), 65–70. https://doi.org/10.1080/08964289.2013.847815

Cully, J. A., & Teten, A. L. (2008). *A therapist's guide to brief cognitive behavioral therapy*. 111.

Darcy, A., Daniels, J., Salinger, D., Wicks, P., & Robinson, A. (2021). 5. Evidence of human-level bonds established with a digital conversational agent: Cross-sectional, retrospective observational study. *JMIR Formative Research*, *5*(5), e27868. https://doi.org/10.2196/27868

David, D., Cristea, I., & Hofmann, S. G. (2018). Why cognitive behavioral therapy is the current gold standard of psychotherapy. *Frontiers in Psychiatry*, *9*, 4. https://doi.org/10.3389/fpsyt.2018.00004

DeCoster, J. (1998). *Overview of factor analysis*. Retrieved from http://www.stat-help.com/notes.html

DiGiuseppe, R., Venezia, R., & Gotterbarn, R. (2018). What is cognitive behavior therapy? In A. Vernon & K. A. Doyle (Eds.), *Cognitive behavior therapies: A guidebook for*

practitioners (pp. 1–35). American Counseling Association. https://doi.org/10.1002/9781119375395.ch1

de Kock, F. S., Görgens-Ekermans, G., & Dhladhla, T. J. (2014). A confirmatory factor analysis of the General Health Questionnaire–28 in a black South African sample. *Journal of Health Psychology*, *19*(10), 1222–1231. https://doi.org/10.1177/1359105313488972

Delloite. (2021). *Global governance toolkit for digital mental health: Building trust in disruptive technology*.

Demirci, H. M. (2018). *User experience over time with conversational agents: case study of woebot on supporting subjective well-being* [M.S. - Master of Science]. Middle East Technical University.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining 'gamification'. *Proceedings of the 15th international academic mindtrek conference on envisioning future media environments - MindTrek '11*, 9. https://doi.org/10.1145/2181037.2181040

Dobson, D., & Dobson, K. S. (2009). *Evidenced-based practice of cognitive-behavioral therapy. New York: Guilford.*

Dobson, K. S. (2008). Cognitive therapy for depression. In *Adapting cognitive therapy for depression: Managing complexity and comorbidity.* (pp. 3–35). Guilford Press.

Dozois, D. J. A. (2001). *Historical and philosophical bases of the cognitive-behavioral therapies*. 30.

Dozois, D. J. A., & Beck, A. T. (2008). Cognitive schemas, beliefs and assumptions. *risk factors in depression*, 24(1).

Drapeau, A., Marchand, A., & Beaulieu-Prevost, D. (2012). 1. (C) Epidemiology of psychological distress. In L. LAbate (Ed.), *Mental Illnesses—Understanding, Prediction and Control*. InTech. https://doi.org/10.5772/30872

Endsley, P., Weobong, B., & Nadkarni, A. (2017). The psychometric properties of GHQ for detecting common mental disorder among community dwelling men in Goa, India. *Asian Journal of Psychiatry, 28*, 106–110. https://doi.org/10.1016/j.ajp.2017.03.023

Etikan, I. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics, 5*(1), 1. https://doi.org/10.11648/j.ajtas.20160501.11

Eysenbach, G. (2005). The Law of Attrition. *Journal of Medical Internet Research, 7*(1), e11. https://doi.org/10.2196/jmir.7.1.e11

Farrer, L. M., Gulliver, A., Katruss, N., Fassnacht, D. B., Kyrios, M., & Batterham, P. J. (2019). A novel multi-component online intervention to improve the mental health of university students: Randomised controlled trial of the Uni Virtual Clinic. *Internet Interventions, 18*, 100276. https://doi.org/10.1016/j.invent.2019.100276

Fenn, K., & Byrne, M. (2013). The key principles of cognitive behavioural therapy. *InnovAiT: Education and Inspiration for General Practice, 6*(9), 579–585. https://doi.org/10.1177/1755738012471029

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). 1.(C) Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), e19. https://doi.org/10.2196/mental.7785

Flett, J. A. M., Conner, T. S., Riordan, B. C., Patterson, T., & Hayne, H. (2020). App-based mindfulness meditation for psychological distress and adjustment to college in incoming university students: A pragmatic, randomised, waitlist-controlled trial. *Psychology & Health*, *35*(9), 1049–1074. https://doi.org/10.1080/08870446.2019.1711089

Fricker, R. D. (2017). Sampling methods for online surveys. In N. G. Fielding, R. M. Lee, & G. Blank, *The SAGE Handbook of Online Research Methods* (pp. 162–183). SAGE Publications Ltd. https://doi.org/10.4135/9781473957992.n10

Gaffney, H., Mansell, W., & Tai, S. (2019). Conversational agents in the treatment of mental health problems: mixed-method systematic review. *JMIR Mental Health*, *6*(10), e14166. https://doi.org/10.2196/14166

Gelaye, B., Tadesse, M. G., Lohsoonthorn, V., Lertmeharit, S., Pensuksan, W. C., Sanchez, S. E., Lemma, S., Berhane, Y., Vélez, J. C., Barbosa, C., Anderade, A., & Williams, M. A. (2015). Psychometric properties and factor structure of the General Health Questionnaire as a screening tool for anxiety and depressive symptoms in a multi-national study of young adults. *Journal of Affective Disorders*, *187*, 197–202. https://doi.org/10.1016/j.jad.2015.08.045

Goldberg, D. P., Gater, R., Sartorius, N., Ustun, T. B., Piccinelli, M., Gureje, O., & Rutter, C. (1997). The validity of two versions of the GHQ in the WHO study of mental illness in general health care. *Psychological Medicine*, *27*(1), 191–197. https://doi.org/10.1017/S0033291796004242

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: an Overview. *Current psychiatry reports*, *21*(11), 116. https://doi.org/10.1007/s11920-019-1094-0

Gravetter, F. J., & Forzano, L.-A. B. (2012). *Research methods for the behavioral sciences* (4th ed). Wadsworth.

Hafen, M., Reisbig, A. M. J., White, M. B., & Rush, B. R. (2006). Predictors of depression and anxiety in first-year veterinary students: A preliminary report. *Journal of Veterinary Medical Education*, *33*(3), 432–440. https://doi.org/10.3138/jvme.33.3.432

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. *BJOG: An International Journal of Obstetrics & Gynaecology*, *125*(13), 1716–1716. https://doi.org/10.1111/1471-0528.15199

Harrer, M., Adam, S. H., Fleischmann, R. J., Baumeister, H., Auerbach, R., Bruffaerts, R., Cuijpers, P., Kessler, R. C., Berking, M., Lehr, D., & Ebert, D. D. (2018). 7. Effectiveness of an internet- and app-based intervention for college students with elevated stress: randomized controlled trial. *Journal of Medical Internet Research*, *20*(4), e136. https://doi.org/10.2196/jmir.9293

Herbert, J. D., Gaudiano, B. A., & Forman, E. M. (2013). The importance of theory in cognitive behavior therapy: A perspective of contextual behavioral science. *Behavior Therapy*, *44*(4), 580–591. https://doi.org/10.1016/j.beth.2013.03.001

Hill, C. R., & Hughes, J. N. (2007). An examination of the convergent and discriminant validity of the Strengths and Difficulties Questionnaire. *School Psychology Quarterly*, *22*(3), 380–406. https://doi.org/10.1037/1045-3830.22.3.380

Hochheimer, C. J., Sabo, R. T., Krist, A. H., Day, T., Cyrus, J., & Woolf, S. H. (2016). Methods for evaluating respondent attrition in web-based surveys. *Journal of Medical Internet Research*, *18*(11), e301. https://doi.org/10.2196/jmir.6342

Hofmann, S. G., Asnaani, A., Vonk, I. J. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, *36*(5), 427–440. https://doi.org/10.1007/s10608-012-9476-1

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

Hwang, W. J., Ha, J. S., & Kim, M. J. (2021). 9. Research trends on mobile mental health application for general population: a scoping review. *International Journal of Environmental Research and Public Health*, *18*(5), 2459. https://doi.org/10.3390/ijerph18052459

Inkster, B., Sarda, S., & Subramanian, V. (2018). An empathy-driven, conversational artificial intelligence agent (wysa) for digital mental well-being: real-world data evaluation mixed-methods study. *JMIR MHealth and UHealth*, *6*(11), e12106. https://doi.org/10.2196/12106

Jhangiani, R. S., Chiang, I.-C. A., Cuttler, C., & Leighton, D. C. (2019). *Research Methods in Psychology*. 446.

Kajitani, K., Higashijima, I., Kaneko, K., Matsushita, T., Fukumori, H., & Kim, D. (2020). 11. Short-term effect of a smartphone application on the mental health of university students: A pilot study using a user-centered design self-monitoring application for mental health. *PLOS ONE*, *15*(9), e0239592. https://doi.org/10.1371/journal.pone.0239592

Kaminer, D., & Shabalala, N. (2019). Developing a student mental health policy for a South African university: Consultation, contestation and compromise. *South African Journal of Higher Education*, *33*(5). https://doi.org/10.20853/33-5-3597

Kampel, L., & Orman, J. (2017). E-mental health for psychological distress in university students: A narrative synthesis on current evidence and practice. *International Journal of Innovation, 3*(3), 13.

Kim, Y. J., Cho, M. J., Park, S., Hong, J. P., Sohn, J. H., Bae, J. N., Jeon, H. J., Chang, S. M., Lee, H. W., & Park, J.-I. (2013). The 12-Item General Health Questionnaire as an effective mental health screening tool for general korean adult population. *Psychiatry Investigation, 10*(4), 352. https://doi.org/10.4306/pi.2013.10.4.352

Knapstad, M., Sivertsen, B., Knudsen, A. K., Smith, O. R. F., Aarø, L. E., Lønning, K. J., & Skogen, J. C. (2021). 6. Trends in self-reported psychological distress among college and university students from 2010 to 2018. *Psychological Medicine, 51*(3), 470–478. https://doi.org/10.1017/S0033291719003350

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., Singh, I., & Group, N. Y. P. A. (2019). Can your phone be your therapist? young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support. Biomedical Informatics Insights, 11, 1178222619829083. https://doi.org/10.1177/1178222619829083

Kumar, S., Nilsen, W. J., Abernethy, A., Atienza, A., Patrick, K., Pavel, M., …., & Swendeman, D. (2013). Mobile health technology evaluation: the mhealth evidence workshop. *American Journal of Preventive Medicine, 45*(2), 228–236. https://doi.org/10.1016/j.amepre.2013.03.017

Laerd Statistics (2018). Spearman's rank-order correlation using SPSS statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Laerd Statistics (2015). Independent-samples t-test using SPSS statistics. *Statistical tutorials and software guides*. Retrieved from https://statistics.laerd.com/

Lagan, S., D'Mello, R., Vaidyam, A., Bilden, R., & Torous, J. (2021). Assessing mental health apps marketplaces with objective metrics from 29,190 data points from 278 apps. *Acta Psychiatrica Scandinavica*, *144*(2), 201–210. https://doi.org/10.1111/acps.13306

Lakens, D. (2021). *Sample Size Justification* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/9d3yf

Lattie, E. G. (2019). 10. Technology and college student mental health: Challenges and opportunities. *Frontiers in Psychiatry*, *10*, 5.

Lau, N., O'Daffer, A., Yi-Frazier, J. P., & Rosenberg, A. R. (2021). Popular evidence-based commercial mental health apps: analysis of engagement, functionality, aesthetics, and information quality. *JMIR MHealth and UHealth*, *9*(7), e29689. https://doi.org/10.2196/29689

Lee, R. A., & Jung, M. E. (2018). Evaluation of an mHealth App (DeStressify) on university students' mental health: Pilot trial. *JMIR MENTAL HEALTH*, 16.

Lecomte, T., Potvin, S., Corbière, M., Guay, S., Samson, C., Cloutier, B., Francoeur, A., Pennou, A., & Khazaal, Y. (2020). Mobile apps for mental health issues: meta-review of meta-analyses. *JMIR MHealth and UHealth*, *8*(5), e17458. https://doi.org/10.2196/17458

Lewis-Beck, M., Bryman, A., & Futing Liao, T. (2004). *The SAGE Encyclopedia of Social Science Research Methods*. Sage Publications, Inc. https://doi.org/10.4135/9781412950589

Linardon, J., & Fuller-Tyszkiewicz, M. (2020). Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of Consulting and Clinical Psychology*, *88*(1), 1–13. https://doi.org/10.1037/ccp0000459

Longyear, R. L., & Kushlev, K. (2021). Can mental health apps be effective for depression, anxiety, and stress during a pandemic? *Practice Innovations*, *6*(2), 131–137. https://doi.org/10.1037/pri0000142

Lovejoy, C. A. (2019). Technology and mental health: The role of artificial intelligence. *European Psychiatry*, *55*, 1–3. https://doi.org/10.1016/j.eurpsy.2018.08.004

Lukens, E. P., & McFarlane, W. R. (2004). Psychoeducation as evidence-based practice: Considerations for practice, research, and policy. *Brief Treatment and Crisis Intervention*, *4*(3), 205–225. https://doi.org/10.1093/brief-treatment/mhh019

Lundquist, A. R., Lefebvre, E. J., & Garramone, S. J. (2014). Smartphones: fulfilling the need for immediacy in everyday life, but at what cost? *International Journal of Humanities and Social Science*, *4*(2), 10.

Ly, K. H., Ly, A.-M., & Andersson, G. (2017). 6. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet Interventions*, *10*, 39–46. https://doi.org/10.1016/j.invent.2017.10.002

Malavolta, I. (2016). Web-Based hybrid mobile apps: State of the practice and research opportunities. *2016 IEEE/ACM International Conference on Mobile Software Engineering and Systems (MOBILESoft)*, 241–242. https://doi.org/10.1109/MobileSoft.2016.050

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2021). Smartphone psychological therapy during covid-19: a study on the effectiveness of five popular mental health apps for anxiety and depression. *Frontiers in Psychology*, *12*, 775775. https://doi.org/10.3389/fpsyg.2021.775775

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). 13. Apps with maps—anxiety and depression mobile apps with evidence-based frameworks: systematic search of major app stores. *JMIR MENTAL HEALTH*, 11.

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). 16. Clinical or gimmickal: The use and effectiveness of mobile mental health apps for treating anxiety and depression. *Australian & New Zealand Journal of Psychiatry*, *54*(1), 20–28. https://doi.org/10.1177/0004867419876700

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). 18. Smartphone psychology: New approaches towards safe and efficacious mobile mental health apps. *Professional Psychology: Research and Practice*, *51*(3), 214–222. https://doi.org/10.1037/pro0000278

Mboya, I. B., John, B., Kibopile, E. S., Mhando, L., George, J., & Ngocho, J. S. (2020). Factors associated with mental distress among undergraduate students in northern Tanzania. *BMC Psychiatry*, *20*(1), 28. https://doi.org/10.1186/s12888-020-2448-1

McCloud, T., Jones, R., Lewis, G., Bell, V., & Tsakanikos, E. (2020). 15. Effectiveness of a mobile app intervention for anxiety and depression symptoms in university students: randomized controlled trial. *JMIR MHealth and UHealth*, *8*(7), e15418. https://doi.org/10.2196/15418

McGowan, T. C., & Kagee, A. (2013). Exposure to traumatic events and symptoms of post-traumatic stress among South African university students. *South African Journal of Psychology*, *43*(3), 327–339. https://doi.org/10.1177/0081246313493375

McLachlan, K. J. J., & Gale, C. R. (2018). The effects of psychological distress and its interaction with socioeconomic position on risk of developing four chronic diseases.

Journal of Psychosomatic Research, *109*, 79–85. https://doi.org/10.1016/j.jpsychores.2018.04.004

Melton, L. (2017). Brief introduction to cognitive behavioral therapy for the advanced practitioner in oncology. *Journal of the Advanced Practitioner in Oncology*, *8*(2). https://doi.org/10.6004/jadpro.2017.8.2.6

Mennin, D. S., Ellard, K. K., Fresco, D. M., & Gross, J. J. (2013). United We Stand: Emphasizing commonalities across cognitive-behavioral therapies. *Behavior Therapy*, *44*(2), 234–248. https://doi.org/10.1016/j.beth.2013.02.004

Michaelis, M., Burgess, S., Junne, F., Rothermund, E., Gündel, H., Zipfel, S., Wolf, M., & Rieger, M. A. (2021). Mental health applications for primary and secondary prevention of common mental disorders: attitudes of german employees. *Frontiers in Psychiatry*, *12*, 508622. https://doi.org/10.3389/fpsyt.2021.508622

Milne-Ives, M., Lam, C., De Cock, C., Van Velthoven, M. H., & Meinert, E. (2020). Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. *JMIR MHealth and UHealth*, *8*(3), e17046. https://doi.org/10.2196/17046

Mirowsky, J., & Ross, C. E. (2002). Measurement for a human science. *Journal of Health and Social Behavior*, *43*(2), 152. https://doi.org/10.2307/3090194

Monnier, D. (2020). Woebot: A continuation of and an end to psychotherapy? *Psychothérapies*, *40*(2), 71–78. Cairn.info.

Morris, R. R., Schueller, S. M., & Picard, R. W. (2015). Efficacy of a web-based, crowdsourced peer-to-peer cognitive reappraisal platform for depression: randomized controlled trial. *Journal of Medical Internet Research*, *17*(3), e72. https://doi.org/10.2196/jmir.4167

Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. *Health Psychology*, *34*(Suppl), 1209–1219. https://doi.org/10.1037/hea0000306

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

Neary, M., & Schueller, S. M. (2018). 19. 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

Nerdrum, P., Rustøen, T., & Rønnestad, M. H. (2006). Student psychological distress: a psychometric study of 1750 Norwegian 1st-year undergraduate students. *Scandinavian Journal of Educational Research*, *50*(1), 95-109.

Norton, B. J., & Strube, M. J. (2001). Understanding statistical power. *Journal of Orthopaedic & Sports Physical Therapy*, *31*(6), 307–315. https://doi.org/10.2519/jospt.2001.31.6.307

O'Dea. (n.d.). *Smartphone users in South Africa 2014-2023*. Statista. Retrieved 1 March 2020, from https://www.statista.com/statistics/488376/forecast-of-smartphone-users-in-south-africa/

Ormrod, J. E., & Leedy, P. D. (2015). Practical research: Planning and design. *Teaching Sociology*, *18*(2), 248. https://doi.org/10.2307/1318509

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533–544. https://doi.org/10.1007/s10488-013-0528-y

Park, Y. S., Konge, L., & Artino, A. R. (2020). The positivism paradigm of research: *Academic Medicine*, *95*(5), 690–694. https://doi.org/10.1097/ACM.0000000000003093

Payne, G., & Williams, M. (2011). *The landscape of quantitative methods*. SAGE Publications, Ltd. https://doi.org/10.4135/9781446268384

Price, M., Gros, D. F., McCauley, J. L., Gros, K. S., & Ruggiero, K. J. (2012). Nonuse and dropout attrition for a web-based mental health intervention delivered in a post-disaster context. *Psychiatry: Interpersonal and Biological Processes*, *75*(3), 267–284. https://doi.org/10.1521/psyc.2012.75.3.267

Prochaska, J. J., Vogel, E. A., Chieng, A., Kendra, M., Baiocchi, M., Pajarito, S., & Robinson, A. (2021). 12(1). A therapeutic relational agent for reducing problematic substance use (woebot): development and usability study. *Journal of Medical Internet Research*, *23*(3), e24850. https://doi.org/10.2196/24850

Proudfoot, J., Clarke, J., Birch, M.-R., Whitton, A. E., Parker, G., Manicavasagar, V., Harrison, V., Christensen, H., & Hadzi-Pavlovic, D. (2013). 22. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: A randomised controlled trial. *BMC Psychiatry*, *13*(1), 312. https://doi.org/10.1186/1471-244X-13-312

The [output/code/data analysis] for this paper was generated using Qualtrics software, Version 2022 of Qualtrics. Copyright © 2022 Qualtrics. Qualtrics and all other Qualtrics product or service names are registered trademarks or trademarks of Qualtrics, Provo, UT, USA. https://www.qualtrics.com

Ramachandran, M., Suharwardy, S., Leonard, S. A., Gunaseelan, A., Robinson, A., Darcy, A., Lyell, D. J., & Judy, A. (2020). 74: Acceptability of postnatal mood management through

a smartphone-based automated conversational agent. *American Journal of Obstetrics and Gynecology*, *222*(1), S62. https://doi.org/10.1016/j.ajog.2019.11.090

Resnik, D. (2015). Bioethical issues in providing financial incentives to research participants. *Medicolegal and Bioethics*, 35. https://doi.org/10.2147/MB.S70416

Rowe, M., & Sauls, B. (2020). The use of smartphone apps in clinical practice: A survey of South African physiotherapists. *Open Access*, 8.

Ruggiero, G. M., Spada, M. M., Caselli, G., & Sassaroli, S. (2018). A historical and theoretical review of cognitive behavioral therapies: from structural self-knowledge to functional processes. *Journal of Rational-Emotive & Cognitive-Behavior Therapy*, *36*(4), 378–403. https://doi.org/10.1007/s10942-018-0292-8

Sanderson, W. C., Arunagiri, V., Funk, A. P., Ginsburg, K. L., Krychiw, J. K., Limowski, A. R., Olesnycky, O. S., & Stout, Z. (2020). The nature and treatment of pandemic-related psychological distress. *Journal of Contemporary Psychotherapy*, *50*(4), 251–263. https://doi.org/10.1007/s10879-020-09463-7

Sarwar, M., & Soomro, T. R. (2013). Impact of smartphone's on society. *European journal of scientific research*, *98*(2), 216-226.

Schweizer, M. L., Braun, B. I., & Milstone, A. M. (2016). Research methods in healthcare epidemiology and antimicrobial stewardship—quasi-experimental designs. *Infection Control & Hospital Epidemiology*, *37*(10), 1135–1140. https://doi.org/10.1017/ice.2016.117

Segopolo, M. T., Selemogwe, M. M., Plattner, I. E., Ketlogetswe, N., & Feinstein, A. (2009). A screening instrument for psychological distress in Botswana: validation of the setswana

version of the 28-Item General Health Questionnaire. *International Journal of Social Psychiatry, 55*(2), 149–156. https://doi.org/10.1177/0020764008093448

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.*, xxi, 623–xxi, 623.

Sharma, G. (2017). Pros and cons of different sampling techniques. International journal of applied research, 3(7), 749-752.

Sharp, J., & Theiler, S. (2018). 2. A review of psychological distress among university students: pervasiveness, implications and potential points of intervention. *International Journal for the Advancement of Counselling*, *40*(3), 193–212. https://doi.org/10.1007/s10447-018-9321-7

Shurick, A. A., Hamilton, J. R., Harris, L. T., Roy, A. K., Gross, J. J., & Phelps, E. A. (2012). Durable effects of cognitive restructuring on conditioned fear. *Emotion*, *12*(6), 1393–1397. https://doi.org/10.1037/a0029143

Silver, L. & Johnson, C., 2018. *Internet connectivity seen as having positive impact on life in Sub-Saharan Africa*, Pew Research Center's Global Attitudes Project. United States of America. Retrieved from https://policycommons.net/artifacts/617116/internet-connectivity-seen-as-having-positive-impact-on-life-in-sub-saharan-africa/1597871/ on 05 Nov 2021. CID: 20.500.12592/cc42cq.

Skjuve, M., & Brandtzæg, P. B. (2018). Chatbots as a new user interface for providing health information to young people. In: Andersson, Yvonne; Dahlquist, Ulf; Ohlsson, Jonas (eds). *Youth and news in a digital media environment* – Nordic-Baltic perspectives. 193–212.

Son, C., Hegde, S., Smith, A., Wang, X., & Sasangohar, F. (2020). Effects of COVID-19 on college students' mental health in the United States: interview survey study. *Journal of Medical Internet Research*, *22*(9), e21279. https://doi.org/10.2196/21279

Soroka, L., & Kurkova, K. (2019). Artificial intelligence and space technologies: legal, ethical and technological issues. *Advanced Space Law, 3*. https://doi.org/10.29202/asl/2019/3/11

Schultz, L.M. P. Sprent & N.C. Smeeton (2007). *Applied Nonparametric Statistical Methods* (4th ed.). *Psychometrika* 75, 579–580 (2010). https://doi.org/10.1007/s11336-010-9166-4

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tchounwou, P. (2004). Environmental research and public health. *International Journal of Environmental Research and Public Health*, *1*(1), 1–2. https://doi.org/10.3390/ijerph2004010001

Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). 20. 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

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: a review of the psychiatric landscape. *The Canadian Journal of Psychiatry*, *64*(7), 456–464. https://doi.org/10.1177/0706743719828977

Vogel, E. A., Chieng, A., Robinson, A., Pajarito, S., & Prochaska, J. J. (2021). Associations between substance use problems and stress during covid-19. *Journal of Studies on Alcohol and Drugs*, *82*(6), 776–781. https://doi.org/10.15288/jsad.2021.82.776

Wagner, C., Kawulich, B., & Garner, M. (2012). *EBOOK: Doing social research: A global context*. McGraw Hill.

Wang, X., Markert, C., & Sasangohar, F. (2021). Investigating popular mental health mobile application downloads and activity during the covid-19 pandemic. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 001872082199811. https://doi.org/10.1177/0018720821998110

Weisel, K. K., Fuhrmann, L. M., Berking, M., Baumeister, H., Cuijpers, P., & Ebert, D. D. (2019). Standalone smartphone apps for mental health—A systematic review and meta-analysis. *Npj Digital Medicine*, *2*(1), 118. https://doi.org/10.1038/s41746-019-0188-8

Weiss, L. A., Westerhof, G. J., & Bohlmeijer, E. T. (2016). Can we increase psychological well-being? the effects of interventions on psychological well-being: a meta-analysis of randomized controlled trials. *PLOS ONE*, *11*(6), e0158092. https://doi.org/10.1371/journal.pone.0158092

Wheaton, B. (2007). The twain meet: Distress, disorder and the continuing conundrum of categories (comment on Horwitz). *Health: An Interdisciplinary Journal for the Social Study of Health, Illness and Medicine*, *11*, 303–319. https://doi.org/10.1177/1363459307077545

White, H., & Sabarwal, S. (2014). Quasi-experimental design and methods. *Methodological briefs: Impact evaluation*, *8*(2014), 1-16.

Wiles, R., Crow, G., Heath, S., & Charles, V. (2008). The management of confidentiality and anonymity in social research. *International Journal of Social Research Methodology*, *11*(5), 417–428. https://doi.org/10.1080/13645570701622231

Wilson, S. R. (2018). *Selecting a research methdology*. SAGE Publications, Inc. https://doi.org/10.4135/9781483381411

Woebot Health. (2021). Woebot Health. Retrieved 10 May 2022, from https://woebothealth.com/

World Health Organization. (2018). *Mental health: Strengthening our response*. https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response

Yaghubi, H., Karimi, M., Omidi, A., Barouti, E., & Abedi, A. (2012). Validity and factor structure of the General Health Questionnaire (GHQ-12) in university students. *Journal of Behavioral Sciences*, 6 (2), 153-160.

Yoon, F. B., Huskamp, H. A., Busch, A. B., & Normand, S.-L. T. (2011). Using multiple control groups and matching to address unobserved biases in comparative effectiveness research: An observational study of the effectiveness of mental health parity. *Statistics in Biosciences*, *3*(1), 63–78. https://doi.org/10.1007/s12561-011-9035-4

Zwarun, L. (2018). *Matched Groups Design*. SAGE Publications, Inc. https://doi.org/10.4135/9781483381411

APPENDIX A: INSTITUTIONAL APPROVAL LETTER

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Humanities
Fakulteit Geesteswetenskappe
Lefapha la Bomotho

Humanities 100. Since 1919

15 November 2021

Dear Mr KL Gernandt

Project Title:	The effect of a CBT-based mental health app (Woebot) on the psychological distress levels of university students
Researcher:	Mr KL Gernandt
Supervisor(s):	Dr B Moteleng
Department:	Psychology
Reference number:	15073778 (HUM029/0221)
Degree:	Masters

I have pleasure in informing you that the above application was **approved** by the Research Ethics Committee on 15 November 2021. Data collection may therefore commence.

Please note that this approval is based on the assumption that the research will be carried out along the lines laid out in the proposal. Should the actual research depart significantly from the proposed research, it will be necessary to apply for a new research approval and ethical clearance.

We wish you success with the project.

Sincerely,

Prof Karen Harris
Chair: Research Ethics Committee
Faculty of Humanities
UNIVERSITY OF PRETORIA
e-mail: tracey.andrew@up.ac.za

APPENDIX B: PERMISSION FOR RESEARCH

Research permission request Inbox ×

Keanan <keanangernant@gmail.com> Fri, Oct 15, 2021, 1:35 PM

to Matete

Dear Dr Madiba,

RE: Research permission: The effect of a CBT-based mental health app on the psychological distress levels of university students

This serves as a formal request to undertake research with second-and third-year humanities students aged between 18-24 years at the University of Pretoria.

My research topic is: "The effect of a CBT-based mental health app on the psychological distress levels of university students". This research data will be collected as part of a full dissertation for the Department of Psychology at the University of Pretoria. I am currently a registered academic master's student in Psychology.

The research proposal will include the following sections: brief description and statement of the research question, introduction and rationale, research hypothesis, summary of preliminary literature review, theoretical framework, proposed research methodology and ethical considerations.

The purpose of the study is to determine the effect of a mental health mobile application (Woebot) on the psychological distress levels of university students over a period of three weeks.

The proposed research will involve quantitative research based on a quasi-experimental matched comparison group design. The study flyer will be advertised to University of Pretoria students on social media sites including Academia.edu, LinkedIn, and Facebook which will contain a link to the questionnaire.

The questionnaire consists of a Qualtrics survey battery intended to collect information such as the participant's email address, demographic information, and psychological distress levels over the last two weeks. The information and outcome scores will determine whether a participant is suited for inclusion in the study based on the inclusion criteria and cutoff scores for each measure.

A total of 100 students in the intended sample size for the study will be divided into two groups. One of the groups will receive the experimental condition (Woebot app) and the other group will receive a nature conservation assignment (ecological conservation and sustainability). The app will be made available for the control condition after the three-week period is over through an email notification. All research data for this study will be collected using Qualtrics, Excel and SPSS version 26.

The final draft of the research findings will be made available to you. **Attached to this letter is an approval letter for the research study received the UP Survey Coordinating Committee on the 14th of October 2021 and a permission sheet to sign electronically if you agree with the use of students that meet the inclusion criteria of this study, namely second-and third-year humanities students aged between 18 and 24 currently enrolled at the University of Pretoria**. If you require any further information relating to my study, you can email me at keanangernant@gmail.com. My research supervisor is Dr. Barnard Motileng who can be reached through email at benny.motileng@up.ac.za.

Kind Regards

RESEARCH PERMISSION SIGNATURE LINE

Signature: Date: 15 October 2021

X

Matete Madiba
Doctor

Signature-Researcher Date

_____ _____ _____01/07/21_____

K L Gernandt MA Psychology

(Cell: 0823860264)

(Email: keanangernant@gmail.com)

Signature-Supervisor Date

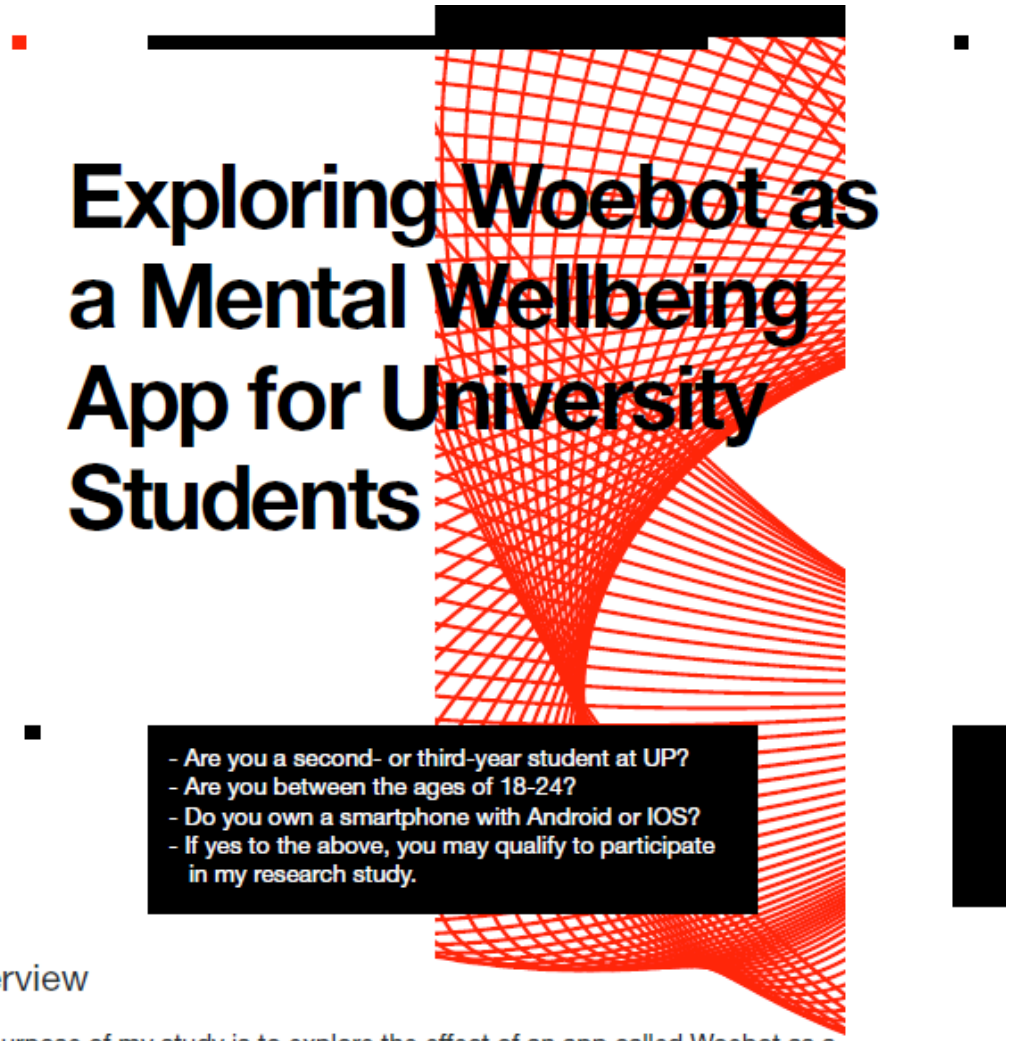
_ _____22/07/21_____

Dr Barnard Buti Motileng

(Cell: 0798820444)

(Email: benny.motileng@up.ac.za)

APPENDIX C: PARTICIPANT SAMPLING FLYER



Overview

The purpose of my study is to explore the effect of an app called Woebot as a mental wellbeing resource for university students over three weeks. Woebot is a chatbot app that helps people monitor their moods and learn more about themselves. The only participation requirements are that you use the Woebot app as regularly as you can and to complete a short app usage survey at the end of each week. The study is entirely online, and your identity will not be disclosed to anyone, and your data will be kept strictly confidential.

Please take our short survey to find out more by clicking the link below:

https://pretoria.eu.qualtrics.com/jfe/form/SV_87D2AA1sqt3CfoW

APPENDIX D: PARTICIPANT INFORMATION SHEET

Hello, my name is Keanan Gernandt. I am currently a psychology master's student at the Faculty of Humanities at the University of Pretoria. You are being invited to take part in my research study. Before you decide to participate in this study, it is crucial that you understand why the research is being done and what it will involve. Please take some time to read the following information carefully, as it explains the details of this research project. Please feel free to contact me using the contact details provided if anything is unclear or you need more information.

WHAT IS THE PURPOSE OF THE STUDY?

The purpose of this study is to investigate the effect of a mental wellbeing smartphone app called Woebot. Woebot is a chatbot app that helps people track their moods and learn more about themselves. This research study could help us determine the app's usefulness as a mental wellbeing resource for students at South African universities.

WHY HAVE YOU BEEN INVITED TO PARTICIPATE?

You are eligible to participate if you are a second or third-year student currently enrolled at the University of Pretoria between 18 and 24.

You have also met the following criteria:

- Own a smartphone with the Android or iOS operating system.
- You have daily internet access.

WHAT IS THE NATURE OF YOUR PARTICIPATION IN THIS STUDY?

Your participation in this study will entail one of two conditions.

- If you are part of the active participation group, you will receive the Woebot app assignment. This will include downloading and interacting with the Woebot app on a regular basis for 21 days. At the end of each week, you will be emailed a short survey to record your app usage over the previous week.
- If you are a part of the comparison group, you will only be required to complete the initial survey, followed by another survey that will be emailed to you in three weeks. You will not have to use the Woebot app at all during the three-week study period.

CAN YOU WITHDRAW FROM THIS STUDY EVEN AFTER HAVING AGREED TO PARTICIPATE?

Participating in this study is voluntary, and you are under no obligation to consent to participation. An agreement to participate will be acknowledged by signing an informed consent accompanying this participation information sheet. If you do not decide to participate in the study, you won't be penalised, and there will be no consequences for doing so.

WILL THE INFORMATION THAT YOU CONVEY TO THE RESEARCHER BE KEPT CONFIDENTIAL?

Anonymity and confidentiality will be ensured by assigning code numbers to you and not using your name for any research notes and documents. The reporting of findings will also be anonymous, and only me and my supervisor will have access to the data.

WHAT ARE THE POTENTIAL BENEFITS OF TAKING PART IN THIS STUDY?

This study may help university-based app developers create more targeted apps to assist university students. By participating in this study, you will contribute to the advancement of research into the practical use of apps such as Woebot for mental wellbeing.

WHAT ARE THE ANTICIPATED RISKS FROM TAKING PART IN THIS STUDY?

There are no anticipated risks in this research. However, a psychologist at the University of Pretoria will be on standby if any emergencies arise.

WHAT WILL HAPPEN IN THE UNLIKELY EVENT THAT SOME FORM OF DISCOMFORT OCCUR AS A RESULT OF TAKING PART IN THIS RESEARCH STUDY?

Should any distress occur during the study, you may directly contact the psychologist on standby Dr Sharon Sibanda 012 420 3685 or sharon.sibanda@up.ac.za. You may also contact any of the following emergency mental healthcare service providers if you need immediate assistance:

- UP careline Call: (0800 747 747)
- SA Federation for Mental Health Telephone: +27(11) 781 1852 or Email: info@safmh.org
- Cipla 24hr Mental Health Helpline (0800 456 789)
- Adcock Ingram Depression and Anxiety Helpline (0800 70 80 90)

HOW WILL THE RESEARCHER(S) PROTECT THE SECURITY OF DATA?

Electronic information will be stored for a minimum of 5 years. Future use of the stored data will be subject to further Research Ethics Review and approval if applicable. Any information or physical data about research participants will be kept in a locked cabinet, and electronic data will be kept in a password-protected file in the Department of Psychology.

WHAT WILL THE RESEARCH DATA BE USED FOR?

Data gathered from participants will be used to fulfil the requirements of a psychology master's dissertation and to be published as a journal article or conference paper.

WILL I BE PAID TO TAKE PART IN THIS STUDY?

No, you will not be paid to take part in this study. The study is not meant to generate any income only to increase research knowledge.

HAS THE STUDY RECEIVED ETHICS APPROVAL?

Yes, this study has received written approval from the Research Ethics Committee of Faculty of Humanities, University of Pretoria.

HOW WILL YOU BE INFORMED OF THE FINDINGS/RESULTS OF THE RESEARCH?

If you want to see the research study findings, you can send me an email request at keanangernant@gmail.com and I will send you the findings after the final research report has been reviewed and approved.

WHO SHOULD YOU CONTACT IF YOU HAVE CONCERNS, COMPLAINTS OR QUERIES RELATING TO THE STUDY?

If you have any questions about this study or your rights as a research participant or experience any adverse effects from your participation in this study, you can contact my supervisor or me (contact details provided below).

Thank you for reading this information sheet and agreeing to participate in this study!

Researcher Contact Information

Name Surname Keanan Gernandt

Contact number 0823860264

Email address: keanangernant@gmail.com

Supervisor Contact Information

Name Dr Barnard Moteleng

Contact number 012 442 02907

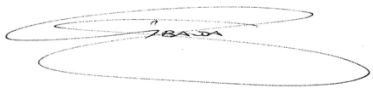
Email address: benny.motileng@up.ac.za

APPENDIX E: PSYCHOLOGIST ON STANDBY PERMISSION LETTER

Dear Ms Sibanda

My name is Keanan Gernandt I am conducting research in the Department of Psychology at the University of Pretoria. This research is being conducted as part of my postgraduate degree (Masters in Psychology). The title of this study is "The effect of a CBT-based mental health app on the psychological distress levels of university students". The purpose of the study is to determine the effect of a mental health mobile application (Woebot) on the psychological distress levels of university students over a period of three weeks. This study will be conducted under the supervision of Dr Benny Motileng (Research Psychologist, University of Pretoria (UP). I am requesting your permission to be on standby for psychological support during the study period in the event that a student reports experiencing significant distress and needs urgent assistance. Your help would be greatly appreciated. If you require any further information relating to the study, you can email me at u15073778@tuks.co.za or contact me on 0823860264 or my supervisor Dr Motileng at benny.motileng@up.ac.za

Kind regards

Signature-Psychologist  Date _____01/07/21_____

PS: 0113212

Signature-Researcher Date

_____ _____ _____01/07/21 ____

APPENDIX F: WOEBOT INTERVENTION INSTRUCTIONS

Woebot App Instructions

Welcome and thank you for taking part in my research study!

Your participation in the study is scheduled to begin on **Monday, 11 April 2022 and will end on Monday 2 May 2022.**

The smartphone app that you will be using is called **Woebot.**

Woebot is an automated conversational agent (chatbot) app who helps you monitor mood and learn about yourself. Drawing from a therapeutic framework known as Cognitive Behavior Therapy, Woebot asks people how they're feeling and what is going on in their lives in the format of brief daily conversations. Woebot also talks to you about mental health and wellness and sends you videos and other useful tools depending on your mood and needs at that moment. You can think of Woebot as a choose-your-own-adventure self-help book that can store all your entries and gets more specific to your needs over time.

Weekly Woebot Usage Instructions

- You will be asked to use the app as many of the days out of the week as you can for 3 weeks starting on Monday 11 April 2022.
- The average daily interaction with Woebot is only a few minutes, but you can use the app as much as you want each day.
- At the end of each week, I will email you a short survey to ask how many days you used the Woebot app. I will not ask for personal information or exact app usage details.

- You can choose to record this manually or you can view the mood tracker in the app, which will show you how many days you checked in with the app.
- Although I obtained official permission from Woebot labs to use their app for this study, I will not have access to the details of your app use or any information you might share on the app.
- If you would like to view Woebot's privacy policy, here is the link: https://woebothealth.com/privacy-webview/#:~:text=Woebot%20operates%20and%20uses%20appropriate,%2C%20unauthorized%20disclosure%2C%20or%20access.
- ***I will also ask that you please not use any other mental wellbeing apps during this time. However, if you do happen to use any, please indicate this in the short survey that will be sent to you at the end of each week**.

PHASE 1: WOEBOT APP INSTALLATION AND SETUP

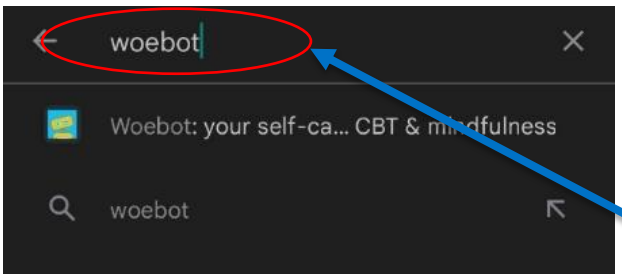
If you have an **Android phone**, please click on this link:

- https://play.google.com/store/apps/details?id=com.woebot&hl=en_ZA&gl=US

If you have an **Apple phone**, please click on this link:

- https://apps.apple.com/us/app/woebot-your-self-care-expert/id1305375832

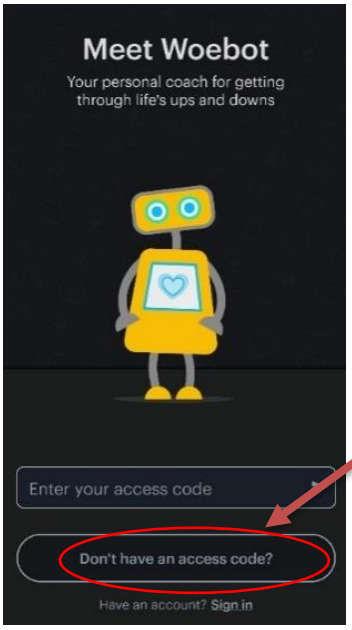
Alternatively, you can search for the app as indicated below:



STEP 1

A. Please open the app store on your smartphone (Google Play for Android or the App Store for iOS).

B. In the app search bar please type in 'Woebot' as indicated to the left.

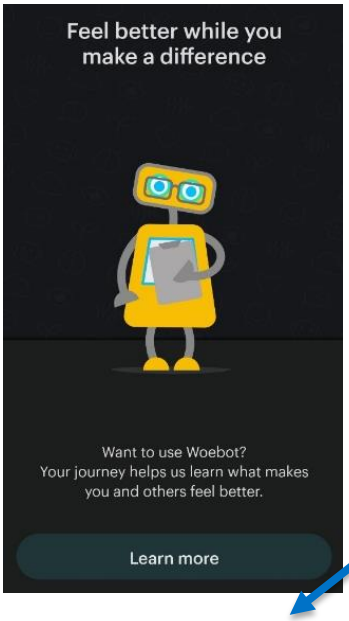


STEP 2

A. Next, please install the Woebot app on your phone. Ensure that you have at least **15Mb** of free space on your phone.

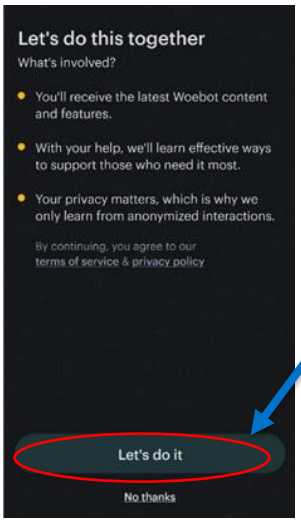
B. Please **open the app** after you have installed it.

C. On the opening screen select the option '**Don't have an access code**?'.



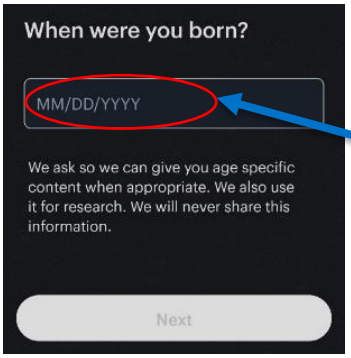
STEP 3

Select the '**Learn more**' option.



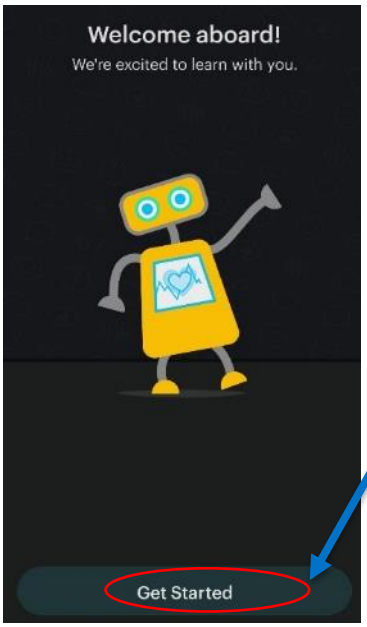
STEP 4

Read 'What's involved' and select the '**Let's do it'** option.



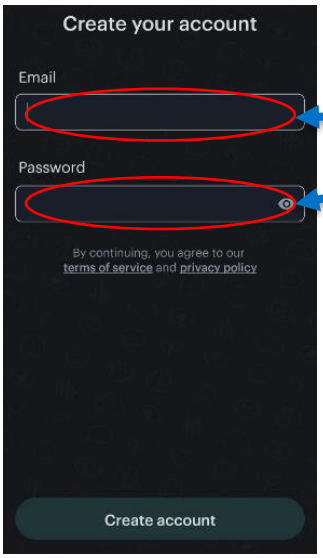
STEP 5

Please enter your age.



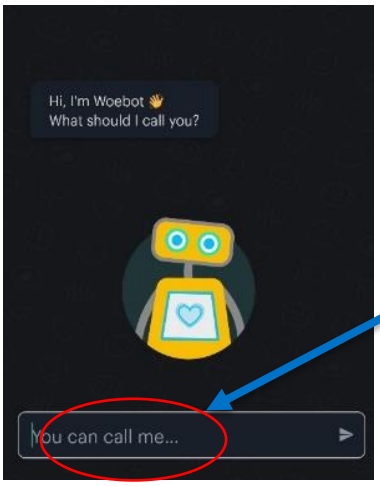
STEP 6

On this screen, please select the '**Get Started**' option.



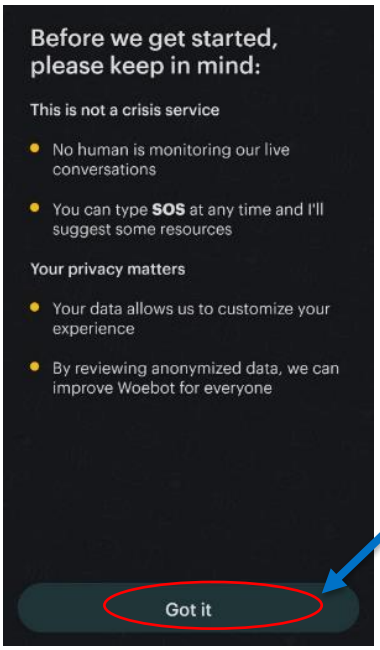
STEP 7

On this screen you will be asked to create an account. Please enter your email address and a password. You can either use your personal email address or your UP email address.



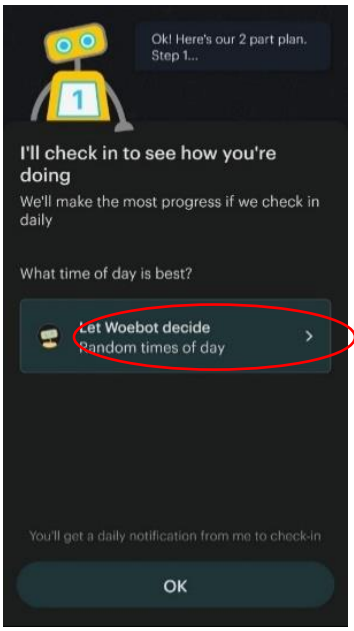
STEP 8

On this screen, please enter your name or the name you would like Woebot to call you



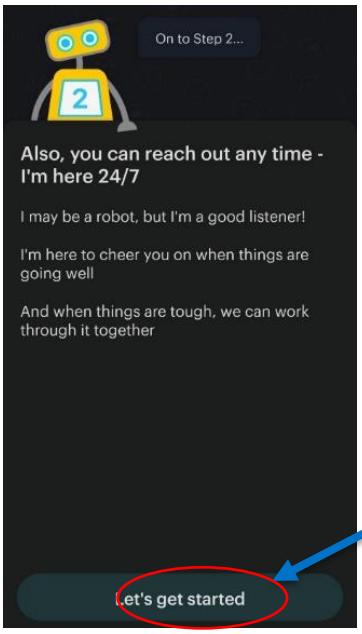
STEP 9

Read 'Before we get started…' and select 'Got it'.



STEP 10

On this screen, please choose a time of the day to receive the notification from Woebot. If you cannot decide, select the 'Let Woebot decide' option.

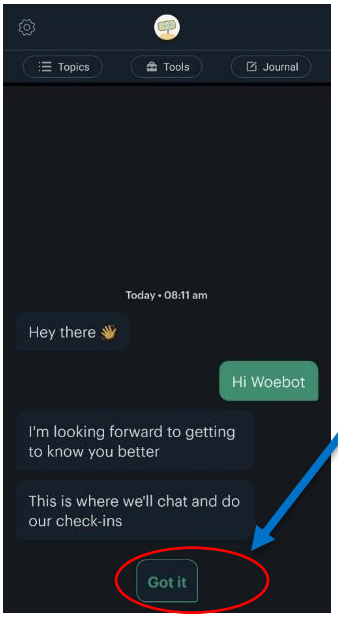


STEP 11

Select 'Let's get started'

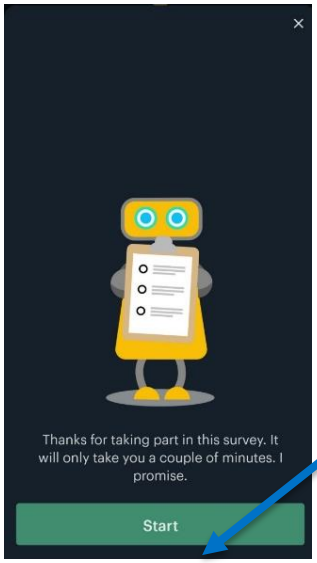
PHASE 2: WOEBOT INTERACTION INSTRUCTIONS

- After installing and setting up the app you are now ready to begin your interaction with the Woebot app.



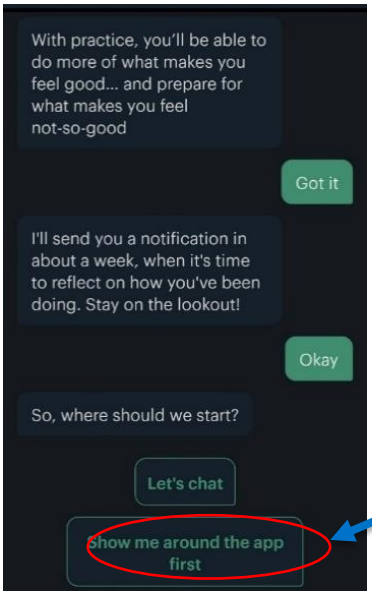
STEP 1

This screen presents the main interface where you will be chatting with Woebot. Select the 'Got it' option to get your first interaction going.



STEP 2

Woebot will then ask you to complete a short survey to check-in on how you've been doing recently. Please select **Start** to begin the survey



STEP 3

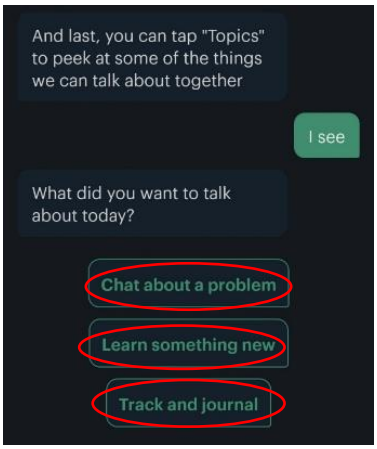
After you have completed the survey, please choose the '**Show me around the app first'** option.

DAILY WOEBOT INTERACTION OPTIONS

- Woebot will usually give you a few interaction options you can choose from each day namely, **chat about a problem**, **learn something new** and **track and journal**.
- These options are described below.
- ***It is up to you to choose how you would like to interact with Woebot each day.***

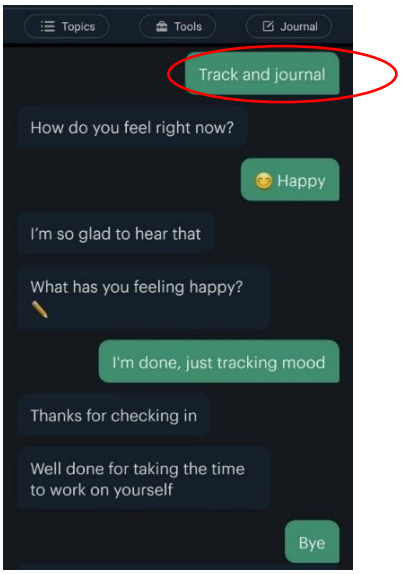
Option A: 'Chat about a problem'

- You can choose '**chat about a problem**' if you would like some help for something like anxiety, depression, loneliness, procrastination etc.



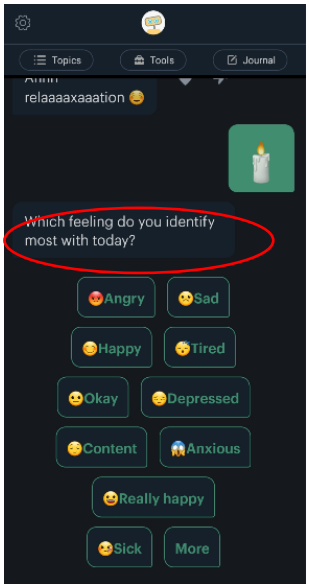
Option B: 'Learn something new'

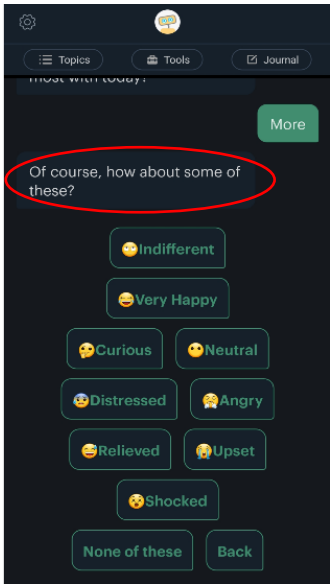
If you choose to '**learn something new**', Woebot will provide interesting and practical short lessons on a variety of topics related to mental wellbeing.



Option C: 'Track and Journal'

If you select the 'track and journal' option, you can check in your current mood and track your patterns over time.



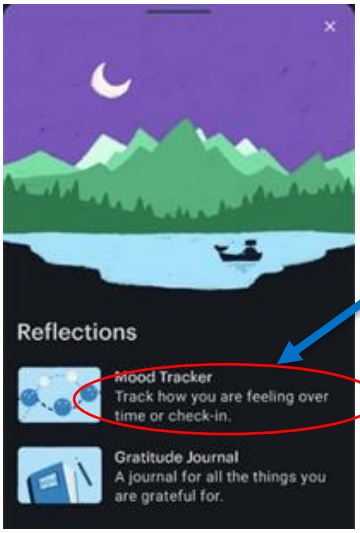


Woebot will regularly ask you to check in your current mood as shown here. **Please do this as often as possible throughout each week.**

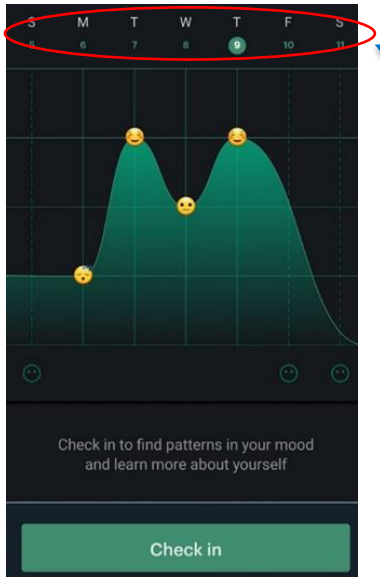
PHASE 3: END OF WEEK SURVEY INSTRUCTIONS

1) At the end of each week, you will receive an email with a link titled '**Woebot App Weekly Usage Survey**' in which you will be asked about your daily usage of the Woebot app over the past 7 days.

2) If you can't remember how many days you interacted with the app, check the **mood tracker** on the Woebot app, which will show a graph with emojis indicating the days you checked in with the app as shown below.



The mood tracker can be found under the **Reflections** tab.



The mood tracker shows how many days you checked in with Woebot last week.

3) You can also use the **Woebot Usage Log** on the next page to track the days you used the Woebot app.

4) Please complete the **end of week questionnaire**, even if you did not use the app at all during the week, as it is essential for data collection purposes.

Woebot Usage Log

Please underline or highlight the days you interacted with the Woebot app in the past week

WEEK 1 (Monday 11 April - Sunday 17 April)						
Monday 11 April	Tuesday 12 April	Wednesday 13 April	Thursday 14 April	Friday 15 April	Saturday 16 April	Sunday 17 April

WEEK 2 (Monday 18 April - Sunday 24 April)						
Monday 18 April	Tuesday 19 April	Wednesday 20 April	Thursday 21 April	Friday 22 April	Saturday 23 April	Sunday 24 April

WEEK 3 (Monday 25 April - Sunday 1 May)						
Monday 25 April	Tuesday 26 April	Wednesday 27 April	Thursday 28 April	Friday 29 April	Saturday 30 April	Sunday 1 May

*** On Monday 2 May you will receive an email with a final questionnaire. Please complete this survey.**

THANK YOU VERY MUCH FOR TAKING PART IN MY STUDY!

APPENDIX G: QUALTRICS PRE-STUDY SURVEY

Thank you for taking the time to complete this survey!

If you would like to find out more about the details of the study please download and read the **participation information sheet** below.

ATTACHED PDF DOCUMENT: PARTICIPATION INFORMATION SHEET

Q1 Are you currently a registered **second** or **third-year** student at the University of Pretoria?

☐ Yes

☐ No

Q2 If you chose "Yes," please specify which year you are currently in.

☐ Second Year

☐ Third Year

Q3 Please select the faculty you are in

- ○ Faculty of Economic and Management Sciences
- ○ Faculty of Education
- ○ Faculty of Education
- ○ Faculty of Engineering, Built Environment and Information Technology
- ○ Faculty of Health Sciences
- ○ Faculty of Humanities
- ○ Faculty of Law
- ○ Faculty of Natural and Agricultural Sciences
- ○ Faculty of Theology and Religion
- ○ Faculty of Veterinary Science
- ○ Gordon Institute of Business Science (GIBS)

Q4 What is your current age?

○ Age _____

Q5 Do you own a smartphone with Android or IOS operating systems?

☐ Yes

☐ No

Q6 Do you have internet access?

☐ Yes

☐ No

Q7 Please provide your **name and surname** below

_____

Q8 Please provide your **current email address**

_____

Q9 What is your gender?

○ Male

○ Female

○ Non-binary / third gender

○ Prefer not to say

Q10 What is your racial category?

○ African/Black

○ White/Caucasian

○ Coloured

○ Asian/Indian

Instructions; Please indicate how often you experienced the following occurrences recently (**over the past 1-2 weeks**) by selecting the appropriate option. There are no right or wrong answers. Do not spend too much time on any statement.

Q1 I was able to concentrate on what I was doing

○ Better than usual

○ Same as usual

○ Less than usual

- ○ Much less than usual

Q2 I lost sleep over worry

- ○ Not at all
- ○ No more than usual
- ○ Rather more than usual
- ○ Much more than usual

Q3 I felt like I was playing a useful part in things (society, school, work, family etc.)

- ○ More so than usual
- ○ Same as usual
- ○ Less than usual
- ○ Much less useful

Q4 I felt capable of making decisions about things

- ○ More so than usual

- ○ Same as usual
- ○ Less so than usual
- ○ Much less capable

Q5 I felt constantly under strain

- ○ Not at all
- ○ No more than usual
- ○ Rather more than usual
- ○ Much more than usual

Q6 I felt I couldn’t overcome difficulties in my life

- ○ Not at all
- ○ No more than usual
- ○ Rather more than usual
- ○ Much more than usual

Q7 I was able to enjoy my normal day-to-day activities

- More so than usual
- Same as usual
- Less so than usual
- Much less than usual

Q8 I was able to face up to my problems

- More so than usual
- Same as usual
- Less so than usual
- Much less able

Q9 I felt somewhat unhappy and depressed

- Not at all
- No more than usual
- Rather more than usual

- Much more than usual

Q10 I felt like I was losing confidence in myself

- Not at all
- No more than usual
- Rather more than usual
- Much more than usual

Q11 I had thoughts of being a worthless person

- Not at all
- No more than usual
- Rather more than usual
- Much more than usual

Q12 I felt reasonably happy, all things considered

- More so than usual

○ About same as usual

○ Less so than usual

○ Much less than usual

Participation choice; Please indicate whether you would like to take part as a

○ Comparison Group Participant (no app use)

○ Active Participant Group (app use)

Info: If you select **active participant**, you will receive an email in a few days with instructions on how to download and use the Woebot app.

APPENDIX H: QUALTRICS POST-STUDY SURVEY

Woebot Treatment Group Survey

Q1 Well done for completing the final week of the study! This is the final survey of the study. It will ask you about this week's Woebot use and some questions about your recent mental wellbeing. **Please complete this survey as it is essential for data collection purposes.**

Please indicate below how many days out of this past week (**from Monday 25 April until Sunday 1 May)** you used/interacted with the Woebot app.

Q2 If you can't remember how many days you used the app, you can access the **mood tracker** under the reflections tab in the app. For example, below, the user's interaction count for that week was 4 days

Q3 Alternatively, you can use the **Woebot Usage Log** included with the Participation Instructions Sheet to track the number of days you used the app. For example, below, the user interacted with the app **5 days** out of the week.

Q4 Over the **last week** (Monday 25 April until Sunday 1 May) I interacted with the Woebot app

- ○ No days/ 0 days
- ○ 1 Day
- ○ 2 Days
- ○ 3 Days
- ○ 4 Days
- ○ 5 Days
- ○ 6 Days
- ○ Everyday/7 Days

Q5 Did you use any mental health app besides Woebot over the last 7 days?

- ○ Yes
- ○ No

Q6 If yes, what was the app/s called?

_____

Instructions Please indicate how often you experienced the following occurrences recently (**over the past 2 weeks**) by selecting the appropriate option. There are no right or wrong answers. Do not spend too much time on any statement.

(GHQ-12 Questions Again)

Comparison Group Post-Study Survey

Q1 Thank you for taking this survey! This is the 3-week later follow-up survey for the smartphone app (Woebot) study.

The survey should take approximately 2 minutes or less to complete.

Q1 Did you use any mental health/wellbeing apps over the last 3 weeks?

- ○ Yes
- ○ No

Q2 If yes, what was the app/s called?

_____

Instructions Please indicate how often you experienced the following occurrences recently (**over the past 2 weeks**) by selecting the appropriate option. There are no right or wrong answers. Do not spend too much time on any statement.

(GHQ-12 Questions Again)

APPENDIX I: INFORMED CONSENT FORM

Q1 **Please complete the informed consent questions after having read the participation information sheet below**

Q2 I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any consequences or penalties.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q3 I understand that information collected during the study will not be linked to my identity, and I give permission to the researchers of this study to access the information.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q4 I understand that this study has been reviewed by and received ethics clearance from Research Ethics Committee Faculty of Humanities of the University of Pretoria.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q5 I understand who will have access to personal information and how the information will be stored with a clear understanding that I will not be linked to the information in any way.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q6 I give consent that data gathered may be used for a dissertation.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q7 I understand how to raise a concern or make a complaint.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q8 I give permission to be quoted directly in the research publication while remaining anonymous.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q9 I have had sufficient opportunity to ask questions, and I agree to take part in the above study.

- ○ Agree
- ○ Disagree
- ○ Not applicable

Q10 Please Insert your name in the text box down below:

I, _____ (participant name), confirm that the person asking my consent to take part in this research informed me about the nature, procedure, potential benefits and anticipated inconvenience of participation.

_____

Q11 Please provide your signature below by drawing on the SIGN HERE line:

Q12 Person Taking Consent: Keanan Gernandt (Researcher)

Signature:

APPENDIX J: WEEKLY WOEBOT USAGE REPORT

Q1 Well done for completing the first week of the study! Please indicate below how many days out of this past week (**from Monday 11 April until Sunday 17 April)** you used/interacted with the Woebot app.

Q2 If you can't remember how many days you used the app, you can access the **mood tracker** under the reflections tab in the app. For example, below, the user's interaction count for that week was 4 days

---

Q3 Alternatively, you can use the **Woebot Usage Log** included with the Participation Instructions Sheet to track the number of days you used the app. For example, below, the user interacted with the app **5 days** out of the week.

Q4 Over the **last week** (From Monday 11 April until Sunday 17 April) I interacted with the Woebot app

- ○ No days/ 0 days
- ○ 1 Day
- ○ 2 Days
- ○ 3 Days
- ○ 4 Days
- ○ 5 Days
- ○ 6 Days
- ○ Everyday/7 Days

Q5 Did you use any mental wellbeing app besides Woebot over the last 7 days?

- ○ Yes
- ○ No

Q6 If yes, what was the app/s called?

_____