I Bet There's an App for That:

Using Mental Health Apps for Reducing Anxiety and Depression

Jamie M Marshall

M.Psych. (Clinical), B.A. (Hons.) (Psych.), Grad. Dip. Soc. Sci., MAPS FCCLP

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

University of New England

Armidale, NSW, Australia

1st May 2021

Candidate Certification

I certify that the work in this thesis, entitled "***I Bet There's an App for That: Using Mental Health Apps for Reducing Anxiety and Depression***", has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree to any other university or institution other than the University of New England.

I also certify that the thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and in the preparation of the thesis itself has been appropriately acknowledged.

In addition, I certify that all information sources and literature used are indicated in this thesis.

1st May 2021

Acknowledgements

I would like to thank my Principal Supervisor, Prof Debra Dunstan, and Co-Supervisor, Dr Warren Bartik, for their support and guidance over the many years that it took to bring this thesis to fruition. They showed me that there was always a will and a way through any obstacle when I struggled to come up with either.

Thank you to the participants who took part in the research, and for their patience in receiving text messages from me every day for a period of 16 weeks.

Thank you to the relevant University of New England staff for always helping me to navigate the various necessary bureaucratic processes. Also, thank you to the University of New England management and School of Psychology staff for making it possible for me to undertake this research degree with resource and financial support.

Finally, thank you to all my family members and friends for their support during my PhD journey, particularly Stephanie, Casimir, Clare, Greg, and Bev.

List of Publications

This thesis is submitted as a "Thesis by Publication". The following peer-reviewed publications have emanated from the research contained in this thesis:

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Clinical or gimmickal: The use and effectiveness of mobile mental health apps for treating anxiety and depression. *Australian and New Zealand Journal of Psychiatry, 54*(1), 20-28. https://doi.org/10.1177%2F0004867419876700

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

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2019). The digital psychiatrist: In-search of evidence-based apps for anxiety and depression. *Frontiers in Psychiatry, 10*, Article 831. https://doi.org/10.3389/fpsyt.2019.00831

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Apps with maps – Anxiety and depression mobile apps with evidence-based frameworks: Systematic search of major app stores. *JMIR Mental Health, 7*(6), Article e16525. https://doi.org/10.2196/16525

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Positive psychology mobile applications for increasing happiness and wellbeing – A systematic app store review. R U appy? *European Journal of Applied Positive Psychology, 4*(12), 1-10. https://www.nationalwellbeingservice.org/volumes/volume-4-2020/volume-4-article-12/

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Effectiveness of using mental health mobile apps as digital antidepressants for reducing anxiety and depression: Protocol for a multiple baseline across-individuals design. *JMIR Research Protocols, 9*(7), Article e17159. https://doi.org/10.2196/17159

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). The Role of digital mental health resources to treat trauma symptoms in Australia during COVID-19. *Psychological Trauma: Theory, Research, Practice, and Policy, 12*(S1), S269-S271. http://dx.doi.org/10.1037/tra0000627

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Treating psychological trauma in the midst of COVID-19: The role of smartphone apps. *Frontiers in Public Health, 8*, Article 402. https://doi.org/10.3389/fpubh.2020.00402

Note to Examiners

This thesis contains seven publications that have been integrated as chapters into one manuscript. As the University of New England requires that these chapters be presented as published in the relevant journal, some explanatory notes will assist the examiner navigate its contents.

Firstly, some published chapters, and chapters submitted for publication, are for journals that use American English, while other published chapters, and chapters submitted for publication, are for journals that use traditional English. This explains the shifts between traditional and American English spelling throughout the thesis chapters. Similarly, the placement of Tables and Figures within manuscripts can vary between journals and this will be observed in some chapters. However, the numbering of Tables and Figures has been altered to reflect the chronological order of how they appear in this thesis. Furthermore, the formatting and subheading requirements of journals vary, and this explains why subheading styles may vary from chapter to chapter; for example, some journals require an article to start with the subheading, *Introduction*, while others require no such subheading for the Introduction section. The chapters and sections that have not been submitted for publication conform to the standards of the *Publication Manual of the American Psychological Association 7th Edition* and use traditional English.

Secondly, there is some overlap in content between chapters as a result of the number of publications. I apologise for this, but it was unavoidable.

Thirdly, to reduce space, all references appear in a consolidated reference list towards the end of the thesis rather than having individual reference lists at the end of each chapter. Again, it was recognised that there was repetition of references from chapter to chapter. Similarly, there are two Appendices: Appendix A contains supplementary data for the main

intervention study (Chapter 10 – Study 4); and Appendix B is a consolidated collection of supplementary information and data from the pilot study (Chapter 7 – Study 3) and other chapters.

Fourthly, to reduce space, some minor editing of published chapters has occurred (such as deleting journal-specific notes about authors' contributions and the like). Examiners are encouraged to seek out the published works if interested in comparing with the versions in this thesis.

Finally, there is a reference to a six-month follow-up in some of the published articles. This six-month follow-up data will be analysed, written up and published independently to this thesis.

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Treating Psychological Trauma in the Midst of COVID-19: The Role of Smartphone Apps. *Frontiers in Public Health*, 8. doi:10.3389/fpubh.2020.00402

No proof of publication could be located for the following chapters:

Chapter 4

Study 1: Smartphone Apps as Digital Antidepressants: A Systematic Review of the Independent Research into Apps for Reducing Anxiety and Depression

Chapter 7

Study 3: Can a Smartphone App Make You Feel Super Better? A Pilot Study Utilizing a Multiple Single-Case Design

Contents

List of Tables

List of Figures

Abbreviations and Symbols

Abbreviations

ACT	Acceptance and commitment therapy
AIC	Average inter-item correlation
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
ANZCTR	Australian and New Zealand Clinical Trials Registry
ARIMA	Autoregressive integrative moving average
App	Application (for a mobile device)
AUD	Australian dollars
BAI	Beck Anxiety Inventory
BDI	Beck Depression Inventory
CBT	Cognitive behavioural therapy
CES-D	Center for Epidemiologic Studies Depression Scale
COVID-19	Coronavirus disease
CSI	Clinical significance index
DASS-21	Depression Anxiety Stress Scale-21 short-form version
DBT	Dialectical behaviour therapy
EMDR	Eye movement desensitisation and reprocessing
ERP	Exposure response prevention
GAD	Generalised anxiety disorder
GAD-7	Generalized Anxiety Disorder – 7-Item Scale
HADS-A	Hospital Anxiety Depression Scale – Anxiety
IPT	Interpersonal therapy
ITSA	Interrupted time series analysis
MARS	Meta-Analysis Reporting Standards

mHealth	Mobile health
OCD	Obsessive compulsive disorder
OQ-45.2	Outcome Questionnaire-45 second edition version
PASSR	Protocol for App Store Systematic Reviews
PHQ-9	Patient Health Questionnaire – 9-item Scale
PHQ-15	Patient Health Questionnaire – Somatic Symptoms Scale
PHQ-PD	Patient Health Questionnaire – Panic Disorder Scale
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PSS	Perceived Stress Scale
PTSD	Posttraumatic stress disorder
QIDS-SR	Quick Inventory of Depressive Symptomatology – Self-Report
RCI	Reliable change index
RCT	Randomised controlled trial
SMS	Short Message Service
SNRI	Serotonin-norepinephrine reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
STAI	State Trait Anxiety Inventory
SUDS	Subjective Units of Distress
UK	United Kingdom of Britain
uMARS	User version of the Mobile Application Rating Scale
US	United States of America
USD	United States dollars
VA	U.S. Department of Veterans' Affairs
Y-BOCS-SR	Yale-Brown Obsessive Compulsive Scale – Self-Report

Symbols

M	Mean
N	Frequency
SD	Standard deviation
S_{diff}	Standard difference
S_E	Standard error

Thesis Abstract

Digital mental health resources have been expanding over the past 20 years in line with the increasing use of the Internet. In particular, websites that offer automated therapy for mental health issues, and telehealth services provided over the Internet by mental health practitioners, have evolved into commonplace usage. The increasing ownership of Internet-enabled smartphones throughout the world is giving rise to the opportunity for addressing the growing mental health needs that are not being met by in-person mental health services. This problem has been highlighted in the recent coronavirus disease (COVID-19) pandemic. Smartphone applications (apps) are allowing automated therapy options to expand beyond the desktop or laptop computer, and large numbers of consumers are taking advantage of this. Although many thousands of mental health apps are available, the vast majority do not have evidence of their efficacy or effectiveness. Two key barriers to conducting research are the time and financial constraints associated with randomised controlled trials. This thesis focuses on the issue of evidence for mental health apps by shining a light on the current situation, and providing a possible solution to expanding the overall evidence base. It is proposed that a single-case design with multiple participants is able to feasibly and practically provide such information. The use of a single-case design in this research on five separate mental health apps successfully demonstrated how efficacious and effective these mental health apps can be, even during a rare time of heightened global distress, as posed by the COVID-19 pandemic. Outcomes of this research also provide ideas and guidance about how future studies can incorporate the participation of clinicians in private practice.

Chapter 1: General Introduction and Thesis Outline

I bet there's an app for that is a common statement made nowadays. It is reflective of the dominance that mobile Internet-enabled devices, particularly smartphones, have on our lives. While it is often lamented that such a large proportion of time is taken up with using these devices, the potential positive opportunities that massive smartphone ownership affords are exciting. Smartphone applications (apps) address a number of issues in a manner that has not previously been possible. For example, people can easily stay connected with friends and family when they are physically apart, people can be traced and contacted in the event that they have been in close proximity to a person who has coronavirus disease (COVID-19), and other health issues can be addressed in a timely fashion, such as taking a blood-sugar reading. With approximately 318,000 health apps publicly available (IQVIA, 2017) to over 5.2 billion smartphone owners worldwide (Barboutov et al., 2017), it is understandable that governments, researchers, and clinicians are looking at ways to incorporate the use of smartphones into healthcare.

As a clinical psychologist in private practice, the student researcher has attempted to incorporate the use of apps into his clinical work with clients. However, in doing so, came up against the obstacle of little evidence or potentially biased evidence for the vast majority of apps that were considered possibilities for inclusion into his clinical practice. This PhD research provided an opportunity to address this issue and to begin conversations in the literature on the many fronts facing the widespread acceptance of mental health apps.

Mental Health Service Delivery via Smartphone Apps

Apps differ from websites in an important way: apps can be downloaded on to a smartphone or tablet device and used without having to connect to the Internet, whereas access to websites require an Internet connection (Bushnell, 2020). Some apps also require an Internet connection to operate, and that is because they contain functionality that requires

access to online content. Online health information and programs (i.e., automated websites) have been operating since the late 1990s, and there is now widespread acceptance for the efficacy of many mental health online programs (Barak & Grohol, 2011). Apps, on the other hand, have developed on a different trajectory to websites. Compared to the development of websites in the 1990s to the present day, the development of apps has involved a wider collection of individuals who have the technical knowledge and resources to create apps for specific purposes on a larger scale (Uryutin, 2018), including on various fronts in the fields of medicine and health.

One such health-related use of apps is in the area of mental health. Approximately 10,000 apps fall into this category (Torous, Firth, et al., 2018). Awareness and use of these apps has greatly increased through 2020 with the rise of the COVID-19 pandemic (Basu, 2020). The resultant global upheaval has meant that in-person mental health services have been in greater demand, and current options have not been able to meet this demand (Liu et al., 2020). Although telehealth services have expanded and become widely accepted in a range of healthcare settings, including mental health, there has also been an interest in how automated options, such as apps, may be able to fill the gap in supply.

Mental health apps offer a number of potential benefits. These include: portability and accessibility (having immediate access to help any time); a convenient way of doing homework activities from face-to-face sessions; the ability to set reminders for any number of things (e.g., taking medication); providing interim help for people on waiting lists for face-to-face services; increased anonymity; improved access to treatment for people in rural and remote areas; acceptability by difficult-to-reach groups such as teenagers (Wang et al., 2018); and, reduced burden on primary healthcare services. Importantly, it also affords an opportunity for mental apps to be developed that target First Nations peoples such as *iBobbly* (https://www.blackdoginstitute.org.au/resources-support/digital-tools-apps/ibobbly/), *Kurdiji*

1.0 (https://kurdijiapp.wordpress.com/), and *Stay Strong* (https://www.menzies.edu.au/page/Research/Projects/Mental_Health_and_wellbeing/Development_of_the_Stay_Strong_iPad_App/), and culturally and linguistically diverse (CALD) communities. These populations may experience unique barriers to accessing in-person services, or even barriers to accessing other mental health apps that are aimed at majority populations that do not recognise unique cultural aspects of mental health. All of these potential benefits, along with the massive global smartphone ownership rates, provide a compelling case for having apps that can treat mental ill-health.

The primary way that consumers choose a mental health app to download is by reading the app ratings and reviews in the various app stores (Huang & Bashir, 2017). However, search functions in the app stores are limited, and algorithms that determine which apps are displayed and in what order are the product of corporate intellectual property. There is currently no easy way for a consumer to do an app store search that is based on apps with research evidence, so this leaves the app star ratings and reviews as the principal decision-making factor. The price of a mental health app also has a negative correlation with downloads, such that lower priced or free mental health apps have consistently higher ratings than higher priced apps (Huang & Bashir, 2017). Consumers also choose mental health apps based on advice on social media or by word of mouth (Rubanovich et al., 2017).

The Evidence Base for Mental Health Apps

Reviews on the extent and quality of research into the efficacy of mental health apps are starting to gather pace, as is the research itself. However, these reviews point to a general dearth of research, especially in independent research (Ameringen et al., 2017; Bry et al., 2018; Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Wang et al., 2018). That is, the vast majority of research has been undertaken by individuals and organisations involved in the development of the app being

studied, or who stand to gain financially or otherwise if the app is successful. Nevertheless, these reviews have combined findings that equate to an overall small to moderate effect size for common mental health problems, such as anxiety and depression (Lai & Jury, 2018).

To date, while the limited research on mental health apps has produced positive results relating to efficacy, a high degree of heterogeneity exists across studies. Outcomes have been measured with varying instruments (Lai & Jury, 2018), and research designs and methodologies have differed across studies – some have had placebo (Flett et al., 2019) or waitlist (Lee & Jung, 2018) control groups, while others have had no control group (Paul & Fleming, 2019). Intervention periods have varied considerably, from as little as 10 days (Howells et al., 2016) to as many as 12 weeks (Boisseau et al., 2017). Many published studies lack follow-up data to examine if changes have been maintained over time; an important factor when considering the long-term effects of using mental health apps, especially in comparison to treatment-as-usual therapies. There has also been a lack of prescribed "dose" from researchers and app developers – instructions in how to use the app can vary between little training (Flett et al., 2019) to detailed, and even daily, instructions about how to use the app (Roy et al., 2017). Often, the instruction may be something as simple as "use it however you like" when you feel anxious / depressed / stressed (Kuhn et al., 2017, p. 269). It has previously been shown that if individuals receive basic coaching in how to use a mental health app, they tend to remain more adherent to it (Mohr et al., 2011). For this reason, it would seem necessary to include user training, or at the very least a set of detailed instructions for the user, if that app is to be most effective.

There are many unanswered questions about those mental health apps that are already available for download. For example: What proportion of mental health apps use an evidence-based theoretical framework to inform its operation? What theoretical frameworks are being used? What proportion of mental health apps have been developed with input from

mental health experts, government authorities, or professional associations? And, what proportion of available mental health apps have research evidence for their efficacy? Understanding the current situation of publicly available mental health apps will help to inform the work of researchers, government authorities, and app developers.

There are also issues related to categorising an app as a "mental health app". For example, should an app designed to teach yoga, or to improve one's sleep, or even to relax a consumer by using different relaxing sounds such as cascading waterfalls and creeks, all be classed as mental health apps? Do such apps need to have research evidence to back-up their claims of improving the wellbeing of their users? These are ongoing questions being considered in the literature.

In this thesis, the research focuses on the efficacy and effectiveness of apps that claim to offer a therapeutic treatment for the mental illnesses of anxiety and depression. That is, a comprehensive treatment approach using evidence-based techniques to manage these conditions.

Increasing the Evidence Base Using Single-Case Designs

There may be several and varied reasons why research is lagging behind the uptake of mental health apps by consumers. It can be a financial impediment for app developers if they are required to put their app through a randomised controlled trial (RCT). RCTs can also take long periods of time to conduct, collate and analyse. "Time" can be a critical consideration in the health app marketplace. Over 200 new health-related apps appear in app stores every day (IQVIA, 2017), which no doubt means that some existing apps quickly become obsolete. Perhaps more importantly, new apps in development may miss the opportunity to become available before their competitors, and therefore miss out on any media "buzz" that may accompany the arrival of a new app offering a novel approach.

Although the RCT is generally considered the gold standard methodology in scientific research, including in mental health research, this design may not be the most appropriate to boost the evidence base of a wide variety of mental health apps in an acceptable time period (Clough & Casey, 2015a). In fact, it may hinder research in this space. As a result, alternative research designs have been suggested, including the single-case design (Clough & Casey, 2015a; Mehrotra & Tripathi, 2018).

Barlow et al. (2009) noted that "a series of single-case designs in similar clients in which the original experiment is directly replicated three or four times can produce robust results that may equal or surpass those produced by the experimental group/no-treatment group design" (p. 53). Other researchers in this area have arrived at similar conclusions (Horner et al., 2005; Kazdin, 2017) and identified single-case designs as being complementary to larger group designs, rather than in opposition to them (Buckley et al., 2014; Sheridan, 2014).

Single-case designs have advantages over RCTs. More information about individual participants can be captured in a single-case design, as compared to RCTs where only group means are reported. Therefore, it is possible to make informed hypotheses about the influence of peripheral issues on results (Barlow et al., 2009). Also, data can be collected at more time intervals in single-case designs, more precisely identifying when outcomes change in respect to changes in treatment (Machalicek & Horner, 2018). These designs also offer the opportunity for real-time monitoring and therefore tailoring of interventions to the responses from individuals (Bentley et al., 2019). This is particularly so when modern data-gathering techniques can be automated and digitised, thereby reducing the time burden that gathering such data may have had in the past. Similarly, data can be collated and analysed faster in single-case designs compared to RCTs (Kazdin, 2017); a crucial point in the world of mobile apps where development and listing on app stores happen at rapid speeds.

Single-case designs also have some limitations. One of the most quoted limitations is that results from a single-case study do not have broad generalizability across large populations in a way that larger RCTs may be able to do. Secondly, it can be difficult to find homogenous groups of individuals when attempting to replicate a single-case study, particularly if the individual has a complex presentation, and there may be a debate about whether one's individual presentation is adequately similar to another's. Finally, single-case research is more prone to variability affecting results. For example, if an individual who is the subject of a single-case design experiences a significant negative event (like the death of a family member) during the intervention phase, it will potentially influence the results obtained affecting the judgement of the effectiveness of that intervention. (It is also acknowledged that an intervention can be altered in response to such events in single-case research, which is an advantage over RCTs.) In contrast to this situation, if that individual was in a group research project involving hundreds of participants, the negative event would not necessarily have a substantial impact on the differences between group means.

While the use of single-case designs has been suggested in the literature as an appropriate design for researching mental health apps, no example could be located, and thus is an area of exploration in this thesis.

Mechanism of Action in Mental Health Apps

The limited research into the efficacy and effectiveness of mental health apps has provided broad support for a small to moderate effect size in their ability to treat symptoms of anxiety and/or depression (Lai & Jury, 2018). In clinical psychology, efficacy studies occur under scientific, controlled conditions where participants are screened for their suitability to improve the homogeneity of the experimental group, whereas effectiveness studies are designed to measure interventions in "real-world" clinical settings (Kazdin, 2017). Effectiveness studies are frequently in the form of an independent repetition of a previously

completed experiment (Buckley et al., 2014). While some mental health apps are based on treatment approaches with demonstrated efficacy, little is known beyond this about how and why mental health apps are efficacious. This is a complication that is partly due to the nature of group research designs, a lack of funding, and subsequent slow speeds in the research cycle of data-gathering, to write-up, to publication (Hidalgo-Mazzei & Young, 2019). Questions relating to the mechanism of action of mental health apps include: which individual techniques / theoretical frameworks within an app are working more productively in achieving desired outcomes; which type of client presentation is more responsive to treatment provided by an app; whether a mental health app can be effective as a standalone treatment and/or an adjunct treatment with concurrent face-to-face psychotherapy and/or psychotropic medication; and, if apps are more effective if there is a prescribed "dose"?

There are also further questions about the effectiveness of apps that relate to demographic variables: for example, does gender, age, marital status, chronicity, motivation for change, rapport with the app, ability with technology, and mental health literacy influence treatment outcomes? In terms of symptomatology of anxiety and depression, are apps only effective at a certain severity level of these conditions? These, and other, questions are yet to be comprehensively addressed in mental health app research.

Government Regulation and Consumer Safety

The management of risk of harm from mental health apps, particularly in relation to suicide, is an important issue. If someone experiencing suicidal ideation receives unsolicited messages or phone calls, or taps on inappropriate links contained within apps, they may be in a vulnerable position and experience a negative emotional response. Alternatively, if a person who has turned to an app because they have difficulty talking about their feelings face-to-face becomes suicidal, they may miss the last opportunity to receive helpful, even life-saving, information, before they act on suicidal thoughts. This is why it is recommended that mental

health app developers include evidence-based advice about suicidality, personal safety, and emergency contact information (Bakker et al., 2016).

There are also concerns about the privacy and confidentiality of mental health apps in the way they may use personal information without obtaining transparent and informed consent (Hendrikoff et al., 2019; Neary & Schueller, 2018; Stawarz et al., 2018; Terry & Gunter, 2018). Approximately 70% of health apps do not have a privacy policy that is available to consumers from within the app (Sunyaev et al., 2015). There are no regulations that mandate the requirement for a health app to have a privacy policy, but health and government authorities worldwide are working on ways to improve this situation. In Australia, a set of digital mental health guidelines have recently been released by the Australian Commission on Safety and Quality in Healthcare (https://www.safetyandquality.gov.au/standards/national-safety-and-quality-digital-mental-health-standards). Australia is leading the world in this area, but the guidelines remain voluntary and elsewhere regulation is something that many countries are grappling with.

Involvement of Practising Clinicians

The research in this thesis will examine the effectiveness of mental health apps as a standalone treatment, but will also note whether participants are concurrently being prescribed psychotropic medication and/or being counselled by a mental health clinician. It will therefore also examine the outcomes of mental health apps as an adjunctive treatment. It has been found that more clinicians would use mental health apps if they received training on how to incorporate them into psychotherapy (Armstrong et al., 2018; Miller et al., 2019). It is important for mental health clinicians who decide to include apps into their treatment approach, or decide to recommend the use of apps to their patients and clients, that they are confident the app is, firstly, not going to cause harm. Ideally, however, they also want to be confident that the app is potentially going to benefit their patient or client. Given that mental

health clinicians administer treatments within an evidence-based framework, they would have more confidence in using or recommending an app that had “certified” evidence for its effectiveness and was considered “safe”. However, there is no such widely agreed upon certification system in place anywhere in the world.

The methodology of single-case research designs allows for practice-based research whereby “real-world” data can be gathered by clinicians. One new way to potentially involve clinicians would be to establish an online register, similar to the way RCTs and systematic reviews / meta-analyses are currently registered, that would allow clinicians to add information about an app’s effectiveness based on a patient or client’s response to that app. In addition to clinicians, this centralised registry could potentially be accessed by consumers, researchers, students, and ethics review committees in their efforts to find the most appropriate mental health apps for their purpose. This kind of registry would also provide continually updated and evolving practice-based evidence for mental health apps. For this scheme to be most successful, clinicians would also follow a standardised protocol.

While the centralised registry is an aspirational goal, there would be significant issues that would need to be addressed in order to facilitate its establishment. The first issue centres around who would run such a facility. This could potentially be overseen by a government agency, or a collaboration between universities or health organisations. Within this issue is the question of who would adjudicate and make decisions about apps, and what qualifications and experience would they need. Would some kind of government legislation be required if it was run by a government agency? Funding for the registry would be a key aspect of its ongoing functioning. Would this come from government sources or would the registry have to be self-funded through developers paying to have their apps scrutinised and certified? These barriers are not insurmountable, but would require a collective will to move the idea of a centralised registry forward.

While the advantages of practice-based single-case research for the evaluation of mental health apps lie in the potential of increasing the real-world evidence base for the effectiveness of mental health apps, there are also some potential disadvantages. Firstly, clinicians in practice settings are traditionally time-poor (Hatfield & Ogles, 2004) and will only have limited time to add results and comments to any centralised registry. Secondly, clinicians will only have limited time to monitor the response data for patients or clients who participate in such research. Finally, it is reasonable to assume that most clinicians who participate would only be able to assess a small number of individuals simultaneously at best. However, this research approach has the potential to be highly automated, thereby allowing clinician researchers to be engaged without spending large amounts of time in administration duties associated with the research.

Aims of this Thesis

This thesis has two primary aims that seek to address several research questions. The first aim of the thesis is to describe the current state of play on the research into the efficacy and effectiveness of mental health apps, including the barriers and findings thus far. In doing so, there are seven specific research questions to be explored:

1. What proportion of apps for anxiety and/or depression claim to have research evidence for their effectiveness?
2. What proportion of that research is independent and/or replication research? That is, what proportion of research has not involved a member of the development team of the app being studied, or anyone else who is otherwise connected with the app being studied and who potentially stands to gain financially or professionally from the success of that app?
3. What proportion of apps for anxiety and/or depression have involved a mental health expert in their development?

4. What proportion of apps for anxiety and/or depression have used an evidence-based framework in their development?
5. What are the proportions of specific frameworks that have been used?
6. What proportion of apps for anxiety and/or depression have been developed in affiliation with an academic institution, medical facility, or other government-funded body?
7. What proportion of apps for anxiety and/or depression are free to download?

The second aim of this thesis is to examine if and how a single-case research design can contribute to the existing evidence base for mental health apps and offer a solution to increasing the speed and comprehensiveness of future research. In doing so, there are three specific research questions to be explored:

1. Can a range of mental health apps, employing diverse theoretical orientations, reduce subjective distress and clinically significant symptoms of anxiety and/or depression, and improve functioning, in a sample of heterogeneous participants?
2. Are there specific factors about the participants or the apps that impact on the results?
3. What factors emerge as benefits, facilitators or barriers to using single-case research, conducted by practising clinicians, to develop the evidence base for mental health apps?

The research conducted as part of this thesis uses an established single-case research design, the *multiple baseline across-individuals* design (Barlow et al., 2009), and incorporates a modern, automated approach with the potential to be scaled up to include greater numbers of participants.

The COVID-19 Pandemic

The main intervention study of this thesis was serendipitously impacted by the COVID-19 pandemic. The timing of Study 4 coincided with the “first wave” of Australian

COVID-19 cases in March 2020. Since this time, there has been a huge interest in the option for people to use digital mental health resources in the face of difficulties accessing face-to-face mental health services (Liu et al., 2020). One of the potential advantages of mental health apps is to reduce the burden on primary care services (Azarang et al., 2019), so this period in history has afforded the opportunity for mental health apps (and other digital mental health resources) to realise their potential in this regard.

With little research to back-up claims of efficacy by mental health apps, there is absolutely no research that suggests mental health apps are effective in managing mental health during a time of extreme crisis or stress, such as the one that the world has been living through since the beginning of 2020. This thesis, therefore, is uniquely placed to offer an insight into the abilities of five mental health apps to provide some level of mental health management during a time of heightened community distress.

Conceptual Overview of Thesis

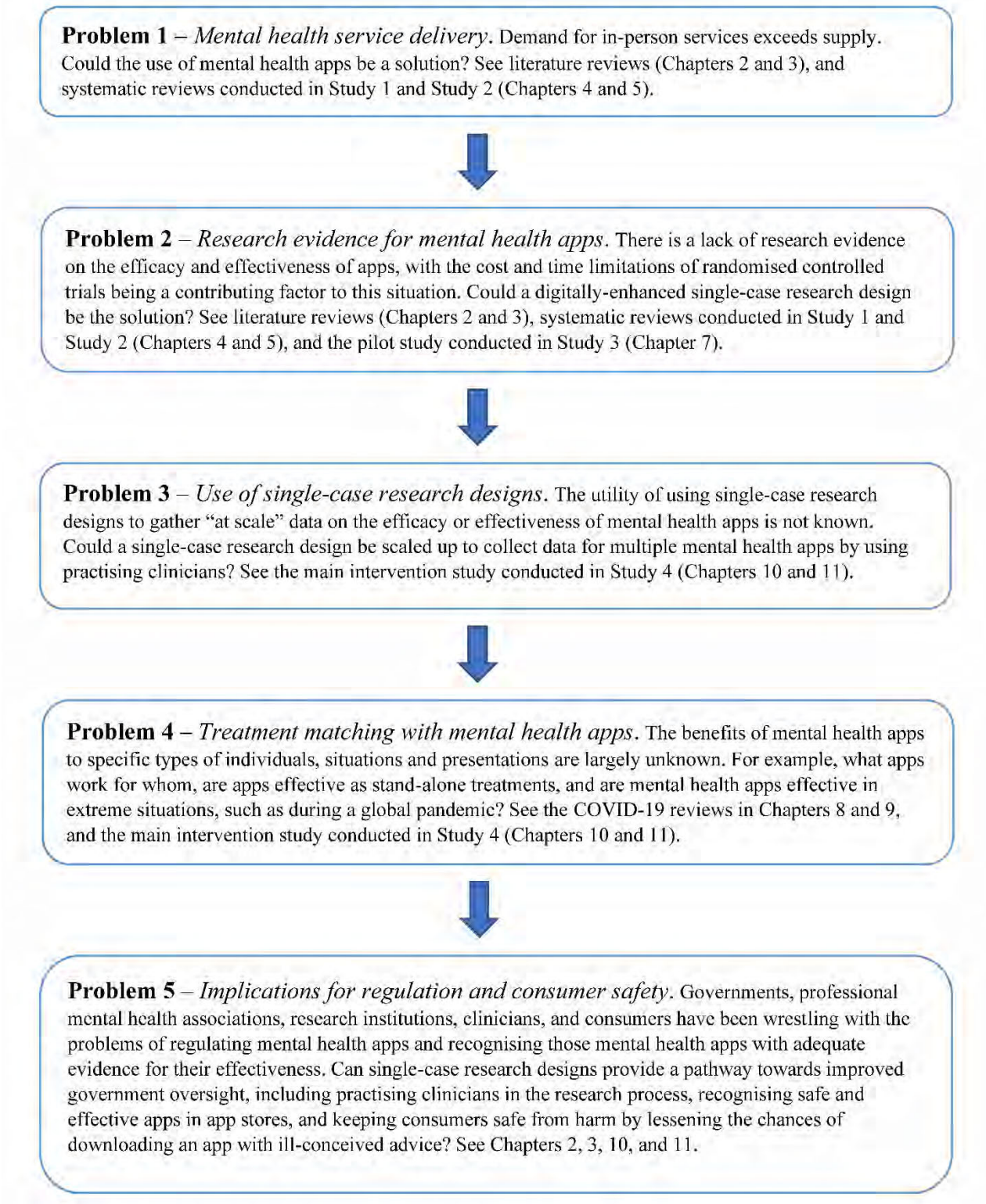
Using a positivist approach in its research philosophy, this thesis seeks to objectively examine the outcomes of symptomatology, life functioning and daily distress for individuals who use an app to manage their mental health over a 10-week period. In using a single-case research design, a large amount of information is able to be gathered and presented to readers so that they can also make subjective interpretations (e.g., with visual inspection of graphs) alongside data with clinical and statistical significance markers. This approach allows the information gathered to potentially explain why a particular mental health app was successful, or not, for a particular individual, and also allows predictions and hypotheses to be drawn for future research.

This thesis outlines the problem of a current lack of research into the efficacy and effectiveness of mental health apps. It then offers a possible solution by outlining a single-case research design methodology, and findings that demonstrate the success of this method.

It is proposed that if such a design were widely adopted for research on mental health apps, it may provide a solution for government health authorities and research institutions, and ultimately have significant benefits for consumers who stand to gain from having access to effective treatments and information in the form of a mental health app.

Figure 1.1

Conceptual Overview of Thesis



Outline of this Thesis

This thesis contributes to the fledgling evidence base for the effectiveness of mental health apps and seeks to offer a practical and feasible solution on how future research can progress. A summary of the contents of this thesis is as follows and is illustrated in Figure 1.2.

Chapter 1 provides an overview of the issue of mental health apps and difficulties with existing research.

Chapter 2 provides an introduction to digital mental health resources. It explores the place of mental health apps within this wider context of other digital resources such as websites, wearable devices, and text-based counselling services. Chapter 2 has been published in the peer-reviewed journal, *Australian and New Zealand Journal of Psychiatry*.

Chapter 3 specifically examines the safety and efficacy issues of mental health apps. Chapter 3 has been published in the peer-reviewed journal, *Professional Psychology: Research and Practice*.

Chapter 4 – Study 1 outlines the first study of this thesis, which is a systematic review of the independent research into apps for reducing anxiety and depression.

Chapter 5 – Study 2A outlines the first part of the second study of this thesis, a systematic review of the major app stores for apps addressing anxiety and depression with published evidence of efficacy. Chapter 5 – Study 2A has been published in the peer-reviewed journal, *Frontiers in Psychiatry*.

Chapter 5 – Study 2B outlines the second part of the second study; a systematic review of the major app stores for apps addressing anxiety and depression that use evidence-based frameworks. Chapter 5 – Study 2B has been published in the peer-reviewed journal, *JMIR Mental Health*.

Chapter 6 outlines the protocol and study design of Studies 3 and 4 of this thesis; a pilot study and the main intervention study. The intervention aspect of this research is carried out using multiple participants in a single-case design. Chapter 6 has been published in the peer-reviewed journal, *JMIR Research Protocols*.

Chapter 7 – Study 3 outlines the third study of this thesis, a pilot study examining the app, *SuperBetter*, using a single-case research design.

Chapter 8 outlines the use of digital mental health resources to treat trauma in Australia during the COVID-19 pandemic. Chapter 8 has been published in the peer-reviewed journal, *Psychological Trauma: Theory, Research, Practice, and Policy*.

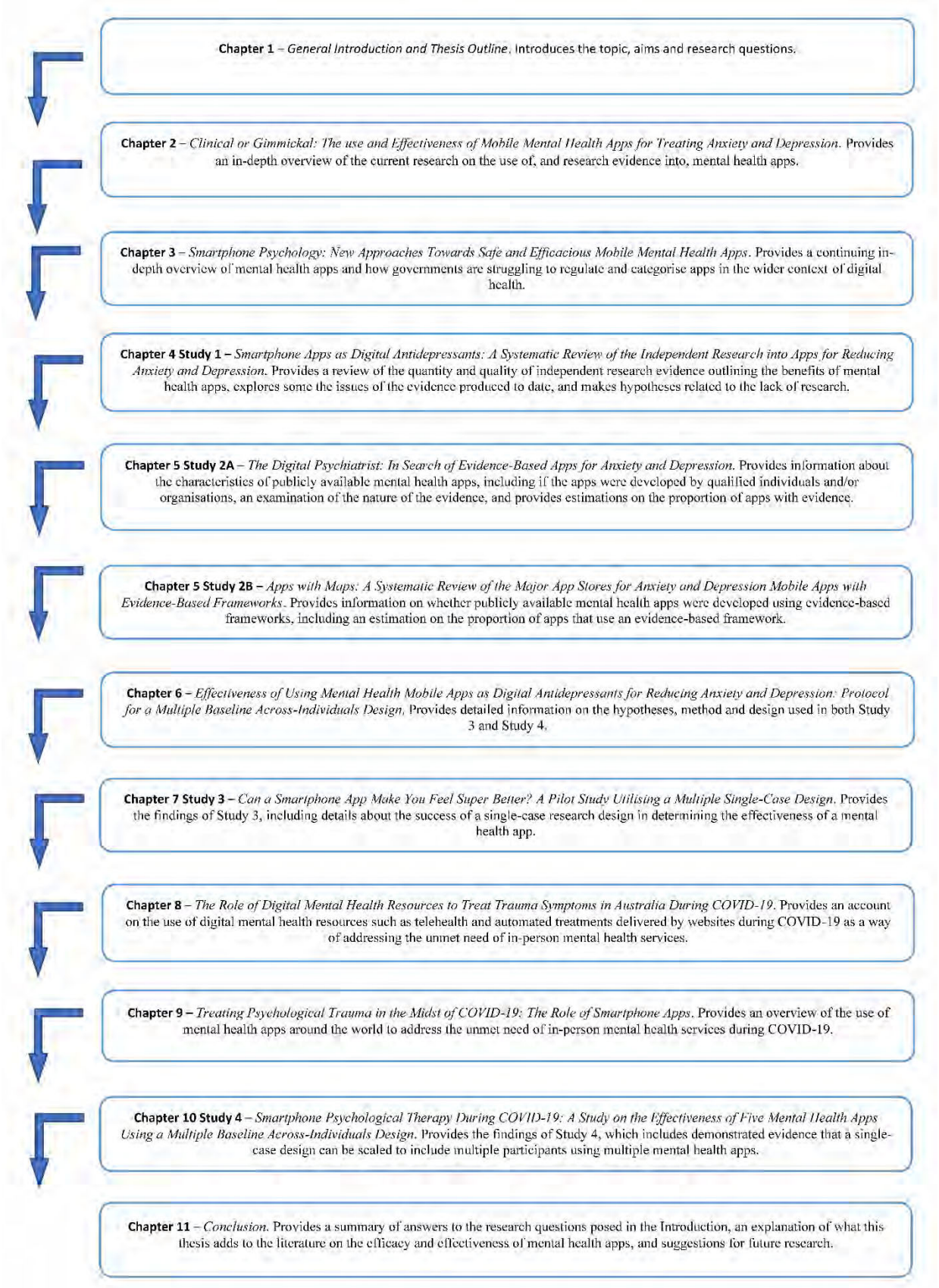
Chapter 9 outlines the use of smartphone apps to treat trauma during the COVID-19 pandemic. Chapter 9 has been published in the peer-reviewed journal, *Frontiers in Public Health*.

Chapter 10 – Study 4 outlines the main intervention study in this thesis, an examination of the effectiveness of five different mental health apps for anxiety and depression.

Chapter 11, the Conclusion, summarises the contribution that this research has made to the literature and identifies avenues of possible future research. It also makes some final statements about how consumers can be further protected from apps that purport to address symptoms of anxiety and depression, but do not have evidence to justify their claims.

Figure 1.2

Outline of Thesis



Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 2

In Chapter 1, a broad overview of the state of research into the efficacy and effectiveness of mental health apps, the state of government regulation of mental health apps, and aspects of mental health apps that are still not known was provided. This overview included the current barriers to research, and how this thesis seeks to explore and offer solutions to these barriers in the form of an alternative experimental design, specifically the single-case research design. In Chapter 2, a nuanced and expansive outline of these issues is given in the context of how it applies to working clinicians, and how the current dearth of research impacts real-world treatment of mental illness, such as anxiety and/or depression.

Chapter 2: Clinical or Gimmickal: The use and Effectiveness of Mobile Mental Health Apps for Treating Anxiety and Depression

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Clinical or gimmickal: The use and effectiveness of mobile mental health apps for treating anxiety and depression. *Australian and New Zealand Journal of Psychiatry, 54*(1), 20-28. https://doi.org/10.1177%2F0004867419876700

Abstract

Objectives: The increase in ownership of smartphones and tablet devices has seen a worldwide government push, championed by the World Health Organization, towards digital healthcare services generally. Mental health has been a strong presence in the digitization of healthcare because of the potential to solve some of the difficulties in accessing face-to-face services. This review summarises the recent history of e-mental health services and illuminates two very different paths. The first is the considerable amount of research that has proven the effectiveness of many online mental health programs for personal computers and laptops, resulting in widespread acceptance of their ability to make a contribution in an individual's recovery from anxiety and depression. The second is associated with the more recent development of apps for smartphones and tablet devices, and the contrasting paucity of research that has accompanied this burgeoning area of e-mental health. This review also outlines the current state of play for research into the effectiveness of mobile mental health apps for anxiety and depression, including issues associated with methodology, and offers

sources of practical advice for clinicians wanting more information about these new digital tools.

Conclusion: Research into the effectiveness of mental health apps is lacking and the majority have no evidence of efficacy. Clinicians need to be aware of what apps have such evidence, and should exercise caution when recommending apps to patients. Suggestions are offered on the direction of future research, including an appeal to further include clinicians in the development and efficacy testing of mental health apps.

Clinical or Gimmickal: The use and Effectiveness of Mobile Mental Health Apps for Treating Anxiety and Depression

Introduction

Smartphones and their applications (apps) are fast becoming devices that are involved in many aspects of our lives, including in the monitoring, assessment, and treatment of physical and mental health. In healthcare, apps can do as little as remind us when to take medication (Santo et al., 2016), to more complex tasks such as monitoring blood glucose levels for people with diabetes (El-Gayar et al., 2013). Society, therefore, must have confidence that apps can accurately do what they say they can do, especially with regards to healthcare. In this respect, it is important to firstly look at the development of digital mental health resources as a whole, and how these have emanated from the broader (ongoing) evolution in digital healthcare. In this context, ***mental health*** apps for smartphones and tablet devices will be examined to ascertain whether these have become a potential viable option for treating mental illness.

The Rise of Smartphones and Tablet Devices

A smartphone is a device that incorporates a mobile phone with computerised capabilities. A tablet device operates in a similar way to a smartphone, but is usually larger and may not necessarily be able to function as a traditional phone. It may, however, be able to communicate with other tablet devices and smartphones in alternative ways such as via Short Message Service (SMS) text messaging, Facetime, Skype and other such programs. Both devices are capable of connecting to the Internet, with accompanying capabilities accessible with a touchscreen. It is this combination of portability, accessibility and convenience that has led to a global explosion in mobile device usage and ownership. Worldwide, the number of mobile device users is expected to pass five billion people in 2019, and figures show

smartphone ownership will continue to rise over the next decade as their reach penetrates less socioeconomically-placed countries (Statista, 2019b).

Close to 80% of Australians (including 80% of teenagers) own a smartphone and 59% of households own a tablet device, with an estimated five million new smartphones purchased by Australian consumers in 2016 (Australian Communications and Media Authority, 2016; Deloitte, 2015). These statistics translate into over 15 million Australians owning a smartphone, and over 11 million owning a tablet device (Nielsen, 2015). More than half the Australian population check their smartphone within 15 minutes of waking (for young people, it is within five minutes) and, collectively, Australians look at their smartphones more than 440 million times a day (Deloitte, 2015), and spend 33 hours using mobile apps each month (Nielsen, 2015).

In New Zealand, approximately 64% of the population owns a smartphone, but this figure is expected to rise to 90% by the end of 2019 (Concilio, 2019). Two thirds of New Zealanders have access to three or more mobile devices, but most prefer to use their smartphone for Internet connection over any other device (Research New Zealand, 2015). Furthermore, while mobile Internet connection has steadily increased over the past five years, broadband connections (usually with devices of a less mobile nature such as personal computers and laptops) have started to decrease (Stats NZ, 2017). This is a further indication of how important mobile devices such as smartphones and tablets have become in daily life.

e-Health: Healthcare Goes Digital

e-Health is the cost-effective and secure use of information and communications technologies in support of health and health-related fields (World Health Organization, 2016). At its core is the ability for patient information to be shared instantly and, potentially, more reliably. The World Health Organization is actively encouraging countries, in the form of a ***National eHealth Strategy Toolkit***, to adopt an e-health policy in the hope that more

efficient, equitable healthcare will be provided to individuals in the future (World Health Organization, 2012).

Both the Australian and New Zealand Federal Governments have embraced e-Health. In a 2008 report, the Australian Federal Government claimed that e-health is a "key enabler and driver of improved health outcomes for all Australians" (Victorian Department of Human Services, 2008, p. 1). e-Health is a significant component of the Australian Federal Government's overall ***National Digital Health Strategy***, and this includes the recent introduction of the ***My Health Record*** system (Australian Digital Health Agency, 2017) which aims to unite the previously disparate state health systems with the desired end result being clinicians across health disciplines and specialties able to communicate more effectively, and therefore improve patient outcomes. The New Zealand Government has introduced a similar digital health strategy, ***Digital Health 2020*** (New Zealand Ministry of Health, 2017), with the aim being for the population to have a personally controlled electronic health record, similar to that of Australia. The ***New Zealand Digital Health Strategy*** describes a broad vision of how digital resources will improve outcomes across all healthcare settings, and aspires to create "a people-powered, smart health system by 2025" (https://www.health.govt.nz/our-work/ehealth/digital-health-2020; New Zealand Ministry of Health, 2019).

Mental Health's Place Within e-Health

Mental health is embedded in Digital Health Strategies, but is yet to find its place. The sensitive nature of information in a mental health setting, and the stigma that can surround mental ill-health, contributes to mental health's often uncomfortable position in the wider digital health sphere. Governments are still trying to define processes, delineations, methodologies, and regulations that provide management and guidance over the use of e-***mental*** health resources.

As a result, many mental health clinicians are confused and uncertain about e-mental health tools without such overarching guidance (Gun et al., 2011; Sinclair et al., 2013). While some mental health clinicians are becoming educated about digital options, many are not engaging (Sprenger et al., 2017). However, consumers are taking it upon themselves to embrace the new technology rather than acting on recommendations from their therapists (Schueller et al., 2018). Therefore, clinicians who ignore digital tools and resources do so at their own risk, as it is becoming increasingly important for them to have an understanding about these tools, and clinical informatics more broadly (Torous, Chan, et al., 2018).

Why do we Need e-Mental Health Services?

Anxiety and depression are the most prevalent mental illnesses worldwide (Vigo et al., 2016) with certain overlapping symptoms and causes for each. There are several evidence-based approaches to treating anxiety and depression symptoms. There are "talking" and behavioural therapies such as cognitive behavioural therapy (CBT) (Butler et al., 2006), interpersonal therapy (Barth et al., 2013), acceptance and commitment therapy (Ruiz, 2012), positive psychology interventions (Bolier et al., 2013), and others. There are other cognitive and "relaxation" therapies such as mindfulness and mindfulness meditation (Hofmann et al., 2010). There are several classes of medication, including antidepressants, such as selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs) (Cipriani et al., 2009), and others. However, it has been estimated that only approximately 30% of people with a probable clinical diagnosis of anxiety and/or depression have received appropriate treatment for these common mental health conditions (Young et al., 2001), and they continue to pose a heavy burden on primary healthcare settings, and on modern society generally (e.g., by having days off work; Gunter and Whittal, 2010).

There are various reasons why many people do not receive evidence-based help for anxiety and depression via the methods listed in the preceding paragraph. The worldwide

mental health system has limited resources in the first instance, with widespread calls for increased funding to help deal with the increase in mental illness across society (World Health Organization, 2018). People living in rural and other underserved areas may not be able to access scarce expert face-to-face services (Sakai et al., 2014). People from low socioeconomic backgrounds may not be able to afford mental health services (Jones et al., 2014). People in certain areas and from certain backgrounds (such as rural areas, cultural minorities, and low socioeconomic backgrounds) may have low mental health literacy that hinders them from getting appropriate help (Gunter & Whittal, 2010; Memon et al., 2016). Still others place stigma on help-seeking for mental health problems (Bowers et al., 2013). It is therefore important to pursue other forms of treatment for anxiety and depression, and e-mental health services offer the potential to address many of these reasons.

Clinical or Gimmickal: The Use and Effectiveness of e-Mental Health Services

e-Mental health services are defined as those digital services such as mobile apps and online web-based programs targeting mental health problems (Christensen et al., 2014). Online web-based programs for anxiety and depression have a large evidence base (for a recent review, see Gratzer et al., 2018), but the evidence for mental health apps is less clear and still emerging (discussed in detail below). e-Mental health services and resources have been identified by the Australian Federal Government as a priority area for future prevention and treatment of mental illness (Department of Health and Ageing, 2012), especially in rural and regional areas where face-to-face mental health services are either unavailable or unaffordable (Hernan et al., 2010). It is widely believed that effective integration and use of e-mental health resources could significantly reduce the social and financial burden of mental illness in Australia (Australian Digital Health Agency, 2017; Black Dog Institute, 2016). Currently, frameworks and standards surrounding e-mental health are being researched and

considered by both the Australian Commission on Quality and Safety in Health Care and the Therapeutic Goods Administration.

Caution needs to be taken when believing that digital interventions could be a panacea for the mental health sector's difficulties. Socioeconomic status is already known to be correlated with risk of mental illness (Lorant et al., 2007). There may be a "digital divide" between lower socioeconomic groups having the finances to own smartphones, where ownership is currently at 71%, compared to those in higher socioeconomic brackets where smartphone ownership is at 95% (Pew Research Center, 2019). For those from low socioeconomic backgrounds who do own smartphones, they have been found to download less apps (Tang, 2015), but there is also conflicting data on this (Rahmati et al., 2012), and to have lower mental health literacy (Kaneko & Motohashi, 2007; Knesebeck et al., 2013). While the gap in smartphone ownership is gradually closing between lower and higher socioeconomic groups (Pew Research Center, 2019), it cannot be assumed that those from lower backgrounds will benefit in the same way from the digital revolution in mental health care as those from higher backgrounds unless it is a cost-effective solution.

Previous research on web-based programs designed for use on personal computers or laptops has found that individuals have more satisfactory ratings for e-mental health resources than clinicians do (Gun et al., 2011; Waller & Gilbody, 2009). There may also be a discrepancy between younger, early career mental health clinicians and older, more experienced clinicians in their knowledge and desire to incorporate digital tools into their practice (Kim et al., 2018). Furthermore, clinicians have indicated specific barriers to using online resources, including concerns about privacy and data security (Hendrikoff et al., 2019; Neary and Schueller, 2018; Stawarz et al., 2018; Terry & Gunter, 2018); lack of time to learn and understand how various online tools work; inability to access relevant training; and concerns about client feedback (Sinclair et al., 2013). It has been recognised that the

successful implementation of digital tools into a patient's treatment is often determined by the clinician receiving recommendations from other clinicians (Hempel et al., 2018), and/or the clinician's general attitude towards such tools (Hickie et al., 2010; Wangberg et al., 2007). Thus, the clinician's perspective would appear to be vital in the successful uptake and wider dissemination of e-mental health resources (Hollis et al., 2018; Martinez-Martin and Kreitmair, 2018). Furthermore, there is a distinct lack of teaching about how to incorporate digital tools into practice in academic mental health programs, including for those training to be psychiatrists (Kim et al., 2018; Torous, Chan, et al., 2018).

Over the last two decades, e-mental health programs (particularly those with a CBT framework) have been found to be effective for both adults and children in reducing anxiety and depression symptoms (e.g., Ebert et al., 2015; Richards and Richardson, 2012). Some studies have compared the effectiveness of e-mental health programs with face-to-face therapy and have found comparable (e.g., Andrews et al., 2010; Richardson et al., 2010) and, in some cases, even more favourable, results (e.g., Merry et al., 2012). However, one of the main differences in results between e-mental health programs and face-to-face therapy is that participant drop-out rates appear to be significantly higher for computerised conditions (Andrews et al., 2010; Richardson et al., 2010), and this issue requires further research. There is now an interest in how stand-alone e-mental health programs (i.e., programs that can be downloaded once and not need ongoing Internet connection or regular therapist involvement) may further enhance the clinicians' digital toolbox.

Mobile digital technology allows users to incorporate their smartphones and tablet devices into the diagnosis and treatment of their mental health conditions in various ways. This includes using physiological and other data automatically gathered by their device (Beiwinkel et al., 2016), to using specific programs to enter data (Schwartz et al., 2016), or getting tips about how they might be able to improve their mood (Bakker et al., 2018a).

Researchers are also looking at how using incidental mobile phone data, such as phone calls made and amount of time the screen is "on", can help predict an individual's mental health diagnosis (Faurholt-Jepson et al., 2019). There may yet be further ways that smartphones can reveal important information about our mental health, and thus potentially lead to more effective treatments (Hidalgo-Mazzei & Young, 2019).

Enter the App

A mobile "app" is a software program that is either pre-loaded or downloaded from a website on to a mobile device such as a smartphone or tablet to fulfil a particular purpose. A significant advantage for apps is that for many, once it is downloaded to one's smartphone or tablet device, it can be used "offline", or without needing to connect to the Internet. There is significant progress being made in the area of online interactive digital platforms that can be accessed via websites. These services are often "app look-a-likes". That is, one can download them as an app, but the app "points" to a website that can provide a greater range of interactive tools, including the possibility of interacting with a therapist. The significant advantages for online interactive platforms are that greater security features can be applied and the amount of content and interactive potential is far greater than an offline, standalone app. These are features that are particularly attractive to clinicians who may wish to use such tools in clinical settings. Nevertheless, there is an attraction for many people who would like to use such apps without needing to connect to the Internet each time they open the app, especially in situations where connectivity to the Internet is unreliable (such as in remote communities where First Nations peoples may reside) and in cases where the cost of regular Internet connectivity is prohibitive (such as with low socioeconomic and/or minority populations that may include First Nations peoples and culturally and linguistically diverse communities). In a recent report (App Annie, 2019), global statistics on app usage revealed that there were 194 billion individual app downloads in 2018, of which China alone

accounted for nearly 50%. These downloads represented over $US100 billion in revenue for the various app stores. The average user in the US, South Korea, Japan and Australia has over 100 apps on their smartphone.

Health apps (including those related to mental health) are one of the fastest growing categories of apps (Khalaf, 2016). Such apps saw an increase of 300% between 2016 and 2018. Over two thirds of adults would be willing to use their smartphone to help manage their health (Makovsky, 2015), and about 60% of people who own a smartphone have downloaded at least one health app (Carras et al., 2014). There are approximately 318,000 health apps currently available, with more than 200 new health apps on average appearing each day (IQVIA Institute, 2017). Of the overall number of health apps, more than 10,000 relate to mental health (Torous, Firth, et al., 2018), but this can include apps that are simply digital versions of existing validated inventories, such as the Kessler Psychological Distress Scale (K10; e.g. NovoPsych), to apps that teach users to breathe more effectively (e.g. Breathe 2 Relax), and apps that simply remind the user when to undertake a specific mental health-related task, such as when to take medication (e.g. Pill Alert Pro).

Individuals find health apps in a number of ways, including through social media and word of mouth, but less so through medical providers (Schueller et al., 2018). This may be because clinicians are reluctant to make such recommendations to their patients. Of the people who download health apps, approximately one third stop using the app after a short time due to the app being “too confusing to use”, and nearly half stop because it either takes too much time to enter data or they simply lose interest (Krebs & Duncan, 2015). Finally, over half the people who have downloaded a health app value ease of use of the health app over trust of the app (Makovsky, 2016), possibly suggesting that research into a health app’s effectiveness is, alarmingly, not an important consideration for many consumers.

Increased desire for mobile apps designed for smartphones and tablet devices, as opposed to web-based programs designed for use with laptops and personal computers, has not seen an equivalent increase in the evidence for their efficacy. That is, the vast majority of previous research into the effectiveness of e-mental health resources has examined programs designed to be used on personal computers and laptops, with a paucity of research on the effectiveness of mental health apps for smartphones and tablets.

I Bet There's an App for That: Research Into Mental Health Apps

Mental health apps offer a number of potential benefits, such as portability and accessibility (having immediate access to help any time); a convenient way of doing homework activities from face-to-face sessions; ability to set reminders for any number of things, e.g. taking medication; people on waiting lists for face-to-face services may be able to get help via a mobile app; increased anonymity; improved access to treatment for people in rural and remote areas; access difficult to reach groups such as teenagers (Wang et al., 2018); and reduce the burden on primary care health services. There is, therefore, much to gain by pursuing research into the efficacy of mobile mental health apps.

Research has shown that people with mental illness increasingly own smartphones and other mobile and "wearable" devices, and are interested in using these to monitor their mental health (Carras et al., 2014; Pung et al., 2018; Torous & Powell, 2015). However, there appears to be a considerable gap between having an interest in such apps and actually using them (Torous, Wisniewski, et al., 2018). Other research has shown that people living with mental illness have varying degrees of interest in and knowledge about mental health apps (Carpenter-Song et al., 2018). Still other research has shown that those with mental illness can have distinctly negative attitudes towards an app's ability to manage the sensitive information often associated with mental health treatment (Hendrikoff et al., 2019). This all adds up to a confusing, uncertain picture about consumer interest in using mental health apps

in the real world (as opposed to research settings). But the potential advantages of having access to effective mental health apps are too great to ignore.

Currently, it appears that the main way that individuals decide which mental health app to download is by using the app ratings and reviews in the app stores (Huang & Bashir, 2017). Huang and Bashir (2017) reveal that the price of a mental health app has a significant negative correlation with downloads. Furthermore, lower priced mental health apps have significantly higher ratings than higher priced apps. Other ways that individuals decide on which mental health apps to download include advice via social media and word of mouth (Rubanovich et al., 2017).

Up until now, research on mental health apps has been poor (Alyami et al., 2017; Bakker et al., 2016; Orman & O'Dea, 2018). Consequently, there is a current push from governments, the mental health research community, and clinicians to improve this situation. For example, the National Institute for Health Research in the UK has developed the *mhabitat: Framework for the effectiveness evaluation of mobile (mental) health tools* (Betton et al., 2017) as a way of guiding app developers towards best practice in producing mental health apps, including highlighting the importance of establishing an evidence base. While this framework provides information about outcome measures to consider in measuring the efficacy of an app, it is not prescriptive in suggesting a specific research design other than recognising that randomised controlled trials (RCTs) are the gold standard, but may not be appropriate for studying the efficacy of mental health apps.

The vast majority of the published research on the efficacy of mental health apps has been carried out by individuals and organisations involved in that app's development (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). In other words, there appears to be little "independent" replication being conducted (or at least published), thus raising the issue of bias and potential conflicts of

interest. The research completed to date has found favourable outcomes for participants using the examined apps (e.g., Dillon et al., 2016; Firth et al., 2017a; Firth et al., 2017b, Miloff et al., 2015), but until increased independent replication takes place, there will continue to be questions about the effectiveness of the studied apps.

Research into the effectiveness of mental health apps is still in its infancy, and there are ongoing discussions in the scientific literature about how best to examine the efficacy of mental health apps (Pham et al., 2016). App development and the online selling of mental health apps through websites such as the Apple App Store and Google Play has increased without significant research to confirm the effectiveness of such apps. This lack of research is attributable to many factors, including limitations inherent in academic settings due to issues such as lack of funding and subsequent slow speeds in the research cycle of data-gathering, to write-up, to publication (Hidalgo-Mazzei & Young, 2019).

The limitations of RCTs in the area of mental health apps are many and varied. App development can take place quickly, and if significant funds have been invested into an app, there is a financial imperative to take it to market as soon as possible so that investors can recoup their investment. This means that the period it may take to conduct an RCT (sometimes years) is less acceptable in such commercial environments. Furthermore, in an environment where RCTs are difficult for investors to justify, replication studies become near impossible. Another limitation of RCTs may be difficulties gaining nuanced knowledge about minority populations such as First Nations peoples and culturally and linguistically diverse (CALD) communities (unless of course such populations are the focus of the RCT in the first place).

Mental health clinicians, who have been trained to follow evidence-based practice within the scientist-practitioner model, may therefore be reluctant to engage with and less willing to recommend mental health apps (Sinclair et al., 2013). Especially as many mental

health apps will not only have no evidence for their effectiveness, but also may use interventions or include content that do not qualify as accepted evidence-based techniques (Huguet et al., 2016; Kertz et al., 2017). Given that mental health consumers will often listen to recommendations by their treatment providers (Pung et al., 2018), increasing the evidence base of such apps may increase the likelihood of clinicians recommending them to patients, and therefore lead to an increase in the uptake of mental health apps (Neary & Schueller, 2018). Only when this occurs will it become more commonplace for clinicians to make a referral to a digital intervention for their patients.

Although there are some published reviews on the effectiveness of apps (e.g., Ameringen et al., 2017; Bry et al., 2018; Donker et al., 2013; Firth et al., 2017a; Firth et al, 2017b; Wang et al., 2018), the ever changing landscape of this area means it is difficult for these reviews to remain current. By the time such reviews appear in peer-reviewed journals, new apps appear in the app stores, and existing apps have often been updated or are no longer available. Nevertheless, it is important to regularly review this space as the international mental health community continues to grapple with understanding the place of mental health apps in the overall assessment and treatment of mental disorders. There are still many questions relating to the efficacy of such apps. For example, there have been studies that have shown that the level of engagement with an app predicts improvements in anxiety and depression symptoms (Bakker & Rickard, 2018). However, other research has raised the possibility that simply getting participants to engage with the novel activity of interacting with their smartphone may be enough to activate decreases in symptomatology (Torous & Firth, 2016), a phenomenon known as the ***digital placebo effect***. Furthermore, there are questions relating to the mechanisms of action of mental health apps, let alone whether they are efficacious. This includes needing to understand which individual techniques within an app are working more productively in achieving desired outcomes, and targeting whether

interventions within an app are evidence-based to begin with. With the research on mental health apps in such a fractured state, it is understandable why clinicians are ambivalent about recommending them.

Resources for Clinicians Wanting More Information About Apps

A description of the features and characteristics of individual mental health apps is outside the scope of this review, but for a summary, see Aryana et al. (2018); and Bakker et al. (2016). The following is a brief overview of some ways that clinicians can seek out information about appropriate digital mental health resources. It should be noted that the vast majority of mental health apps have been developed for individuals to use in isolation, and therefore provide little opportunity for clinicians to integrate these tools into face-to-face therapy (Lan et al., 2018).

One recent effort to provide users of mental health apps with additional information over and above app store star ratings and comments, is the development of PsyberGuide (https://psyberguide.org/), a US-based website managed by One Mind, a research organisation dedicated to brain research under the umbrella of reputable departments of the University of California and Northwestern University (for further details, see Neary & Schueller, 2018). While PsyberGuide is a trusted source of new information to guide clinicians and consumers about choosing appropriate mental health apps, it does not advance the scientific evaluation of apps. That is, it contains reviews by certified mental health professionals, and details of previous research (if any), but does not offer a new way to systematically evaluate an app's efficacy that has been independently supervised by qualified researchers. Perhaps, unsurprisingly, the ratings found on PsyberGuide generally do not correlate with the ratings found in the various app stores (Neary & Schueller, 2018). The information on PsyberGuide is nonetheless useful and trustworthy, as it is managed by a not-for-profit organisation, and offers a new, additional source of information.

The American Psychiatric Association has formulated an app rating framework for clinicians to use when prescribing a mental health app to a patient (Torous, Firth, et al., 2018). It asks clinicians to evaluate four key areas when judging the appropriateness of an app in assessing for: 1. Potential risk and harm, including privacy and data management; 2. Research evidence for efficacy; 3. Ease of use; and 4. Ability to share data with clinicians if necessary. The system does not include previous ratings by other individuals, and it is therefore up to the clinician to "rate" the app from scratch. While this may have future benefits for researchers, it is difficult to see clinicians in primary care or private practice settings regularly going through this time-consuming process. Similar to PsyberGuide, it does not provide a measure of efficacy even though it considers the issue of efficacy to be important.

The Australian Federal Government recently launched the ***Head to Health*** website (https://headtohealth.gov.au/) that provides information about digital mental health options to both consumers and clinicians alike. Similarly, ***reachout.com*** (https://au.reachout.com/tools-and-apps), an Australian not-for-profit organisation, offers advice and ratings from both mental health professionals and consumers. The Australian National University also funded the development of the ***beacon*** website (https://beacon.anu.edu.au/), where digital mental health resources are rated by experts, and research references are provided.

In New Zealand, the Health Navigator website (https://www.healthnavigator.org.nz/) was initially set up to assist general practitioners find information on a variety of health topics, including mental health. It can, however, be used by clinicians and consumers alike to find information on mental health apps, websites and digital mental health programs. The website is supported by the Royal New Zealand College of General Practitioners and many of the District Health Boards. It includes links to different resources that the various District Health Boards may have created and makes these resources accessible to the wider

population. The Health Navigator website also houses the New Zealand Health App Library that provides summary information on mental health apps, as well as health apps generally, including reviews by mental health specialists, general practitioners, consumers, and technological/app development experts. While the Health Navigator, beacon, reachout, and Head to Health websites acknowledge many of the difficulties involved in recommending mental health apps, and each uses a transparent process in deciding which apps to recommend, again they do not provide scientific evidence for the effectiveness of any app, but do recognise other research that may have been completed on the app.

The above sources of additional information on mental health apps are welcome tools for clinicians and consumers, but this type of information is not available at the point of download in the app stores. It would be preferable to have such additional information at this point, as this window may represent the moment of strongest motivation and likelihood to download an app for an individual (Huang & Bashir, 2017). It would also be easier for clinicians and consumers alike if the app stores had more involvement in the process of "certifying" health apps, so that apps with research evidence to back-up their claims of effectiveness stand out from apps that lack such evidence.

Conclusion

There is wide scope for future research in the area of mental health apps. For example, there is an opportunity for mental apps to be developed that target First Nations peoples and culturally and linguistically diverse (CALD) communities. These populations may experience unique barriers to accessing in-person mental health services, or even barriers to accessing other mental health apps that are aimed at majority populations that do not recognise unique cultural aspects of mental health. Clinicians have to be more involved in the app development process, as there is currently a lack of mental health apps with expert input (Alyami et al., 2017; Shen et al., 2015). The involvement of clinicians would improve and increase the

number of apps developed for use in conjunction with face-to-face therapy, an area of huge potential. Clinicians' participation in app research may also increase the number of independent and replication studies, boosting research and assisting mental health apps to gain greater legitimacy, while producing an evidence base from research that is proportionately less heterogeneous than it currently is.

The area of mental health apps, and health apps more broadly, is developing quickly. Technological innovations mean that it has been difficult for research to keep up with development. Studies involving large sample RCTs take time to organise and implement, to manage the large numbers of participants and amounts of incoming data, and to synthesise that data into clinically meaningful results (Pham et al., 2016), and although the RCT is considered the gold standard for demonstrating efficacy in scientific research, there are considerable numbers of researchers who believe that RCTs are not the most appropriate experimental designs for health apps (Betton et al., 2017). If research in the area of mental health apps is to keep up with development, and to include the involvement of clinicians, perhaps newer, more novel ways of research / evaluation are required (Hollis et al., 2018; Menon et al., 2017; Neary & Schueller, 2018; Nicholas et al., 2016; Pham et al., 2016).

While this review has noted the potential of mental health apps to facilitate positive outcomes in conditions such as anxiety and depression, it still seems that more research is needed before these tools can confidently be prescribed by clinicians on a widespread basis.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 3

In Chapter 2, a detailed introduction was given about the current state of play for research into the efficacy and effectiveness of mental health apps for anxiety and depression. However, this is not the only concern related to mental health apps. Other concerns include the need to keep consumers safe if they are using mental health apps to address mental health problems, and this involves a role for government regulators and mental health professional associations. In much the same way as access to medication and devices that purport to manage health conditions are overseen by government regulators, there is a role for governments to provide similar oversight to digital mental health resources. Chapter 3 provides a review of the current situation of government regulation in the area of mental health apps, issues relating to the safety of consumers, and how future research might provide information on these and other issues by embracing alternative experimental research designs to the randomised controlled trial, such as the single-case design.

Chapter 3: Smartphone Psychology: New Approaches Towards Safe and Efficacious Mobile Mental health apps

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

Abstract

Practicing psychologists are being faced with the reality that mobile mental health apps for smartphones and tablet devices are increasing in popularity. This growth area within e-mental health has been well documented in *Professional Psychology: Research and Practice*. This article provides an update on the issues of safety and efficacy in mental health app development, two of the biggest concerns that practicing psychologists have about these new digital tools. Governments and medical authorities are wrestling with how to regulate the health app market to avoid harm to users. At the same time, a lack of research into the efficacy and effectiveness of most mental health apps in the various app stores leaves clinicians and consumers with uncertainty. The vast majority of the limited research to date has been completed by those involved in an app's development. Further independent research and replication is required to demonstrate legitimacy and increase the acceptance of mental health apps as valid sources of therapy. Complicating this issue is disparity about an acceptable methodology for examining the effectiveness of a mental health app. This article proposes a new approach to incorporate multiple baseline single-case designs to increase the amount of evidence and to guide larger scale randomized controlled trials, something that could, and should, include practicing psychologists. This novel approach also proposes that

mental health apps undergo a new “certification” process with the participation of app store marketplaces.

Smartphone Psychology: New Approaches Towards Safe and Efficacious Mobile Mental health apps

The diagnosis and treatment of mental disorders is promoted by the World Health Organization (WHO; 2013) as an activity that requires a highly trained workforce regulated by governments with a prioritized mental health agenda and action plan. The use of mobile mental health applications (apps) is a contentious issue in this context because a large proportion of such apps have been developed by individuals who are not mental health experts (Alyami et al., 2017; Shen et al., 2015). Following an increase in the development of mental health apps – there are over 10,000 available worldwide (Torous, Firth, et al., 2018) – mostly outside of any regulation, there is broad recognition by multiple stakeholders such as governments, research institutions, clinicians, and consumers that regulatory intervention is needed. This is particularly true for apps that offer a comprehensive therapeutic treatment and/or diagnosis for a mental disorder, as opposed to apps that may offer singular or novelty interventions that may or may not be useful in assisting to manage one's mental health, but which may get classified within a broad definition of what constitutes a "mental health" app. Previous reviews by Firth, Torous, Nicholas, Carney, Rosenbaum, et al. (2017), Firth, Torous, Nicholas, Carney, Pratap, et al. (2017), and Lui et al. (2017) found only limited examples of mental health apps with research evidence, therefore it is important that this research gap is addressed using suitable methodologies.

This report builds on previous published material in *Professional Psychology: Research and Practice*; specifically, articles by Clough and Casey (2015a); Clough and Casey (2015b); Jones and Moffitt (2016); Lui et al. (2017); and more recently, Armstrong et al. (2018); Broussard and Teng (2019); and Miller et al. (2019). For a summary, see Table 3.1. Together, these articles represent different strands of an ongoing scientific conversation

about how this emerging area may best assist consumers to get the most effective and efficacious treatment from an app. Following a brief overview of the benefits and shortcomings of apps, this paper critically reviews the issues of safety, regulation, efficacy and effectiveness of mental health apps. It then offers a novel proposal for increasing the evidence base of effectiveness and certifying mental health apps with the involvement of practicing clinicians and the major app stores.

Table 3.1

Recently Published Articles on Mental Health Apps in Professional Psychology: Research and Practice

Article Authors, *Type of Article*	Article Focus	Other Characteristics of the Article	How does the present article add to this?
Clough and Casey (2015a), *Review*	Alternative research designs to RCTs for examining the efficacy of mental health apps, with examples using an app developed by the authors.	Does not advocate one approach over another, but acknowledges that different situations may require different approaches.	Covers recent developments in the areas of safety and regulation, advocates for single-case designs as being the easiest to implement for practicing clinicians, and proposes a certification framework for more accurate identification of efficacious mental health apps in the app stores.
Clough and Casey (2015b), *Review*	Broad overview of the area of mobile health, including history, strengths and limitations, and future directions. The article makes some limited references to mental health within the wider context of mobile health generally.	Designed to give a broad overview of mobile health technology, of most benefit to those with limited or no knowledge about this area.	More focused on regulation, safety and efficacy issues which adds to what is mentioned on these topics in the Clough and Casey (2015b) article. The present article is more specifically focused on mental health.
Jones and Moffitt (2016), *Review*	Aims to inform and encourage developers of health and mental health apps to follow ethical guidelines that are put forward for consideration.	It also addresses issues related to regulation, safety, and law within the context of ethical considerations.	Aimed primarily at clinicians, but also researchers. The present article examines safety, regulation, methodology, and efficacy.
Lui et al. (2017), *Systematic Review*	Examines the empirical status of mental health apps, specific features of apps that may influence efficacy, and what have	While this article does look at efficacy and effectiveness, it focuses on individual features of apps that influence these issues.	Examines how methodology could change so that more evidence is produced in support (or otherwise) of the effectiveness and efficacy of mental health apps. The present article is specific in its suggestions for

	been some of the methodological shortcomings of previous research.		future methodology and certification, and also provides a more recent update on efforts to regulate health apps which have advanced since the Lui et al. (2017) article was written.
Armstrong et al. (2018), *Research Study*	Focuses on the evaluation of a program to teach mental health providers how to use specific apps.	This article showed that mental health providers will utilize mental health mobile apps in their practice if given specific training in how to do so.	Covers recent developments in the areas of safety and regulation, advocates for single-case designs as being the easiest to implement for practicing clinicians, and proposes a certification framework for more accurate identification of mental health apps in the app stores.
Broussard and Teng (2019), *Theoretical Review*	Focuses on different learning theories that may inform the theoretical underpinnings of mental health apps in conjunction with other factors, such as evidence-based practices and techniques to foster engagement.	This article examines how mental health experts can team up with app developers in order to produce more engaging, more effective, more efficacious mobile mental health solutions. Focuses on Experiential Learning Theory and Blooms Taxonomy.	Covers recent developments in the areas of safety and regulation, advocates for single-case designs as being the easiest to implement for practicing clinicians, and proposes a certification framework for more accurate identification of mental health apps in the app stores.
Miller et al. (2019), *Research Study*	Perceptions and use of apps by providers of mental health services to the Veteran's Affairs community.	This study also examined the barriers to using apps for mental health providers to the Veteran's Affairs community.	Covers recent developments in the areas of safety and regulation, advocates for single-case designs as being the easiest to implement for practicing clinicians, and proposes a certification framework for more accurate identification of mental health apps in the app stores.

Benefits and Shortcomings of Mental Health Apps

Mental health apps offer potential value in a number of different ways. They are compact, easily transportable, allow access to instant assistance, and can provide anonymity for people who do not wish to visit a mental health professional in person. Apps may also allow interactive homework activities, and be set up to send information digitally to a therapist. There are apps that set reminders (e.g., take medication, do meditation etc.) and allow individuals to participate in self-help tasks or psychoeducation if on a waiting list for face-to-face services. People using apps in rural areas can potentially access mental health services previously unavailable, and people from lower socioeconomic backgrounds may be able to access affordable treatment via an app when that treatment may be cost-prohibitive to receive in private face-to-face settings. There may be increased access for other groups such as teenagers with more effective early intervention (Wang et al., 2018), and people with less severe mental illness could access an effective app that places less demand on primary care services. Importantly, it also affords an opportunity for mental apps to be developed that target First Nations peoples, and culturally and linguistically diverse (CALD) communities. These populations are over-represented in measures of health and wellbeing, and the development of specific mental health apps could assist in overcoming cultural barriers. In relation to First Nations peoples, the Close the Gap priorities of the Australian Federal Government (https://www.closingthegap.gov.au/) have signalled improved mental health as a key health outcome, including a reduction in the levels of suicide. There have already been apps developed that address these issues, including *iBobbly* (https://www.blackdoginstitute.org.au/resources-support/digital-tools-apps/ibobbly/), *Kurdiji 1.0* (https://kurdijiapp.wordpress.com/), and *Stay Strong*

(https://www.menzies.edu.au/page/Research/Projects/Mental_Health_and_wellbeing/Development_of_the_Stay_Strong_iPad_App/).

The shortcomings of mental health apps include suboptimal use of the available technological features of smartphones (Frank et al., 2018; Hendrikoff et al., 2019; Shah et al., 2018), but also a lack of technology literacy on the part of some users, particularly the elderly, who may struggle to operate apps, and smartphones generally (Mohadis & Ali, 2014). Furthermore, the development of mental health apps has often occurred without expert input, government funding, or the involvement of academic institutions (Alyami et al., 2017; Shen et al., 2015), or adequate training for consumers and clinicians (Armstrong et al., 2018), and without ethical guidelines (Broussard & Teng, 2019). There are also concerns about privacy and confidentiality (Hendrikoff et al., 2019; Neary & Schueller, 2018; Stawarz et al., 2018; Terry & Gunter, 2018), especially in terms of an individual's personal information being used without their permission or knowledge which may pose a threat to anonymity. Approximately 70% of health apps do not have a privacy policy that is available to users from within the app (Sunyaev et al., 2015).

The management of suicide risk is an important and sensitive issue. The use of data sharing is of particular concern in such situations as someone experiencing suicidal ideation who may receive inappropriate unsolicited text messages or phone calls. If an individual is suicidal and is targeted with inappropriate online advertising as a result of their personal information being sold by app developers to marketing companies, they may be in a heightened vulnerable position that may make them more susceptible to content they perceive as negative. Some app users with suicidal ideation may find it difficult talking about their feelings with a mental health clinician face-to-face. Using an app may be the last chance for some individuals to receive helpful information before they act on suicidal thoughts. This is one reason why Bakker et al. (2016), in their review and recommendations for mental health

app developers, consider it important for mental health apps to include advice about suicidality, personal safety, and emergency contact information.

Regulation of Mental Health Apps: Assessing the Risk of Harm

The availability of mental health apps in the various app stores has, up until now, occurred with little government regulation. This remains a challenge for authorities and app developers, given the vast numbers of health apps that are available (Jones & Moffitt, 2016; Wang et al., 2018). Complicating the issue is that many apps may fall under a broad definition of a "mental health" app, but may actually be best described as "self-help", or offering a single or novelty intervention. A valid question arises at this point about what type of app needs to be regulated, and what is the degree of harm that is currently being done by mental health apps that are not? There is currently little in the way of legal evidence that suggests that apps falling under the banner of mental health are doing harm, but the growth in use of these tools suggests that it is important to monitor this space (Armontrout et al., 2018). From a clinician's perspective, there are specific situations when regulation of apps would be more salient. For example, when recommending that clients use an app between sessions or when treatment has finished, the clinician has to be confident that the app is not going to cause harm, and that it ideally is going to benefit the client. Clinicians would have more confidence about recommending an app if they knew it had "certified" evidence for its effectiveness and was considered "safe". Furthermore, there may be legal and ethical ramifications in the event that a client does come to harm as a result of a clinician recommending an app that proved to offer non-evidence-based information (Armontrout et al., 2016).

In the U.S., the Food and Drug Administration (FDA) intends to ban "mobile apps that are medical devices and whose functionality could pose a risk to a patient's safety" (as quoted in Armontrout et al., 2018, p.207). That is, the FDA will regulate on the basis of

individual safety, rather than focusing on efficacy (https://www.fda.gov/medical-devices/digital-health/mobile-medical-applications). Harm, and its opposite construct "safety", have been defined in many different ways and within many different contexts, and it is difficult to find a definition that is both succinct and encompassing (Emanuel et al., 2008). For the purposes of stating a definition for this article, harm refers to both physical and psychological injury that may occur to an individual (Merriam-Webster online dictionary, https://www.merriam-webster.com/dictionary/harm) as a result of using an app that makes claims about improving mental health. However, in regulating mental health apps, there are questions for authorities in examining what constitutes harm when making decisions about whether to ban a mental health app or not in determining its status as a medical device. If it is not considered a medical device, then the issue of whether or not it could cause harm is not considered by the FDA. Therefore, the most important definitional consideration initially is the definition of a mental health app as a medical device.

New efforts are also being made to regulate apps and other e-mental health tools by government bodies in other parts of the world, such as Australia (https://www.safetyandquality.gov.au/our-work/safety-in-e-health/certification-framework-for-digital-mental-health-services/), New Zealand (https://www.health.govt.nz/our-work/ehealth), Canada (https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/activities/announcements/notice-digital-health-technologies.html), and the United Kingdom (https://www.gov.uk/government/publications/health-app-assessment-criteria/criteria-for-health-app-assessment). Thus far, all are focused on the need to eradicate potentially unsafe health apps that may place an individual at risk of harm, and all are dealing with the intricacies about what constitutes harm, and what factors define a mental health app as a "medical device". Equally important, however, is that apps have demonstrated efficacy.

Measuring the Efficacy and Effectiveness of a Mental Health App: Is There an App for That?

In clinical psychology, efficacy studies occur under scientific, controlled conditions where participants are screened for their suitability to improve the homogeneity of the experimental group, whereas effectiveness studies are designed to measure interventions in "real-world" clinical settings (Kazdin, 2017). Effectiveness studies are frequently in the form of an independent repetition of a previously completed experiment (Buckley et al., 2014). Due to their growing popularity amongst consumers, the evaluation of mental health apps in both domains is important, but has lagged behind app development. If the effectiveness of mental health apps can be proven, eventually by independent replication, it may have an impact on theory insofar as being able to provide widely accepted evidence that specific theoretical frameworks can be transferable to mobile devices such as smartphones and tablets, in much the same way it is now widely accepted that such frameworks can be transferable to other digital formats including videoconference psychotherapy, virtual reality for anxiety, and web-based desktop and laptop computer programs (Mohr, Burns, et al., 2013).

Currently there is no standardized method for assessing the efficacy or effectiveness of mental health apps. This issue has been raised previously by researchers (Clough & Casey, 2015b; Lui et al., 2017), but there is little evidence in peer-reviewed journals that much has changed. For example, a recent report considered a total of 17 assessment frameworks, best practice principles, and quality assurance guidelines in relation to the development of health apps from different countries, and indicated that authorities worldwide are grappling with this issue (Nielsen & Rimpilainen, 2018). In research settings, questionnaires such as the Mobile Application Rating Scale (MARS; Stoyanov et al., 2015) have been developed as a means to rate the various features of apps, but the MARS does not focus on efficacy. Instead, it asks 29

questions relating to various aspects of the app, and contains only one question that asks the completer to consider whether the app has any published evidence for its efficacy or effectiveness. Therefore, it does not provide a measure of efficacy or effectiveness.

The most widely accessible app evaluations are consumer ratings and reviews in the app stores. This information is what the majority of consumers use to help them choose a mental health app (Huang & Bashir, 2017). However, these ratings and reviews are poorly monitored by the app stores and there have been reports of fake reviews (Xie & Zhu, 2015). Ratings, therefore, are often likely to have questionable validity and reliability in terms of offering an assessment of effectiveness and efficacy.

In response to this, efforts have been made by mental health experts to address the need for guidance in assessing apps. A collaboration between Canadian health research facilities, app developers, clinicians, and consumers resulted in a set of guiding principles for a mental health app assessment framework (Zelmer et al., 2018). This framework uniquely identifies gender responsiveness, cultural appropriateness, and user inclusion at all levels as crucial elements. Although efficacy was considered the number one criterion, there was no advice about how to study and assess it. The American Psychiatric Association also recently introduced an app rating framework designed to guide clinicians in their recommendations to patients and clients (Torous, Firth, et al., 2018). This involves clinicians considering each app in the areas of: potential risk of harm; data security; evidence of effectiveness; usability; and level of clinician interaction. Again, though, there is no guidance about a suitable methodology for examining efficacy or effectiveness.

There are now many reputable websites worldwide that provide advice about mental health apps and often have “expert” and “consumer” reviews, with information on published evidence if it exists. These include: ***PsyberGuide*** (https://psyberguide.org/); ***Head To Health*** (https://headtohealth.gov.au/); ***reachout.com*** (https://au.reachout.com/tools-and-apps);

beacon (https://beacon.anu.edu.au/); and ***Health Navigator*** (https://www.healthnavigator.org.nz/). While these websites provide valuable information not contained in app store reviews, very few rate an app's efficacy, and none recommend specific methods of assessing an app's efficacy.

Difficulties With Ongoing Research

As mentioned above, there are many apps that are categorized as a "mental health" app, but such apps may only offer singular or novelty interventions that do not qualify as comprehensive therapeutic treatments and/or diagnostic instruments. For example, *Breathe2Relax* is an app with a simple function: to help the user control their breathing by offering visual and auditory cues. It is possible that this would be categorized as a mental health app, but it does not necessarily need to have research done to determine a level of efficacy or effectiveness. It does not claim to offer a diagnosis or comprehensive therapeutic treatment for any mental disorder. These are important issues because of the large number of apps that fall into a similar category offering simple "interventions" for a narrow purpose. We would argue that research needs to focus on apps that claim to offer a comprehensive therapeutic treatment and/or diagnosis for a mental disorder.

To date, while the limited research on mental health apps has produced some positive results relating to efficacy, a high degree of heterogeneity exists across studies. Outcomes have been measured with varying instruments (Lai & Jury, 2018), and research designs and methodologies have differed across studies - some have had placebo (Flett et al., 2019) or waitlist (Lee & Jung, 2018) control groups, while others have had no control group (Paul & Fleming, 2019). Intervention periods have varied considerably, from as little as 10 days (Howells et al., 2016) to as many as 12 weeks (Boisseau et al., 2017). There has been a lack of prescribed "dose" from researchers and app developers – instructions in how to use the app can vary between little training (Flett et al., 2019), to detailed, and even daily, instructions

about how to use the app (Roy et al., 2017). Often, the instruction is simply "use it however you like" when you feel depressed/anxious/stressed (Kuhn et al., 2017, p. 269). It has previously been shown that if individuals receive basic coaching in how to use a mental health app, they tend to remain more adherent to it (Mohr et al., 2011). For this reason, it would seem necessary to include user training for that app to be most effective. Similarly, if clinicians received training in how to incorporate apps into psychotherapy, more would (Armstrong et al., 2018; Miller et al., 2019). Many published studies lack follow-up data to examine if changes have been maintained over time, an important factor when considering the long-term effects of using mental health apps, especially in comparison to treatment-as-usual therapies. Finally, although a number of reviews have been published (e.g., Donker et al., 2013; Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Mehrotra & Tripathi, 2018; Menon, Rajan, & Sarkar, 2017) with combined findings that equate to an overall small to moderate effect size for common mental health conditions such as anxiety and depression (Lai & Jury, 2018), overall the research remains sparse.

Two systematic reviews on apps for treating anxiety and depression (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017) located a total of 20 apps with published research demonstrating significant efficacy, but none of these studies represented independent research. That is, all of the located research was carried out by individuals who were either involved in the development of the app being studied, stood to gain financially from the app, or otherwise had another association with the app. While much of the limited previous research has been to a satisfactory scientific standard, if claims about favorable treatment outcomes are to be accepted, the level of independent research must increase.

A particular challenge for mental health app research is that apps are regularly updated, and it cannot be assumed that research results on one version of an app will yield the same results on updated versions (Torous, Levin, Ahern, & Oser, 2017; Wang et al., 2018). Therefore, there is a need to develop methods of being able to perform similar studies across different versions of an app. One solution to this problem has been proposed by Mohr, Cheung, et al. (2013) using the method known as Continuous Evaluation of Evolving Behavioral Intervention Technologies (CEEBIT) which can allow updated versions of an app to be evaluated separately from older versions. CEEBIT also allows different, newer apps to be tested against other older apps by comparing information related to positive behavior change in response to using the app. CEEBIT is based on a process of continued testing of similar apps that does not necessarily have to end as newer apps or updated versions continue to become available. The focus of this system is the elimination of apps on the basis of inferiority, rather than identifying apps that have reached a certain level of effectiveness.

It is difficult to see any strong pattern emerging in the literature for a research methodology on the efficacy of mental health apps. More randomized controlled trials (RCTs) have been widely called for, but one of the central obstacles to this is the time it takes to conduct such trials, often years. The popularization of mental health apps has taken place in the relatively recent past, so it is understandable that few RCTs have been completed. However, it is the extent to which apps have been embraced, including in the use of mental health apps, that has led to the realization that ongoing research is needed to confirm that the use of apps has proven potential to assist in managing mental health (Lui et al., 2017). In the world of app development, things happen quickly because improvements in technology mean that new features become available that supersede the abilities of existing apps. Therefore, future research needs to come up with alternative methodologies to meet this fluid market and

technology-driven demand for increasing evidence of effectiveness and efficacy in a timely manner (Clough & Casey, 2015b).

Future Research: A New Approach Towards Single-Case Designs and Clinician Researchers

An alternative to RCTs for testing the efficacy of mental health apps is the single-case design (Clough & Casey, 2015a; Mehrotra & Tripathi, 2018). Barlow et al. (2009) noted that "a series of single-case designs in similar clients in which the original experiment is directly replicated three or four times can produce robust results that may equal or surpass those produced by the experimental group/no-treatment control group design" (p. 53). Other researchers in this area have arrived at similar conclusions (Horner et al., 2005; Kazdin, 2017) and have identified single-case designs as being complimentary to larger group designs rather than in opposition to them (Buckley et al., 2014; Sheridan, 2014).

Single-case designs have advantages over RCTs. More information about individual participants can be captured in a single-case design, as compared to RCTs where only group means are reported. Therefore, it is possible to make more informed hypotheses about how peripheral issues may have influenced results (Barlow et al., 2009). Also, data can be collected at more time intervals in singe-case designs, more precisely identifying when outcomes change in respect to changes in treatment (Machalicek & Horner, 2018). These designs also offer the opportunity for real-time monitoring, and therefore tailoring to the responses from individuals (Bentley et al., 2019). Finally, data can be collated and analyzed faster in single-case designs compared to larger RCTs (Kazdin, 2017); a crucial point in the world of mobile apps, where development and listing on app stores happens at rapid speeds.

The methodology of single-case designs allows practice-based research whereby "real world" data can be gathered by practicing clinicians. One new way to potentially involve clinicians would be to establish an online register, similar to the way that RCTs and

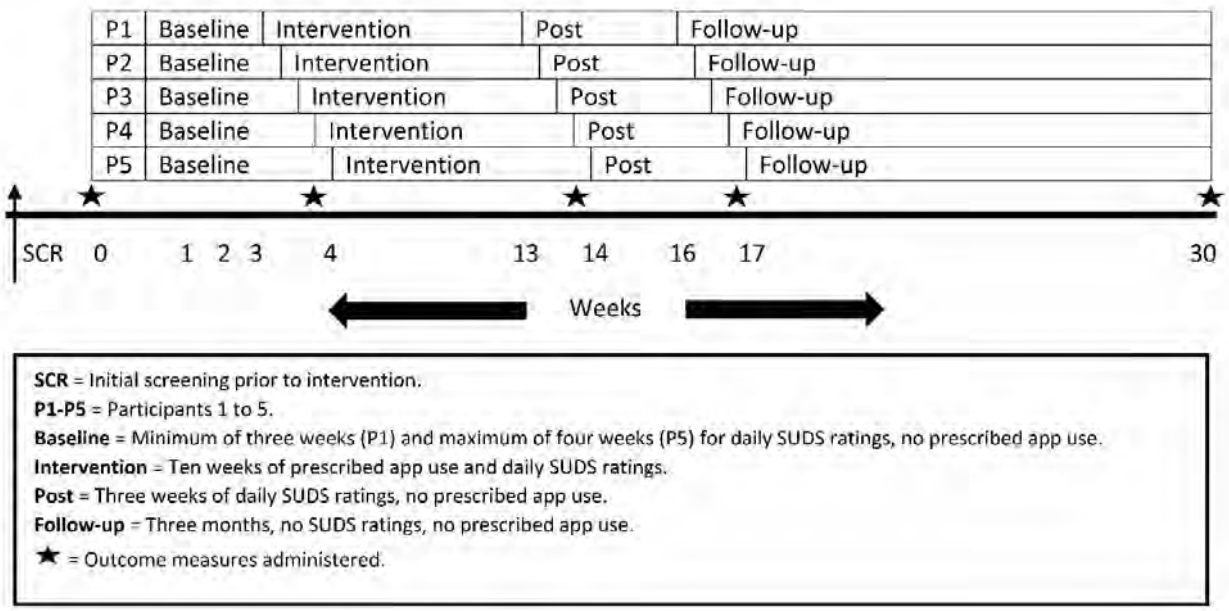
systematic reviews/meta-analyses are currently registered, that would allow clinicians to add information about an app's effectiveness based on a client's response to that app using a standardized methodology. In addition to clinicians, this centralized registry could potentially be accessed by consumers, researchers, students, and ethics review committees in their efforts to find the most appropriate mental health apps for their purpose. This kind of registry would also provide continually updated and evolving practice-based evidence for mental health apps. For this scheme to perhaps be most successful, clinicians may have to follow a standardized protocol.

As a practical exercise for clinicians (and researchers), the following is a simplified demonstration of how such a protocol might look, and how research on a mental health app may progress using the methodology of a multiple baseline across-subjects design (Barlow et al., 2009) in a practice setting. Five individuals are asked to use a specific app. A baseline period is established to confirm a series of stable readings across time (and this allows each individual to act as their own control). The app is then introduced to the clients, with instructions on how to use it. For example, "Use the app for at least ten minutes per day, five days per week, for ten weeks" (based on the idea of an equivalent 50-minute face-to-face session with a therapist once per week). There would then be a post-intervention period of equivalent time to the initial baseline period, with a follow-up period of three months. Individuals could provide regular data on their emotional state by a Subjective Units of Distress (SUDS) scale that is automatically texted to their smartphone each day, asking them to reply with a number out of ten that corresponds to their emotional state at that time, also known as ecological momentary assessment (Shiffman et al., 2008). Several online survey and statistical platforms are able to be programmed so that text messages and questionnaires are automatically sent, and returning data are automatically assigned to individual participants. In this way, a clinician is able to track an individual's progress, and eventually

make a judgement on the effectiveness of the app based on its demonstrated clinical significance (Jacobson et al., 1999; Jacobson & Truax, 1991) according to accepted interpretation standards in single-case design (Barlow et al., 2009). More detailed self-report inventories could also be sent digitally at specific points throughout the various phases for a more in-depth analysis of specific symptomatology. See Figure 3.1 for a visual representation of how such a real-world experiment might take place. The clinician researcher could then add comments/results to the centralized registry as described in the preceding paragraph.

Figure 3.1

An Overview of a Multiple Baseline Across-Individuals Design



While the advantages of practice-based single-case research for the evaluation of mental health apps lie in the potential of increasing the real-world evidence base for the effectiveness of these apps, there are also some potential disadvantages of this approach (Barlow et al., 2009). Firstly, clinicians in practice settings are traditionally time-poor (Hatfield & Ogles, 2004), and will only have limited time to add results and comments to any centralized registry. This also suggests that clinicians will only have limited time to monitor the response data for clients who participate in such research, and it is therefore reasonable to assume that most clinicians who participate would only be able to assess a small number of individuals simultaneously at best. Further considerations may also include the need for clinician researchers to engage in training in how to use certain data collection platforms such as Qualtrics, and the extra time it may take to collate and match participant data from multiple platforms. However, this research approach has the potential to be highly automated,

thereby allowing clinician researchers to be engaged without spending large amounts of time in administration duties associated with the research. Secondly, results from a single-case study do not have broad generalizability across large populations in a way that larger RCTs may be able to do. Thirdly, it can be difficult to find homogenous groups of individuals when attempting to replicate a single-case study, particularly if the individual has a complex presentation, and there will always be a debate of whether one individual's presentation is adequately similar to another's. Fourthly, it will be difficult to evaluate what proportion of positive outcomes may have been due to the app, and what proportion may have been due to any other treatment being provided by the clinician. Finally, single-case research is more prone to variability affecting results. For example, if an individual who is the subject of a single-case design experiences a significant negative event (like the death of a family member) during the intervention phase, it will potentially influence the results obtained that may inform the judgment of the effectiveness of that intervention. (It is also acknowledged that an intervention can be altered in response to such events in single-case research, which is an advantage over RCTs.) Whereas, if that individual was in a group research project involving hundreds of participants, the negative event would not necessarily have a substantial impact on the results if the study is looking at group mean differences.

The application of this type of research design to the evaluation of mental health apps could be further enhanced with the co-operation of the major app stores.

Future Research: A New Certification Framework

The current way that mental health apps are represented in the various app stores is inadequate on a number of levels, but particularly in regard to risk of harm and efficacy. While governments have started to act on the risk of harm, the possibility of clinician-led research could provide one solution for improving the amount of evidence for the effectiveness of mental health apps. If clinicians report back to a centralized registry, this will

have to be maintained and administered by some authority. Similar examples of centralized repositories in recent years have been successfully run by academic institutions either by receiving government grants or by providing the finances themselves, and in return receiving the kudos and respect for doing so. It is difficult to see where else the money may come from to finance this type of scheme. One goal for such testing is to create libraries of mental health apps that have independent evidence for their effectiveness, a pursuit widely called for in the literature (e.g., Mehrotra & Tripathi, 2018).

The two largest app marketplaces, the Apple App Store and Google Play store would need to be willing to re-categorize mental health apps in ways that consumers can clearly distinguish between apps that have acceptable scientific research, and those that do not. Apple and Google may be more willing to do this if there is a financial benefit. If a mental health app became certified by an independent clinician researcher, it would give that app a marketing credential over and above similar apps without such certification. The cost of this testing process would be the responsibility of the app developer and would have to be marketed to them as potentially offering greater financial returns for having undergone certification.

By re-classifying mental health apps as certified in this way, Apple and Google would also be playing a role in educating the public, because consumers may not be aware of the importance and need to establish evidence to back-up claims of effectiveness. With the involvement of the app stores, consumers could conceivably become educated in how to critically assess the evidence base of an app, the source of the app (mental health expert, academic institution, government authority, or a non-expert individual with questionable motives), learn to be more familiar with the contents of privacy policies, and to be educated on the importance of confirming the existence of a privacy policy. Under this model, consumers could also become more aware of the limits of mental health apps, and be directed

towards other help if necessary (e.g., to a General Practitioner or mental health professional). Apple and Google could also display contact details for emergency suicide support services on pages listing mental health apps.

Conclusion

Currently, searching the app stores for a mental health app that is efficacious and safe is problematic. There are limited numbers of mental health apps with research evidence to begin with, but those that do are difficult to differentiate from the many more that do not. Therefore, it may be time that the two largest app marketplaces, the Apple App Store and Google Play store, categorized mental health apps in ways that consumers and clinicians can clearly distinguish between those that have acceptable scientific research, and those that do not.

The literature is clear: there is a need for users of mental health apps to be able to identify those that are safe and have proven effectiveness (Firth et al., 2018). The implications for practicing psychologists in this regard is that they need to be certain the apps they are incorporating into therapy by way of recommendations to clients and patients will do no harm, and that there is evidence that justifies their use in the first place. To this end, research into the effectiveness of mental health apps needs to become a standard and visible form of credentialing. This, however, involves a financial consideration. Evaluation costs money and it is not known if consumers would pay a nominal fee to download an app that had official certification for its effectiveness over that of an app that was free but without such certification. What is more certain is that research support would increase the legitimacy of apps as tools to effectively treat mental health issues.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 4 – Study 1

In Chapter 3, a detailed review was provided of the current situation of government regulation in the area of mental health apps, issues relating to the safety of consumers, and how future research might provide information on these and other issues by embracing alternative experimental designs to the randomised controlled trial, such as the single-case design. This completed a comprehensive introduction on the current status quo of mental health apps. That is, there is very little research considering there are so many mental health apps and so many people using them. There also appears to be very little government regulation, with health authorities struggling to keep up with new innovations and the many mental health apps that are being developed and made publicly available at a frenetic rate. Finally, there seems to be a major barrier to more mental health apps being put through a scientific evaluation process: the randomised controlled trial is usually too expensive and too time-intensive for app developers to consider if they want to beat their competitors to the marketplace. In Chapter 4, an updated systematic review of the peer-reviewed literature is conducted to locate independent research or replication studies on the efficacy and/or effectiveness of mental health apps.

Chapter 4 – Study 1: Smartphone Apps as Digital Antidepressants: A Systematic Review of the Independent Research into Apps for Reducing Anxiety and Depression

Marshall, J. M., Dunstan, D. A., & Bartik, W. (Submitted). Smartphone apps as digital antidepressants: A systematic review of the independent research into apps for reducing anxiety and depression. *Journal of Medical Internet Research*.

Abstract

Background: The development and use of mobile apps to treat anxiety and depression is growing. Previous reviews have revealed that while research in this area is increasing, there is still only limited research available overall. Given the popularity of mental health apps amongst consumers, there is an urgent need for further research into the efficacy and effectiveness of these digital tools. When examined more closely, the limited number of previous efficacy studies in this area have largely been completed by researchers with an association to the app being examined. No previous review could be located that specifically focused on the current level of independent research and replication in this space.

Objective: We aimed to systematically review the literature for independent and/or replication studies on apps for anxiety and/or depression, and provide a meta-analysis for comparison with findings of other reviews that have incorporated all research. If results from independent research resembled the overall pattern of research that included studies by those who have a development and/or financial interest in the apps being studied, this would provide compelling evidence that claims of effectiveness across all studies were applicable and well-founded.

Methods: Following the PRISMA framework, five databases (ProQuest, PubMed, Google Scholar, ScienceDirect, and the Cochrane Library) were systematically searched using 50 key word search terms. Inclusion criteria included: the app had to offer a therapeutic treatment for anxiety and/or depression rather than a narrowly defined skill (such as hypnosis, or deep breathing etc.), or symptom monitoring, thought recording, or diagnostic tool (although apps could have any of these tools as part of offering a therapeutic treatment), and no author could have any association or relationship with the app being studied.

Results: Four articles describing independent research on five apps were located. All articles found statistically significant results for improvements in anxiety symptoms, but only one study contained significant results for reductions in depression. Given the low number of studies, it was decided that a meta-analysis would not produce meaningful data and was therefore not undertaken. This decision was supported following application of the Cochrane risk of bias tool, and the Newcastle-Ottawa Scale which rated each study as "Poor Quality".

Conclusions: While the independent research suggested that apps may improve symptoms of anxiety, there was less evidence for depression. Given the low number of studies, and their variable quality, caution is required in interpreting results. This review has nevertheless illuminated the need for further independent research and replication in order to give mental health apps more legitimacy.

Trial Registration: This systematic review is registered with PROSPERO, registration number CRD42019137821 (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=137821).

Smartphone Apps as Digital Antidepressants: A Systematic Review of the Independent Research into Apps for Reducing Anxiety and Depression

Smartphones and their applications (apps) are used in many aspects of our lives, including in the monitoring, assessment, and treatment of physical and mental health. There are currently 2.5 million apps available on Google Play, and 1.8 million in the Apple App Store, the two biggest app marketplaces (Statista, 2019a). Calculating how many health apps are included in these numbers is difficult to gauge, as definitions vary about what a "health" app is, and what a "mental health" app is. However, health apps (including those related to mental health) are one of the fastest growing categories of apps. There are somewhere between 97,000 (Nehabaluni, 2018) and 350,000 (Research2Guidance, 2017) health apps, with an average of more than 200 new health apps made publicly available each day (IQVIA Institute, 2017).

Of the total number of health apps available 10,000 relate to mental health (Torous, Firth, et al., 2018). This number makes it difficult for consumers to know how to choose a mental health app, especially as using the limited search capabilities in the app stores can produce unpredictable results (Alyami et al., 2017; Shen et al., 2015). Consumers currently use the app store reviews and ratings as their main source of advice about which app to download (Huang & Bashir, 2017), but this is an unreliable method (Xie & Zhu, 2015).

There are now several websites administered by government departments, academic institutions, and not-for-profit organizations that provide information about mental health apps. For example, PsyberGuide (One Mind, 2019), the National Health Service (National Health Service, 2019), Head To Health (Australian Government Department of Health, 2019a), Beacon (Australian National University, 2019), Reachout (Reachout.com, 2019), Health Navigator (The Royal New Zealand College of General Practitioners, 2019), and

others. There are also more sophisticated frameworks that have been developed to help evaluate mental health apps, including by the American Psychiatric Association (Torous, Firth, et al., 2018), and others (Zelmer et al., 2018). There are at least 17 app-rating frameworks globally (Nielsen & Rimpilainen, 2018). As consumers rely mostly on reviews and ratings by other consumers, and also by recommendations made on social media and via word of mouth (Rubanovich et al., 2017), it is uncertain to what extent consumers rely on those app-rating mechanisms mentioned above. The more sophisticated frameworks are arguably of more value to clinicians, but it is uncertain to what extent clinicians rely on such frameworks. Many clinicians only use and recommend digital mental health resources if they receive such recommendations from other clinicians (Hempel et al., 2018). There are still sizable numbers of clinicians not engaging in digital mental health resources (Schueller et al., 2018). Consumers, up until now, have not appeared to show an overwhelming interest in the level of evidence or government oversight of mental health apps – this call has come from researchers, governments, medical and mental health organizations, and clinicians. It may be up to willing clinicians, as the ones who have face-to-face contact with consumers, who lead the charge in educating the public about the need and importance of research in this area.

Why do we Need Mental Health Apps?

Although there are currently several effective evidence-based approaches to treating anxiety and depression, these are often not accessible to people for varying reasons (Gunter & Whittal, 2010). Mental health apps provide one possible solution to this problem, given that smartphones are owned by approximately 70% of the world's population (Barboutov et al., 2017), with this figure expected to continue to rise for at least the next decade (Statista, 2019b). As the level of smartphone ownership grows, so does the opportunity for apps to offer mental health treatment that many people do not receive face-to-face. There are many other potential advantages that mental health apps offer, including: anonymity, enhanced

activities to reinforce strategies learned in face-to-face therapy, receiving reminders about taking medication and completing tasks, and the opportunity for people with less severe symptoms to undergo treatment without taking up scarce primary care resources. Digital mental health options, including apps, may also allow specific populations to access treatment they may not otherwise be able or willing to, such as First Nations peoples, culturally and linguistically diverse (CALD) communities, people in rural locations (Hernan et al., 2010), low socioeconomic groups, and teenagers (Wang et al., 2018). For all of these reasons, it is important to research the efficacy and effectiveness of mental health apps.

Previous Research

While being helpful in many ways, there is virtually no guidance in the above-mentioned websites and frameworks about the best way to measure the efficacy of a mental health app. New approaches to examining mental health apps have previously been outlined in the literature (Clough & Casey, 2015), but a consensus is yet to be reached. Previous research reflects this uncertainty. A recent meta-analysis of RCTs into the efficacy of apps to treat symptoms of anxiety and depression (Linardon et al., 2019) found that they are more effective for improving symptoms of depression than anxiety. Specifically, apps offered greater statistically significant improvements in depression compared to control conditions such as waitlists, placebo apps, and information / education apps. While apps also offered greater statistically significant improvements in social anxiety and generalized anxiety compared to control conditions, there was no such significant difference with in-person or other computerized treatments. It is also important to recognize that apps with adjunctive support from clinicians had greater effect sizes for both anxiety and depression compared to studies without such support. Reviews between 2017 and 2019 saw an increase in RCTs into the efficacy of mental health apps found in the literature from 20 studies (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al.,

2017) to 66 studies (Linardon et al., 2019). While the increase in RCTs into the efficacy of mental health apps for anxiety and depression is pleasing, it is still recognized that this may not be the most efficient method for boosting the evidence base (Torous et al., 2021).

Although several past reviews of mental health apps have largely found them to be effective in reducing symptoms of anxiety and depression (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Linardon et al., 2019), there has been considerable heterogeneity across studies. For example, in some studies participants were given daily guidance on how to use the app (Roy et al., 2017) while in others there was little guidance (Flett et al., 2019); in some studies participants used the app for as little as 10 days (Howells et al., 2016) while in others the app was used for as long as 12 weeks (Boisseau et al., 2017); some studies measured outcomes at follow-up points as long as 12 months after the intervention period was over (Roy et al., 2017) while other studies had no long-term follow-up at all (Paul & Fleming, 2019). Therefore, there is considerable uncertainty about what to measure, as much as how to measure, in deciding if an app is efficacious. For example, research on mental health apps is yet to look at factors or measures such as productivity at work, level of service utilization, and interpersonal functioning, and appears to have focused on levels of psychological distress. There is also the possibility of the digital placebo effect (Torous & Firth, 2016) and Hawthorne effect (McCambridge et al., 2014) in previous research, but have rarely been mentioned. Even with these difficulties and uncertainties, there has been a dearth of research on the efficacy of apps.

Two literature reviews on mental health apps for anxiety (Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017) and depression (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017) found a total of 20 apps with research published in peer-reviewed journals. Every research project in these reviews involved individuals that had an association with the app

being studied. That is, there were no examples of independent research on apps for anxiety or depression that were found in the literature. Nevertheless, the little research that has been done has revealed some other noteworthy characteristics.

Firstly, there are concerns about the methods used to evaluate the efficacy of apps. In part this arises because different apps may function in different ways; therefore, it may not be possible to evaluate each in the same way. For instance, many mental health apps recommend users access the app when symptoms of psychological distress are present. Such apps are taking a "mental health first aid" approach to intervention and lack precise instruction about duration and dosage (Fleming et al., 2018). In contrast, some apps take a "dosage" approach and may give instructions, such as "use the app for 15 minutes each day for six weeks." This is more in line with evidence-based treatment guidelines for the use of psychotropic medication and frequency / length of psychotherapy sessions that includes assessment, treatment, and maintenance phases (Mrazek & Haggerty, 1994). With mental health apps taking such different approaches, it is difficult to examine them systematically. The other issue related to research design is that often by the time a large scale, randomized controlled trial has been completed, the app may have been updated, or disappeared from app stores altogether, as more popular apps with improved features come on to the market.

Secondly, the content of apps and their developmental process varies greatly. Most have not been developed with the involvement of mental health experts, government funding, or academic institutions (Alyami et al., 2017; Shen et al., 2015). There appears to be a need for greater training in how to use mental health apps for both consumers and clinicians (Armstrong et al., 2018), and a better understanding of the ethical concerns of mental health apps (Broussard & Teng, 2019).

Finally, with a number of quality systematic reviews (as mentioned above) on the efficacy of mental health apps already in existence, it seems that there is now a need to delve

deeper into specific aspects of these apps. This paper has identified one such need to focus on the independent research into the efficacy of mental health apps, as no such review could be located in the literature.

Independent Research

Pure independent research free from the potential influence of external forces, including researchers having an interest in the outcome of that research, is an idealized and unrealistic outcome of most scientific practice (Ioannidis, 2005). After all, most researchers are paid by institutions, and their research may be generated or made possible by government funding, or funding from other sources, and may be based on the outcome of their own previous research. A possible buffer to this is public funding or open-source funding, so the resource is freely available for the public good. However, minimizing potential bias due to possible conflicts of interest remains an important goal in medical science, and is doubtless true of any branch of scientific endeavor. This is especially relevant when examining the effectiveness of a product, whether it be mental health apps, vitamin supplements (Gaziano et al., 2012), or weed killer (Donley & Gillam, 2018). In cases of product testing, if the developer is involved in the testing process, there will always be questions about the legitimacy of results obtained from that process if there is no independent research backing it up (Young et al., 2008).

A relevant example concerns studies of antidepressant medication. A recent meta-analysis showed that early research on antidepressants (much of it funded by pharmaceutical companies that developed them) produced higher efficacious and effective results than more recent studies (that showed significantly lower effect sizes) by researchers who have no association to the medications being tested (Cipriani et al., 2018). This also found that 82% of all studies on antidepressant medications were at moderate or high risk of bias using the Cochrane Bias Tool due to the involvement of pharmaceutical companies. One recent study

even found that a popular antidepressant medication, sertraline, had no clinically significant effect on symptoms of depression, in contrast to the results obtained by previous research that had involved the medication's developer (Lewis et al., 2019).

The above is but one example related to mental health that highlights the need for further independent research. The same could be said for research on psychological treatments, whether it be person-delivered or Internet-delivered, or a combination of both. Increasing the levels of independent research in all these areas will potentially improve the legitimacy of any treatment, and this is why independent research is so important. However, the need to increase independent research in all areas of mental health is rarely a topic of discussion in the literature.

For the purposes of this manuscript, the authors have considered independent research in the area of mental health apps as research conducted by individuals, institutions and organizations who were not involved in the development of the app, and who do not stand to gain financially or otherwise from it. While much of the research conducted by those who have an association with the app being studied may be of acceptable quality, it is important that independent research in this area increase to further legitimize the evidence base and lessen concerns regarding bias. As stated above, this is important in all forms of product testing. Given that the development of mental health apps is still in its early stages, there is the opportunity to advocate for and produce more independent research to bolster findings of other research that may have or will be produced by individuals with an association to the app being studied. Having a base of independent research will strengthen trust from both consumers and clinicians. It will also contribute to an acceptance that mental health apps are a legitimate mental health treatment / management option than what may otherwise be if the proportion of independent research did not increase beyond its current levels. However, independent evidence does not automatically equate to "good" or reliable evidence – it still

needs to adopt accepted scientific standards, as is the case for non-independent research. All research, independent or otherwise, should be analyzed for quality.

There are barriers that have contributed to the current low level of independent research in the area of mental health apps. There may be app developers who are not willing to co-operate with independent researchers, and this may lead to legal reasons why independent researchers do not examine certain apps (and the legal issues may vary from country to country). This may therefore lead to difficulties for independent researchers to access relevant data that may have been accumulated by app creators during the development and testing phases. (However, it can also be supposed that there would be many app developers who would welcome the involvement of independent researchers to assess their app.) There may also simply be a current lack of interest from researchers in evaluating "other people's" apps, and a greater interest and desire to develop and test their own apps. This may also explain why little government funding has been distributed to independent researchers; perhaps independent researchers are simply not asking for such funding. Therefore, it may be easier for developers to commission their own "independent" research on their app, rather than waiting for others to do so.

The Current Study

The aim of the present study was to systematically review the independent research into the effectiveness of mental health apps for reducing symptoms of anxiety and/or depression. The study built on previous published research that had shown mobile mental health apps to be effective in reducing such symptoms (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). However, no other published study has uniquely focused on independent / replication studies.

Given that previous research had not adequately examined the issue of 'dosage' or frequency of use, this study also sought to identify those from the shortlisted articles that

recommend to participants about how to use the app, and whether there were differences in instructions across studies. Finally, the methodologies of independent app studies were examined for levels of similarity / disparity.

The justification for looking at independent research is that its presence in any field of scientific enquiry is seen as advantageous, offering strength to the body of overall research due to being less prone to accusations of bias (see above comments). However, it must also be assessed for quality. In terms of mental health apps, it is hoped that a study examining the independent evidence base may promote the need for more, and encourage a wider discussion in the e-mental health literature about its place in this research area.

Methods

Search Strategy

The methodology of this systematic review was informed by the MARS (Meta-Analysis Reporting Standards) (Appelbaum et al., 2018), AMSTAR (Assessment of Multiple Systematic Reviews) (Shea et al., 2007), and PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) (Moher et al., 2009) protocols. This systematic review is registered with PROSPERO, registration number CRD42019137821.

Four researchers, including the lead author, searched five databases: ProQuest, PubMed, Google Scholar, ScienceDirect, and the Cochrane Library. All researchers searched all five databases for all 50 search terms. If one researcher shortlisted an article for consideration, it was placed on the master list. By having all four researchers doing the same search and shortlisting every article even if only selected by one researcher, the odds of collecting all relevant articles were improved compared to the situation of less researchers doing the search. The lead author then examined each shortlisted Abstract to ascertain if it met the complete inclusion criteria, and downloaded the entire article if the Abstract did not provide the necessary information to make this decision. (See Figure 4.1 for the PRISMA

flow chart of article selection.) It was planned that disagreements about inclusion or exclusion of studies would be resolved through discussion until consensus was reached. From these discussions, there were no disagreements about shortlisted articles. The date range covered 1st January 2000 to 30th June 2019 (smartphones were not widely available prior to the year 2000). A total of 50 key word searches were completed (see Textbox B.1 in Appendix B for the complete list).

Inclusion Criteria

The inclusion criteria were:

1. The article needed to be in a peer-reviewed journal and either in English or translated into English;
2. The article needed to be an intervention study with outcomes for either anxiety and/or depression symptoms (i.e., study protocols, conference abstracts, usability studies, and feasibility studies without outcome measures for anxiety and/or depression were excluded);
3. The app needed to offer a therapeutic treatment rather than a narrowly defined skill (such as hypnosis, or deep breathing etc.), or symptom monitoring, thought recording, or diagnostic tool (although apps could have any of these tools as part of offering a therapeutic treatment);
4. The study could not examine an app as an adjunct to any other type of therapy (i.e., the focus of the research needed to look at how using an app could exclusively improve symptoms, although there could be a control group with another type of therapy);
5. It is preferable that participants were randomized to a treatment group - studies that did not do this were not excluded, but it was noted; and
6. No author could have any association or relationship with the app being studied as per our definition of independent research above.

Data Analysis

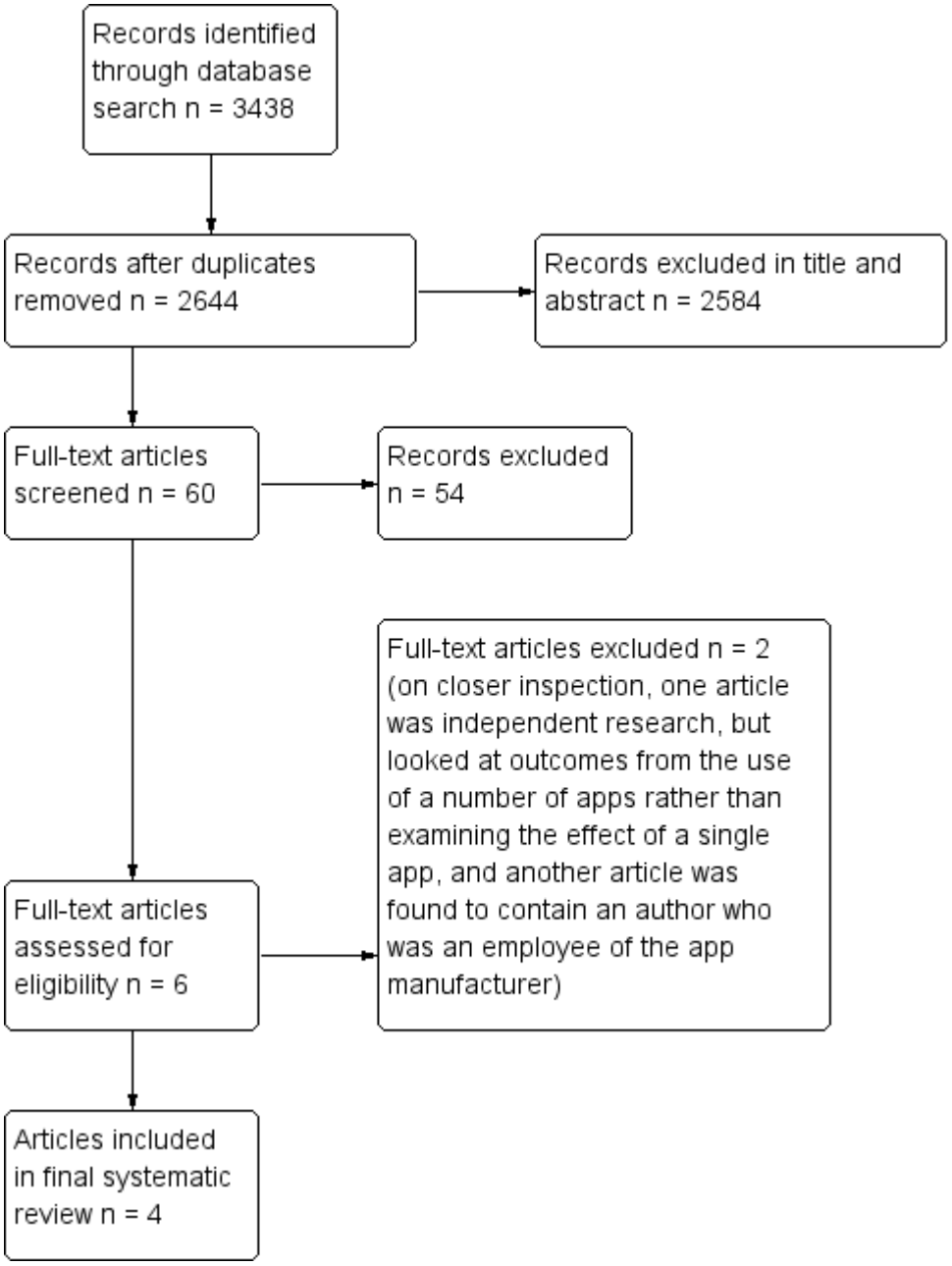
A meta-analysis was planned using results from all included studies. The data was extracted from the included articles by the lead author, and the authors of all included articles were contacted for further statistical information. As there were so few studies found to be eligible, and the studies were highly heterogenous, it was decided that a meta-analysis would not produce meaningful results. Given the low number of studies, this allowed a more thorough assessment of each article, which is often not possible in systematic reviews where larger numbers of articles are included. Descriptive statistics were used to analyze each article. Each of the final shortlisted studies were evaluated for quality using the Newcastle-Ottawa Scale (Luchini et al., 2017).

Results

An initial search of the databases revealed a total of 60 articles for closer examination. This was eventually shortened to six articles, mainly due to one or more of the authors being involved in the app's development, or the app itself being inappropriate for review because it did not meet the inclusion criteria. The list was further shortened to four articles when it was confirmed that: Roy et al. (2017), although involving independent research, had studied the effectiveness of using a number of mental health apps simultaneously and without any way to examine each app on its own; and, Howells et al. (2016) contained an author who was found to be an employee of the app manufacturer. See Figure 4.1 for details. No replication studies were found in the literature.

Figure 4.1

PRISMA Flowchart of Study Selection



There were only four articles for inclusion in the meta-analysis. As there was considerable variability in their methodologies (see Tables 4.1, 4.2, and 4.3 for further information on the methodologies of included studies), it was deemed by the researchers that a meta-analysis would not produce meaningful results (Valentine et al., 2010). This decision was supported following application of the Newcastle-Ottawa Scale which showed that each article fell into the "Poor Quality" category due to the absence of a long-term follow-up measure. Furthermore, when additional information about the statistical analyses was sought from each of the corresponding authors, only one author (Flett) supplied this information. Had a meta-analysis been carried out without using known data from the other three articles, estimates would have been used. For a summary of the four included studies, see Tables 4.1, 4.2, and 4.3.

Boisseau et al. (2017) had a single intervention group that obtained pre to post data for 15 participants, all of whom were U.S. community and hospital outpatients with a diagnosis of obsessive compulsive disorder (OCD) and a mean age of 36.6 years. The group used the *LiveOCDFree* app to provide exposure and response prevention (ERP) therapy, with instructions to use the app for at least one hour per day for 12 weeks (84 days). They found statistically significant improvements for anxiety, but not depression.

Flett et al. (2019) had one control group (n = 67) and two intervention groups of New Zealand undergraduate students with a mean age of 20.1 years. The intervention groups used either the *Headspace* (n = 67) or *Smiling Mind* (n = 58) apps to provide a mindfulness-based treatment, with instructions to use the app for 10 minutes per day for the first 10 days, then use it at the participant's discretion for the next 30 (for a total of 40 days). Both apps produced statistically significant reductions in anxiety. *Headspace* also produced a statistically significant reduction in stress. Neither app produced statistically significant reductions in depression.

Lee and Jung (2018) had a control group of 86 participants and an intervention group of 77 participants, all Canadian undergraduate students with a mean age of 20.6 years. The intervention group used the *Destressify* app to provide mindfulness-based therapy, with instructions to use the app for 5 days per week for 4 weeks (28 days). Statistically significant reductions in anxiety were achieved from pre- to post-treatment, but there was no significant change in depression or stress.

Paul and Fleming (2019) had a single intervention group that obtained pre to post data for 16 participants, all of whom were U.S. undergraduate students with a mean age of 19.8 years. The group used the *MindShift* app to provide anxiety-focused CBT therapy, with instructions to use the app for at least 15 minutes per day, 5 days per week, for 3 weeks (21 days). Statistically significant reductions were found in somatic and generalized anxiety, and depression, from pre- to post-treatment, but not for panic.

All four studies used different measures of anxiety and depression, with the only common measure for two studies (Flett et al., and Lee & Jung) being the Perceived Stress Scale to measure stress.

Risk of bias was measured using the Cochrane risk of bias tool. High risk of bias was found in relation to lack of randomization and concealment of intervention for two of the studies. In two other studies, unclear risk of bias was found in relation to outcome data. For three studies there was risk of bias in relation to selective reporting, due to the lack of response to a request for further statistical information. No study identified participants as being paid. See Table 4.1 for further details.

Table 4.1

Quality Assessment [g-i] for Included Studies

Study	1[a]	2[b]	3[c]	4[d]	5[e]	6[f]
Boisseau et al. (2017)	-	-	-	-	#	+
Flett et al. (2018)	+	#	#	#	+	#
Lee & Jung (2018)	+	+	#	#	#	#
Paul & Fleming (2019)	-	-	-	#	+	#

[a]1 = Random sequence generation.

[b]2 = Allocation concealment.

[c]3 = Blinding of participants and personnel.

[d]4 = Blinding of outcome assessment.

[e]5 = Incomplete outcome data.

[f]6 = Selective reporting.

[g]+ = Study was adequate (low risk).

[h]- = Study was inadequate (high risk).

[i]# = Study had unclear risk.

Table 4.2

Summaries of Articles Included in Review: Sample Type, Participants, and Aims

Study	Sample type	N (intervention)	N (control)	Mean age (years)	Study aims
Boisseau et al. (2017)	U.S. community *and* hospital outpatients.	21 (but finished with 15)	N/A	36.6	Focus was on OCD, but also examined general anxiety and depression.
Flett et al. (2018)	NZ undergraduate students.	72 for Headspace (but finished with 67), 63 for Smiling Mind (but finished with 58)	75 (but finished with 67)	20.1	Depression, anxiety, stress, flourishing.
Lee & Jung (2018)	Canadian undergraduate students.	77	86	20.6	Stress, anxiety, depression.
Paul & Fleming (2019)	U.S. undergraduate students.	18 (but finished with 16)	N/A	19.8	Focus was on anxiety, but also examined depression.

Table 4.3

Summaries of Articles Included in Review: Design, Methodology, Outcome Measures, and App Names

Study	Design	Intervention details / methodology	Notes and other intervention aspects[a]	Outcome measures used[b-m]	App name
Boisseau et al. (2017)	12 weeks app use – 1 hour per day for 12 weeks.	The app provides Exposure and Response Prevention (ERP) treatment for OCD. Participants downloaded the online app user guide and watched the tutorial video within the app.	Only available in the Apple App Store ($46·99). Assessment occurred at three time points: Baseline, Mid-treatment (6 weeks), Post-treatment (12 weeks). No follow-up data or control group.	BDI, BAI, and Y-BOCS-SR	*LiveOCDFree*
Flett et al. (2018)	40 days total app use – 10 mins per day for the first 10 days, then use the app at participant discretion for remaining 30 days.	Three-arm randomized controlled trial. Both intervention apps are based on mindfulness. Control group used an app named Evernote to test for the digital placebo effect.	Apps available in both the Apple App Store and Google Play (free). Assessment occurred at three time points: Baseline, 10 days, 30 days. No follow-up data.	CES-D, HADS-A, PSS, and Flourishing Scale	1. *Headspace* 2. *Smiling Mind*

Lee & Jung (2018)	4 weeks app use – 5 days per week.	Participants were randomized into the control or intervention group. They were told to use the "Core Plan" on the app, and told not to engage with the other features of the app. Intervention app is based on mindfulness and CBT. Control group was waitlisted.	App available in both the Apple App Store and Google Play (free for basic version, or $9.99 for Pro version). No follow-up data.	PSS, STAI, and QIDS-SR	*Destressify*
Paul & Fleming (2019)	3 weeks app use – minimum of 15 mins per day, 5 days per week.	The app has an anxiety focus, and is underpinned by CBT. Participants were only included if they were not receiving any other treatments for anxiety. Participants were given a tutorial on using the app.	App available in both the Apple App Store and Google Play (free). Assessment occurred at 3 time points through the study: Baseline, after 1 week, and after 3 weeks. No follow-up data or control group.	PHQ-15, GAD-7, PHQ-9, and PHQ-PD	*MindShift*

[a]$ amounts in Australian Dollars as at August 2019.

[b]BDI = Beck Depression Inventory.

[c]BAI = Beck Anxiety Inventory.

[d]Y-BOCS-SR = Yale-Brown Obsessive Compulsive Scale – Self-Report.

[e]CES-D = Center for Epidemiologic Studies – Depression Scale.

[f]HADS-A = Hospital Anxiety Depression Scale – Anxiety.

[g]PSS = Perceived Stress Scale.

[h]STAI = State Trait Anxiety Inventory.

[i]QIDS-SR = Quick Inventory of Depressive Symptomatology – Self-Report.

[j]PHQ-15 = Patient Health Questionnaire – Somatic symptoms scale.

[k]GAD-7= Generalized Anxiety Disorder – 7-item scale.

[l]PHQ-9 = Patient Health Questionnaire – 9-item scale.

[m]PHQ-PD = Patient Health Questionnaire – Panic Disorder scale.

Table 4.4

Summary of Statistics of Each Article in Review – Part A[a]

Study	Hedges' g	Standard error	Lower limit	Upper limit
Boisseau et al. (2017)				
Flett et al. (2018); *Headspace*	.17 (depression), .15 (anxiety), .17 (stress), .01 (flourishing).	Depression = 1.47 Anxiety = .63 Stress = .95 Flourishing = 1.21	.946 (depression) .286 (anxiety) .742 (stress)	4.165 (depression) 1.742 (anxiety) 2.455 (stress)
Flett et al. (2018); *Smiling Mind*	.12 (depression), .12 (anxiety), -.05 (stress), -.07 (flourishing).	Depression = 1.48 Anxiety = .64 Stress = .95 Flourishing = 1.23	.227 (depression) .336 (anxiety) -.192 (stress)	3.805 (depression) 1.824 (anxiety) 2.192 (stress)
Lee & Jung (2018)				
Paul & Fleming (2019)				

[a] Blank boxes indicate that these statistics were not available from the articles, and the authors did not respond to requests for this information.

Table 4.5

Summary of Statistics of Each Article in Review – Part B

Study	*p* (for intervention period only)	Effect size (for intervention period only)
Boisseau et al. (2017)	.019[a] (YBOCS-SR)	η^2_p = .25 (YBOCS-SR)
	.058 (BDI)	η^2_p = .21 (BDI)
	.044[a] (BAI)	η^2_p = .23 (BAI)
Flett et al. (2018); *Headspace*	Depression = > .05	Cohen's d = .17 (depression),
	Anxiety = < .05[a]	Cohen's d = .23 (anxiety),
	Stress = < .01[a]	Cohen's d = .30 (stress),
	Flourishing = > .05	Cohen's d = .17 (flourishing).
Flett et al. (2018); *Smiling Mind*	Depression = > .05	Cohen's d = .13 (depression),
	Anxiety = < .05[a]	Cohen's d = .29 (anxiety),
	Stress = > .05	Cohen's d = .18 (stress),
	Flourishing = > .05	Cohen's d = .01 (flourishing).
Lee & Jung (2018)	Stress = .06	η^2_p = .02 (stress)
	Depression = .09	η^2_p = .02 (depression)
	Anxiety = .02[a]	η^2_p = .05 (anxiety)

Paul & Fleming (2019)	Somatic anxiety = < .01[a]	η = .34 (somatic anxiety)
	General anxiety = < .01[a]	η = .51 (general anxiety)
	Depression = < .01[a]	η = .30 (depression)
	Panic = .99	η = .00 (panic)

[a] Identified as statistically significant in the article.

Discussion

Principal Findings

A systematic review of the literature on apps designed to reduce symptoms of anxiety and/or depression identified four articles that met the inclusion criteria. These examined five different apps. All four articles found that the studied app produced a statistically significant reduction in symptoms of anxiety and concluded that each app had potential to be an effective treatment option. However, a statistically significant reduction in symptoms of depression was demonstrated in only one study. The effect sizes were generally small across all studies, yet all studies concluded that there is sufficient evidence to suggest that research into the effectiveness of apps for reducing symptoms of anxiety and depression should continue. The magnitude of these effect sizes corresponds to previous research (Linardon et al., 2019). This may suggest that the role of apps in the future may be most useful as a standalone treatment for people with milder symptoms of anxiety and depression, and/or potentially may be most effective as an adjunctive treatment. These are aspects to be explored in future research.

The selected studies included a total of 386 participants: 233 in intervention groups and 153 in control groups. All but 21 of the intervention participants were undergraduate students. No replication studies were found. Overall, these findings support the argument that further independent research using diverse samples of clinical and community groups is needed.

The five apps in the articles for inclusion in this review were: *LiveOCDFree*, *Headspace*, *Smiling Mind*, *Destressify*, and *MindShift*. *LiveOCDFree* is based on ERP treatment and is designed for use by people with OCD. *Headspace* and *Smiling Mind* are both apps that teach mindfulness skills, and in particular, aim to train users in mindfulness

meditation without specifically targeting either anxiety or depression. *Destressify* teaches CBT and mindfulness skills and specifically targets stress and anxiety using reference to neuroscience concepts such as neuroplasticity. Finally, *MindShift* teaches CBT skills with an emphasis on managing anxiety symptoms.

This literature review has illuminated a number of shortcomings in the existing independent research on mental health apps. Notably, there is a high degree of heterogeneity, which reaffirms observations made in previous reviews of the overall research (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). This makes it difficult to compare studies and reveals a disparate approach by researchers to the evaluation of these digital tools. For example, two of the four studies did not use a control group and, as noted above, three used undergraduate students with a mean age between them of 20.2 years, which is a narrow snapshot of the wider community. Of the two articles that used a control group, none included a "treatment-as-usual" group, such as using a face-to-face evidence-based therapy or medication. Instead, Flett et al.'s control group used a placebo app, and Lee and Jung's control group was waitlisted. The failure to use a treatment-as-usual control group appears to be a common limitation of many mental health app studies to date. Another difficulty in analyzing the previous independent research is that reported results used different statistics and effect sizes, and there were different procedures for the use of the apps. For instance, length of treatment (or dosages) varied between studies, as did the quality of instructions provided. Recommended dosages ranged from 21 to 84 days, and instructions varied from minimal guidance to daily text-messaged directions. No study included actual app usage data. Being independent research, it may be that the researchers were unable to access such data without having the involvement of the app developers and the permission of participants. There was also no long-term follow-up data in any of the studies. All of which points to, amongst other things, the need for more quality independent

research to accompany the overall evidence base, and for a more homogenous approach to be taken by more researchers when it comes to testing the efficacy and effectiveness of mental health apps.

In contrast to this variability, all five apps in this review claimed to use evidence-based frameworks, including CBT, mindfulness, and ERP. Although it is important that a mental health app is developed from an evidence-based perspective, there may be other factors that impact on an app's efficacy and effectiveness. For example, therapeutic alliance (Tang & DeRubeis, 1999), motivation of patient (Addis & Jacobson, 2000), and chronicity (Hamilton & Dobson, 2002) have been shown to effect the success of face-to-face psychotherapy. Similarly, factors such as usability / ease of use, aesthetics, and level of gamification may also impact on the success of a mental health app (Rickard et al., 2016). Although the content of an app may be similar to an older web-based program designed for use on a desktop computer, it is still unknown if the delivery of the program by smartphone changes the experience of that content. Intuitively, the content should have similar effects, but research has yet to prove this. For example, does the size of the smartphone screen as opposed to the size of a desktop computer screen have any impact on effectiveness? Does the ability for a smartphone app to be accessed in a park differ to the effectiveness of a web-based program on a desktop computer in an office or a bedroom? There is still much to study in this area.

Future Research

None of the five apps in this review used a behavior treatment model, which is a shortcoming of many publicly available mental health apps (Naslund et al., 2017). Frameworks such as the health belief model, theory of planned behavior, transtheoretical model, and social cognitive theory may provide ideas for future mental health app development that enhance an app's efficacy (Naslund et al., 2017).

With such differences in methodologies, future research will need to look at how a single methodology can be applied to a broad range of mental health apps. For example, standardizing the dose to a set time period, standardizing the amount of training/guidance given to users, recruiting diverse samples of participants, and reducing threats to internal validity by demonstrating a clear and continuing response to treatment.

One solution may be the use of single-case experimental designs (Barlow et al., 2009). Such designs can utilize strong statistical analyses, offer flexibility in the administration of treatment, offer insight into the response characteristics of individual participants, and can be accomplished relatively quickly compared to larger scale randomized controlled trials. Furthermore, it may be an opportunity for practicing clinicians to participate in the research process. If practicing clinicians used a standardized protocol to examine the effectiveness of mental health apps in their clinical workflow, and were able to contribute their findings to a centralized database, it would not only increase the overall evidence base for mental health apps, but also the independent evidence (Marshall et al., 2020e).

There may also be a place for an independent testing authority in the area of health apps generally (Marshall et al., 2020e). There are already independent testing organizations for a wide cross-section of other products, such as vitamin and mineral supplements (Labdoor, 2019), to furniture (SGS, 2020), and more. There are potential barriers to this kind of testing process, such as who will fund the research, what qualifications would researchers within the testing authority require, and how are apps chosen to be tested (if there are more than 10,000 related to mental health alone) (Torous, Firth, et al., 2018). However, it is worth having discussions about such an entity, whose establishment may be welcomed by app developers who genuinely believe in the ability and potential of their product to help in the management of mental illness.

Limitations

There are several limitations in this research. Firstly, this review only examined articles related to apps that claimed to offer a therapeutic treatment for anxiety and/or depression. Articles that dealt with apps which did not specifically measure outcomes related to this were not included. It is important to acknowledge this because there are apps that have deliberately attempted to market themselves without referring to such potentially stigmatizing terms as “anxiety” and “depression”, yet may still be beneficial for people experiencing these conditions (Bakker et al., 2016). Secondly, this review only searched for peer-reviewed articles in English, and this may have potentially ruled out other relevant but unpublished studies thus creating potential publication bias, and may have missed relevant articles in a language other than English. There are other databases that could have been selected in place of the ones that were used, but time and resources meant that no further databases were searched beyond the five mentioned. Thirdly, this review did not consider apps that acted as an adjunct to other types of treatment, and this is important because apps potentially offer consumers a more enhanced face-to-face therapy experience. Finally, given the nature of the current mental health app environment, new apps are being released daily and research into their effectiveness may be gathering pace, so the state of play of any category of apps can change quickly.

Conclusions

In conclusion, the present review has extended our understanding of the previous research by highlighting the lack of published independent research, and the non-existence of published replication studies in the area of mental health apps. Given the huge uptake by and impact of mental health apps on consumers, it is necessary that with so few published independent studies existing an important finding in the area of mental health app research like this deserves to be clarified as such in the literature, as it reflects a weakness in the

evidence base. While much of the existing research is of high quality, the vast majority is undertaken by individuals and organizations involved in the development of the apps. However, the area of e-mental health is still a young, albeit quickly growing, focus of research. There is time to encourage further independent mental health app research to provide balance and context to research completed by app developers, and to avoid many of the criticisms levelled at early antidepressants research that was funded by pharmaceutical companies. This older research is now being reviewed by independent researchers with new findings that suggest a moderate to high risk of bias is present in much of that research (Cipriani et al., 2018).

There are more app developers claiming that mental health apps now have a substantial evidence base and have proven efficacy, but until more independent research is done, and until replications are published, there will always be the possibility that bias exists in the research due to app developers providing the vast majority of the evidence. Further research, using a standardized methodology, is required to validate more apps and generate confidence in consumers and clinicians of the efficacy of apps in reducing symptoms of anxiety and/or depression. This increase in independent research would strengthen the existing and future findings of other quality research that may have been or will be completed by individuals with associations to the app. App developers should want this fact about the lack of independent research to be known because it is in their interests that more independent research is completed so that their apps can be further validated with more compelling claims of efficacy and effectiveness. In this way, the potential for genuinely effective mental health apps to reduce the current burden on primary care services, and create access to those unable or unwilling to attend other service options, will be realized.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 5 - Study 2A

In Chapter 4 – Study 1, the literature was searched for peer-reviewed published research on the efficacy and effectiveness of apps for alleviating symptoms of anxiety and/or depression. In Chapter 5, in order to get a complete perspective, it was deemed appropriate to also do a systematic search of the major app stores as a separate study. This would give a clearer indication of how many apps designed to alleviate symptoms of anxiety and/or depression are publicly available per se, and would provide an opportunity for locating apps that claim research evidence to then independently check for the existence of that evidence if it was not a study that was located through the literature search of Study 1. Study 2 also afforded the opportunity to develop a new protocol for conducting app store searches, as no such protocol could be located in the literature that can be used in the same manner as the PRISMA (Moher et al., 2009) and AMSTAR (Shea et al., 2007) protocols can be used to guide literature reviews. In Study 2A, the focus is on the evidence base of mental health apps for anxiety and/or depression, the proportion of apps that have been developed with expert or government input, and what proportion of mental health apps are free to download.

Chapter 5 – Study 2A: The Digital Psychiatrist: In Search of Evidence-Based Apps for Anxiety and Depression

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2019). The digital psychiatrist: In-search of evidence-based apps for anxiety and depression. *Frontiers in Psychiatry, 10*, Article 831. https://doi.org/10.3389/fpsyt.2019.00831

Abstract

One of the biggest growth areas in e-mental health resources has been the development and use of mobile mental health apps for smartphones and tablet devices. Such apps are being downloaded at increasing rates, but there have been questions about their efficacy and the research methodologies used to examine this. A review of the major app marketplaces, the Apple App Store and Google Play store, was conducted to locate apps claiming to offer a therapeutic treatment for depression and/or anxiety, and have research evidence for their effectiveness, according to their app store descriptions. App store descriptions were also analyzed to determine whether the app had been developed with mental health expert input; whether they had been developed in association with a government body, academic institution, or medical facility; and, whether or not they were free to download. Overall, 3.41% of apps had research to justify their claims of effectiveness, with the majority of that research undertaken by those involved in the development of the app. Other results indicated that 30.38% of shortlisted apps claimed to have expert development input; 20.48% had an affiliation with a government body, academic institution, or medical facility; and, 74.06% were free to download. Future research must consider other methodologies that may facilitate

more research being completed on a greater number of apps, and future development needs to incorporate greater levels of input by mental health experts. Ways in which app stores could play a key role in encouraging more scientific research into the effectiveness of the mental health apps they sell are discussed.

The Digital Psychiatrist:
In Search of Evidence-Based Apps for Anxiety and Depression

Introduction

Smartphones and Apps

Smartphones are mobile/cellular phones capable of connecting to the Internet and performing similar functions to computers. Smartphone applications (apps) are software programs designed to perform specific tasks on smartphones. Through the use of apps, smartphones can perform many functions, including monitoring, assessing, and treating physical and mental health conditions. On a global scale, there are over 5.2 billion people who own a smartphone (Barboutov et al., 2017).

The proliferation of smartphone apps has led to digital solutions for a myriad of situations. The largest app marketplaces are the Apple App store and Google Play store. Developers pay to have their apps available for download from these stores, which receive commissions from downloaded apps. Apps appear in searches based on search terms, but information about the complex algorithms used to determine what apps appear higher in a search than others is not publicly known. In 2018, global expenditure on downloaded apps was approximately $US92.1 billion (newzoo, 2018), highlighting the lucrative business of being an app marketplace.

Apps Versus Online Interactive Digital Platforms

A significant advantage for apps is that for many, once it is downloaded to one's smartphone or tablet device, it can be used "offline", or without needing to connect to the Internet. There is significant progress being made in the area of online interactive digital platforms that can be accessed via websites. These services are often "app look-a-likes". That is, one can download them as an app, but the app "points" to a website that can provide a

greater range of interactive tools, including the possibility of interacting with a therapist. As an example, a popular Australian mental health app is *This Way Up* (https://thiswayup.org.au/), but once downloaded, it requires an Internet connection to connect to a more powerful online platform with greater interactive features than is feasible to include in an offline, standalone app. The significant advantages for online interactive platforms are that greater security features can be applied and the amount of content and interactive potential is far greater than an offline, standalone app. These are features that are particularly attractive to clinicians who may wish to use such tools in clinical settings. Nevertheless, there is an attraction for many people who would like to use such apps without needing to connect to the Internet each time they open the app, especially in situations where connectivity to the Internet is unreliable (such as in remote communities of First Nations peoples) and in cases where the cost of regular Internet connectivity is prohibitive (such as with low socioeconomic and/or minority populations that may include First Nations peoples and culturally and linguistically diverse communities).

Health and Mental Health Apps

Health apps (including those related to mental health) are one of the fastest growing categories of apps (Khalaf, 2016). Over two-thirds of American adults are willing to use their smartphone to help manage their health (Makovsky, 2015), and about 60% of people who own a smartphone have downloaded at least one health app (Carras et al., 2014). There are approximately 318,000 health apps currently available (IQVIA Institute, 2017). Of these, more than 10,000 relate to mental health (Torous, Firth, et al., 2018).

Research has shown that people with mental illness increasingly own smartphones and other mobile devices, and are interested in using these to monitor their mental health (Carras et al., 2014; Pung et al., 2018; Torous & Powell, 2015). However, there is a large gap between interest in and use of such apps (Torous, Wisniewski, et al., 2018), and variability in

knowledge about the capabilities of existing mental health apps (Carpenter-Song et al., 2018). Further, people living with a mental illness can have distinctly negative attitudes towards an app's ability to manage sensitive information associated with mental health (Hendrikoff et al., 2019). These findings indicate that consumers' attitudes and responses to mental health apps are diverse.

Currently, the main ways individuals choose mental health apps are via ratings and reviews in the app stores (Huang & Bashir, 2017), or through comments made through social media or word of mouth (Rubanovich et al., 2017). However, price is also important, showing a negative correlation with downloads, and that lower priced mental health apps have consistently higher ratings than higher priced apps (Huang & Bashir, 2017). Over half of those who have downloaded a health app value ease of use over trustworthiness of the app (Makovsky, 2016), suggesting that demonstrated effectiveness is not an important consideration for many consumers.

Potential Benefits of Mental Health Apps

There are many potential benefits of using mental health apps for alleviating depression and anxiety. These include portability, immediacy and accessibility. These features may be of particular benefit to rural populations; people on waiting lists for face-to-face services; or, difficult-to-engage groups such as teenagers (Wang et al., 2018). Other advantages include affordability, convenience, and anonymity. These benefits may have specific relevance to lower socioeconomic groups who find traditional treatment cost-prohibitive (Smith, 2015), or those who fear stigmatization. Mobile apps also offer convenient ways of practicing strategies learned in face-to-face therapy, and may incorporate reminders that can be set to increase treatment and medication compliance. By offering effective options to those with milder psychiatric symptoms, the burden on traditional mental

health services could be reduced (Newman et al., 2011). These potential benefits provide compelling reasons to pursue research on the efficacy of mental health apps.

The Current Study

The focus of this app store search was on apps that claimed to address depression and/or anxiety symptoms. These conditions are the most common forms of mental illness (Vigo et al., 2016), and many of their symptoms and causes overlap. While the potential benefits of apps for depression and anxiety are many, there are also possible disadvantages and dangers in using untested and potentially ill-informed apps that may risk doing harm to an individual (e.g., suggesting an individual undertake an activity that is not evidence-based that may have adverse consequences such as drinking alcohol to relax, or taking long-term higher doses of addictive medication to improve sleep). Presently, there is no reliable approximation of the proportion of apps for depression or anxiety that are available in the app stores which have scientific evidence of efficacy, or that have been developed with input from mental health professionals or in collaboration with a government body, academic institution, or medical facility. Further, given that individuals rate free apps more highly than those that incur a cost, and that one of the potential benefits of mental health apps is that people from lower socioeconomic backgrounds may have access to mental health treatment they might not otherwise be able to afford, it is important to know what proportion of mental health apps are available free of charge to the consumer. To address these questions, this mini-review examined apps that claim to offer a therapeutic intervention for reducing symptoms of depression or anxiety and are listed in the Apple App Store and Google Play store under one of the search terms used. The research questions were: What proportion of apps for depression and/or anxiety

1. … claim to have research evidence for their effectiveness?
2. … have involved a mental health expert in their development?

3. … have been developed in affiliation with an academic institution, medical facility, or other government-funded body?

4. … are free to download?

Methods

The methodology for this mini-review was informed by the AMSTAR (Shea et al., 2007) and PRISMA (Moher et al., 2009) protocols for systematic reviews. The coding regime used was based on Alyami et al. (2017) and Shen et al. (2015), both of whom conducted app marketplace searches for mental health apps.

Four researchers, including the lead author, searched the Apple App Store and Google Play store, in December 2018. Previous research had shown that these two marketplaces account for over 90% of available apps for depression (Shen et al., 2015).

A total of 19 key word searches were used for each marketplace across all categories: mental health, depression, anxiety, wellbeing, happiness, psychological distress, positive psychology, suicide, mental illness, CBT, cognitive behaviour therapy, cognitive behavior therapy, ACT, acceptance and commitment therapy, DBT, dialectical behaviour therapy, dialectical behavior therapy, IPT, and interpersonal therapy.

Apps were shortlisted and data were extracted based on their descriptions in the app stores (that is, apps were not downloaded individually). Data relating to organization affiliation and the involvement of mental health professionals was coded as shown in **Table 5.1**. Information about the cost of the app and whether the app claimed any research on effectiveness were also noted from descriptions in the app store. Researchers then located all the claimed research, verifying its existence, through searches made in Google Scholar. To be consistent with how a consumer would go about finding an app, the supporting literature was ***not*** identified first.

Table 5.1

Coding Used for App Store Search

Variable	Code	Description
Organizational affiliation	UNI	University. Produced in association with a university or other academic institution.
	MEDC	Medical Center. Produced in association with a medical institution or hospital.
	GOVT	Government. Produced in affiliation with a government institution.
	INST	Institution. An explicit association (i.e., foundation, center, non-government organization, etc.).
	OTHER	Other. There is a clear but unclassifiable affiliation (e.g., a company), but not a ".com".
	INSUFF	Insufficient. The affiliation cannot be confirmed by available information.
Content source	EXP	Expert. Developed by / with an accredited medical or allied health professional, or a recognized institution.
	EXT	External source. From a specific external source (e.g., DSM, recognized inventory, association etc.), but not based on or inspired by a theory / practice (e.g., CBT).
	LAY	Layperson. Source identified but no credential mentioned. Non-medical expertise clearly indicated by detailed bio or qualifier (e.g., years of experience).
	PLE	Person lived experience. Indication that the app is developed by a person with lived experience.
	INSUFF	Insufficient. No direct information provided about origin of information.

Inclusion criteria were:

1. The app language is English;

2. Reported research is in English and published in a peer-reviewed journal;

3. Supporting research is an intervention study with outcomes measuring depression and/or anxiety symptoms;

4. The app offers a therapeutic treatment, not just symptom monitoring, thought recording, or diagnostic tools (although apps could have any of these as part of a therapeutic treatment). A therapeutic treatment was viewed as one that takes a comprehensive and broad approach to treating anxiety and/or depression;

5. The research did not examine the app as an adjunct to other types of therapy, such as receiving therapist support;

An example of statements from a description in the app store that led the app to be identified as one that offered a therapeutic treatment for anxiety and/or depression (for Destressify):

"… skills for dealing with thoughts, emotions, beliefs that induce stress or anxiety …"

"… this comprehensive program …"

"… the core plan consists of 14 key practices …"

"… meditations or mindfulness exercises … schedule these practices and get reminders …"

"… latest neuroscientific research … neural pathways of the brain rewire themselves …"

An example of statements from a description in the app store that led the app to not be included as one that offered a therapeutic treatment for anxiety and/or depression (for Anxiety Panic Attacks Game – the 5th highest ranked app after typing in "anxiety" to the Apple App Store's search box): "Play now the #1 addictive game! Be careful not to pass through the obstacle, or you'll have to start again. When you need relaxation, diversion or

just a moment of distraction enjoy with your hero. Pass through different obstacles, collect stars and buy new characters!"

Apps that were available in both app stores were only counted once.

Results

The app marketplace search resulted in 293 apps shortlisted for closer inspection based on their app store description of offering a therapeutic treatment for reducing depression and/or anxiety symptoms.

Examination of each shortlisted app's description against the inclusion criteria identified ten apps with evidence of a research base for efficacy (see **Table 5.2**), representing 3.41% of the total number of shortlisted apps across both app marketplaces. When analyzing each app store separately, differences between the two were negligible: The Apple App Store had 3.05% of its depression and anxiety apps having published research evidence and Google Play store had 4.17%. Of the ten apps that had research support, in only three cases (1.02% of the shortlisted apps) was the research ***independent*** (i.e., conducted by an institution or individuals that were not involved in the app's development and/or who would not benefit financially from the app).

The published research articles varied in design and methodology. Seven were randomized controlled trials (RCTs) and two were feasibility / pilot studies without control groups. A total of 1017 participants were in intervention groups, and 447 were in control groups. The mean age across all participants was 32.7 years. Intervention periods varied between two weeks and 12 weeks, with the longest period of follow-up being three months.

Table 5.2

Research Evidence for the Effectiveness of Apps That Offer a Therapeutic Treatment for Anxiety and/or Depression

Article	App name	Design	Participants	Outcome	Time / Length	Statistical analysis
Lee and Jung (2018) *	Destressify	RCT	77 participants in intervention group, 86 in waitlist control group; mean age 20.6 years	Stress, anxiety and depression	Intervention period 4 weeks; no long-term follow-up	ANCOVA/MANCOVA; detailed in article
Christoforou et al. (2017)	Agoraphobia Free	RCT	73 participants in intervention group, 69 in other app "control" group (no waitlist group), mean age 39.7 years	Anxiety (agoraphobia)	Intervention period 12 weeks; no long-term follow-up	Linear mixed model; detailed in article
Kinderman et al. (2016)	Catch It	Feasibility / pilot study	285 participants (no control group); mean age 48.0 years	Positive / negative mood	Intervention period 6 weeks maximum, but varied amongst participants; no long-term follow-up	ANOVA; detailed in article
Carey et al. (2016)	Mindsurf	Feasibility / pilot study	23 participants (no control group); no mean age given	Anxiety, depression	Intervention period 2 weeks; no long-term follow-up	ANOVA; inadequate detail in article
Kuhn et al. (2017)	PTSD Coach	RCT	62 participants in intervention group, 58 in waitlist control group; mean age 39.0 years	Anxiety (posttraumatic stress disorder), depression	Intervention period 3 months; 3 months long-term follow-up	Repeated measures ANOVAs; detailed in article
Bakker et al. (2018a)	MoodMission	RCT	56 participants in intervention group 1, 56 in intervention group 2, 50 in intervention group 3, 64 in waitlist control group; mean age 34.0 years	Anxiety, depression	Intervention period 30 days; no long-term follow-up	ANOVA; detailed in article
Roepke et al. (2015)	SuperBetter	RCT	93 participants in intervention group 1, 97 in intervention group 2, 93 in waitlist control group; mean age 40.2 years	Anxiety, depression, life satisfaction	Intervention period 4 weeks; 6 weeks long-term follow-up	Hierarchical Linear Modeling; detailed in article
Stiles-Shields et al. (2019)	Thought Challenger	RCT	10 participants in intervention group 1, 10 in intervention group 2, 10 in	Depression	Intervention period 6 weeks; 4 weeks long-term follow-up	Repeated measures ANOVA; detailed in article

			waitlist control group; no mean age given			
Flett et al. (2019) *	Smiling Mind	RCT	58 participants in intervention group, 67 in placebo control group; mean age 20.1 years	Stress, anxiety, depression, flourishing	Intervention period 40 days; no long-term follow-up	Multiple regression; detailed in article
Flett et al. (2019) *	Headspace	RCT	67 participants in intervention group, 67 in placebo control group (same control group as above study); mean age 20.1 years	Stress, anxiety, depression, flourishing	Intervention period 40 days; no long-term follow-up	Multiple regression; detailed in article

* = Independent research. RCT = Randomized controlled trial. ANOVA = Analysis of variance. ANCOVA = Analysis of covariance. MANCOVA = Multivariate analysis of co-variance.

In contrast, 30.38% (89/293) of apps claimed to have developmental input from a mental health expert and 20.48% (60/293) claimed to be affiliated with an organization, institution, government body, medical center, or university.

Finally, it was found that 74.06% (217/293) of apps offering a therapeutic treatment for anxiety and/or depression were free to download.

Discussion

A search of the Apple App Store and Google Play store revealed a total of 293 apps claiming to offer a therapeutic treatment for depression and/or anxiety. Of those 293 apps, only 3.41% had published research on their effectiveness. One other app store review has previously examined what proportion of mental health apps had published evidence, but focused only on social anxiety apps, and failed to find any such apps with published evidence (Alyami et al., 2017). This is the first app store review to examine the proportion of depression and anxiety apps in general that have published research corroborating claims of efficacy.

This mini-review found that the small amount of research to date has been completed mainly by individuals and organizations who have an affiliation with the app; either through its development, or being on its staff or board, or otherwise being in a position to financially benefit when the app is sold. Only 1.02% of the 293 apps had support shown in ***independent*** research. With so little independent research supporting favorable evaluations, questions of researcher bias and conflicts of interest inevitably arise.

The nine published research articles (examining the ten apps that were first found in the app stores) are of varying quality. Most do not have long-term follow-up data, with one study having a three-month follow-up, and two further studies having 4-week and 6-week follow-up

data. This reveals that even in this small sample of research, taken from a large number of apps found in the first instance, there is very little evidence to date that apps for anxiety and depression can have positive long-term effects on their users. Furthermore, the intervention periods vary enormously between studies, providing little in the way of guidance about optimal dosage/usage. This issue is reinforced by the non-existence of any replication studies that might manipulate dosage/usage.

Less than one third (30.38%) of apps claimed to have development input from a mental health expert, meaning that over two-thirds of apps for treating depression and/or anxiety were developed ***without*** any professional input. Previous research has produced similar findings of professional input of 38.3% (Shen et al., 2015) and 34.21% (Alyami et al., 2017) for depression and social anxiety apps respectively. This indicates that a large segment of mental health apps for depression and anxiety are being developed by individuals who have no mental health training or connection with a relevant organization. It also raises questions about who some of these apps are aimed it, and what type of audience is the target.

Only 20.48% of the apps were affiliated with an organization, university, government body, or medical center. Previous research found similarly low percentages of 35.0% (Shen et al., 2015) and 7.9% (Alyami et al., 2017). While it is concerning that such a small number of apps have an institutional affiliation, it is equally concerning that the overall number of apps with published research evidence is not closer to this 20.48% figure.

In terms of cost, 74.06% of apps for depression and/or anxiety were free to download. Previous research found that 37.4% of depression apps (Shen et al., 2015), and 52.63% of social anxiety apps were free (Alyami et al., 2017). This suggests that the proportion of free depression

and anxiety-related apps may be increasing, which is positive for people where price-point is critical, but only beneficial if they provide valid treatment.

Resources for Clinicians

In the absence of an adequate level of research, there are a number of places that clinicians can turn to for assistance in assessing the suitability of specific mental health apps. Firstly, efforts are being made by governments around the world to regulate mental health apps, primarily on the basis of risk of harm to the user. This is occurring in countries such as: the United States (https://www.fda.gov/medical-devices/digital-health/mobile-medical-applications), the United Kingdom (https://www.gov.uk/government/publications/health-app-assessment-criteria/criteria-for-health-app-assessment), Canada (https://www.canada.ca/en/health-canada/services/drugs-health-products/medical-devices/activities/announcements/notice-digital-health-technologies.html), New Zealand (https://www.health.govt.nz/our-work/digital-health/digital-health-strategic-framework), and Australia (https://www.safetyandquality.gov.au/our-work/safety-in-e-health/certification-framework-for-digital-mental-health-services/). Secondly, there are reputable websites offering information and reviews for both clinicians and consumers about mental health apps, including information on published evidence if applicable, such as: ***PsyberGuide*** (https://psyberguide.org/); the ***NHS Mental Health Apps Library*** (https://www.nhs.uk/apps-library/category/mental-health/); ***Head To Health*** (https://headtohealth.gov.au/); ***reachout.com*** (https://au.reachout.com/tools-and-apps); ***beacon*** (https://beacon.anu.edu.au/); and ***Health Navigator*** (https://www.healthnavigator.org.nz/). Thirdly, there are at least 17 frameworks worldwide for evaluating mental health apps (Nielsen & Rimpilainen, 2018) that can be used by clinicians,

including the American Psychiatric Association's App Evaluation Model (https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model).

Limitations

This mini-review has limitations. First, there are challenges to conducting a search of the Apple App Store and Google Play, with differences in the way each delivers search results. The algorithms used in each case are unknown and remain the product of corporate intellectual property. Search results can differ on different days, and it is impossible to explain how one app can appear earlier in a search compared to another similar app. The app store searches are also limited by their minimal search options. For example, there is no filter by date, or developer, or other options as are available when searching literature using a standard database. However, analyzing the app stores remains an important exercise because this is how most people currently find mental health apps to download. While it was a deliberate methodological decision not to conduct a literature search first, and instead to search the app stores and then search the literature based on what app store descriptions revealed, it should be noted that an independent search of the literature would likely have produced more apps that met criteria as a "therapeutic treatment" for anxiety and/or depression that were not found in the app store search.

A further limitation of this study is that some of the terms used in our searches may not necessarily reflect words that consumers may use. For example, consumers may be more likely to use terms like "stress", "depressed mood", or "worry", instead of terms like "dialectical behavior therapy" or "interpersonal therapy". A third limitation of this research is that results were based on information contained in the app store description. That is, none of the apps were downloaded to ensure their store description actually matched the app's content. Furthermore, if no research is highlighted in the app store description, this does not necessarily mean that

research does not exist. This is similarly true for the issues of level of expert input and association with a relevant organization. To confirm exact numbers of apps with these elements, every app listed in each app store search would also have to be put through a literature search, which was outside the scope of this study.

Conclusion

The proliferation of health apps, and specifically mental health apps, on the app marketplaces, such as the Apple App Store and Google Play store, has occurred without an equivalent proliferation of scientific evidence for their effectiveness. Mental health clinicians who are trained in the scientist-practitioner model of using evidence-based practice may be reluctant to use these new tools in their normal workflow, and may hesitate to recommend them to patients (Sinclair et al., 2013). Patients are usually ready to listen to advice from their mental health treatment providers (Pung et al., 2018), and if the evidence base for mental health apps is widened, clinicians may be more willing to recommend them (Neary & Schueller, 2018).

The other players in a position to improve the overall situation in establishing the effectiveness of a mental health app are the app marketplaces themselves. If Apple and Google were willing to re-categorize mental health apps, and perhaps health apps generally, in a way that recognized when an app had achieved an acceptable standard of effectiveness, this would allow consumers to more easily distinguish between reliable and untested apps. From a financial perspective, Apple and Google (together with the app developers) have something to gain on this front: potentially more downloads of scientifically-tested apps because of their proven effectiveness.

Future development of mental health apps must incorporate more involvement from clinicians and institutions engaged in mental health research, training or service provision. Governments may also be able to offer regulatory oversight and certification of health apps with adequate

scientific evidence for their effectiveness. Until research on the efficacy of mental health apps is broadened, and ways of searching for and recognizing effective mental health apps is improved, questions will continue to be asked about the place of these digital tools in managing mental health conditions.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

Candidate

15/12/2020
Date

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate 15/12/2020 Date

[redacted]

Principal Supervisor 15/12/2020 Date

Research Progression to Chapter 5 - Study 2B

In Chapter 5 – Study 2A, a systematic search of the major app stores was conducted. This gave a clearer indication of how few publicly available mental health apps have research evidence to confirm their efficacy and/or effectiveness, particularly in relation to independent research and the non-existence of any replication studies. It found that few mental health apps have been developed with expert or government input, but, pleasingly, a large proportion of mental health apps were free to download. A new protocol for conducting app store searches was also developed, as no such protocol could be located in the literature that can be used in the same manner as the PRISMA (Moher et al., 2009) and AMSTAR (Shea et al., 2007) protocols can be used to guide literature reviews. This new protocol is described in more detail in Study 2B, where the issue of mental health apps using specific evidence-based frameworks is explored, including determining the proportion of available mental health apps that have an evidence-based framework, and the proportion of which types of frameworks are being used.

Chapter 5 – Study 2B: Apps with Maps: A Systematic Review of the Major App Stores for Anxiety and Depression Mobile Apps with Evidence-Based Frameworks

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Apps with maps – Anxiety and depression mobile apps with evidence-based frameworks: Systematic search of major app stores. *JMIR Mental Health, 7*(6), Article e16525. https://doi.org/10.2196/16525

Abstract

Background: Mobile mental health applications (apps) have become ubiquitous tools to assist people to manage symptoms of anxiety and depression. However, due to the lack of research and expert input that has accompanied the development of most apps, concerns have been raised by clinicians, researchers and government authorities about their efficacy.

Objective: This review aimed to estimate the proportion of mental health apps offering comprehensive therapeutic treatments for anxiety and/or depression available in the app stores that have been developed using evidence-based frameworks. It also aimed to estimate the proportions of specific frameworks being used in an effort to understand which frameworks are having the most influence on app developers in this area.

Methods: A systematic review of the Apple App Store and Google Play store was performed to identify apps offering comprehensive therapeutic interventions that targeted anxiety and/or depression. The PRISMA checklist was adapted to guide this approach.

Results: Of the 293 apps shortlisted as offering a therapeutic treatment for anxiety and/or depression, 162 (55.29%) mentioned an evidence-based framework in their app store descriptions. Of these, 88 (30.03%) claimed to use cognitive behavior therapy techniques, 46

(15.70%) used mindfulness, 27 (9.22%) positive psychology, 10 (3.41%) dialectical behavior therapy, 5 (1.71%) acceptance and commitment therapy, and 20 (6.83%) used other techniques. Of the 162 apps that claimed to use a theoretical framework, only 10 had published evidence for their efficacy.

Conclusions: The current proportion of apps with evidence-based frameworks is unacceptably low, and those without tested frameworks may be ineffective, or worse, provide a risk of harm to users. Future research should establish what other factors work in conjunction with evidence-based frameworks to produce efficacious mental health apps.

Apps with Maps: A Systematic Review of the Major App Stores for Anxiety and Depression Mobile Apps with Evidence-Based Frameworks

Introduction

Background

The practice of psychotherapy is underpinned by evidence-based therapies and interventions. These techniques are used when they are proven to be efficacious via thorough experimental methods. In most developed countries, mental health clinicians are discouraged by professional associations and government regulations from using therapies without such evidence.

New technology is changing the way therapy can be delivered. Therapies previously only delivered face-to-face can now be accessed electronically, and the way that many people do this is via their smartphone. Smartphone apps are software programs that can be downloaded from specialist websites, known as app stores. The two biggest app stores are the Apple App Store and Google Play. Global expenditure on apps in 2018 was approximately US $92.1 billion (newzoo, 2018), which attests to the widespread use of apps and value of smartphones to consumers.

Mental Health: Is There an App for That?

Health, including mental health, apps are emerging as one of the most important categories of apps. A total of 75% of consumers believe that technology is important in managing their health, and 88% are willing to share mobile health (mHealth) data with their health care provider (Accenture, 2018). Approximately 48% of consumers are currently using a health app (Accenture, 2018), and many more have downloaded at least one (Carras et al., 2014). Health apps are in the top-10 list of worldwide categories for consumer spending on apps (App

Annie, 2019). Although definitions about what constitutes a *health* app varies, and this influences the estimates of how many health apps are available, a 2017 report calculated that there were approximately 318,000 health apps available for download (IQVIA Institute, 2017), with 10,000 of these relating to mental health (Torous, Firth, et al., 2018).

People with mental illness are attracted to the possibility of utilizing apps to manage their mental health (Torous & Powell, 2015); however, there is doubt in this population about how apps manage confidential information (Hendrikoff et al., 2019). There is also uncertainty about the available functions of mental health apps and which apps are the most useful (Carpenter-Song et al., 2018). For guidance about what app to download, mental health consumers rely mainly on ratings and reviews in app stores (Huang & Bashir, 2017) or on advice given through social media or word of mouth (Rubanovich et al., 2017). However, if a mental health expert is not involved in the development of an app, and if it has not been developed using an evidence-based theoretical framework, the app may not be effective or, worse, may do harm to its user.

There are many reasons why research on the efficacy of apps for anxiety and/or depression should be prioritized. Firstly, many people are already using mental health apps (Torous, Firth, et al., 2018), and with five billion people around the world currently using a smartphone (Statista, 2019b), there is potential for many more to do so in the future as this figure rises. Apps provide greater accessibility to mental health resources and offer instant mental health assistance. These features are especially relevant for traditionally difficult-to-reach cohorts, such as teenagers (Wang et al., 2018) and people in rural communities (Jones & Moffitt, 2016). Mental health apps may also offer more cost-effective options for lower socioeconomic groups (Smith, 2015) and greater anonymity and flexibility for others. For example, using an app may be more convenient for consumers with limited time to access other therapy, and more

convenient for clinicians who may prefer their clients and patients to do homework activities on their smartphone so that results can be digitally sent back to the clinician. Certain clients in therapy may also just prefer the novel value of doing interactive homework activities on their phone, rather than with a pen and paper. For many people, being able to use their phone for such activities is simply more convenient (eg, it may be easier for them to use a smartphone to do homework while they are travelling on a bus rather than handling awkward pieces of paper in such situations). In this way, apps may be used as adjuncts to other therapies. Apps can also be used in more practical ways, such as setting reminders to assist with treatment and medication compliance. All of these factors create the potential for apps to reduce the burden on existing mental health services (Newman et al., 2011).

Apps and Evidence-Based Frameworks

For a mental health app to be efficacious, it is fundamental that it be underpinned by an evidence-based framework; that is, an established therapy technique for reducing psychological distress, such as cognitive behavioral therapy (CBT). An evidence-based framework provides a road map or blueprint for an app's functions and performance. However, even when an app claims to offer CBT, the contents and functioning of the app may not align with CBT principles (Stawarz et al., 2018), especially if the app has been developed without expert input.

To date, there has been little research that addresses the potential harm a mental health app may do to an individual. It seems conceivable, however, that an app developed without a proven theoretical framework and method of intervention could have the potential to do harm to a user (Moshi et al., 2018). For instance, a recommendation to use herbal supplements could adversely interact with prescribed medication.

Over the last two decades, e-mental health programs with a CBT framework and designed for use on computers have been found to be effective for both adults and children in reducing anxiety and depression (Ebert et al., 2015; Richards & Richardson, 2012) and improving happiness and well-being (Davies et al., 2014; Manicavasagar et al., 2014; Powell et al., 2013). Some studies have compared the effectiveness of e-mental health programs with face-to-face therapy and have found comparable (Andrews et al., 2010; Richardson et al., 2010), and, in some cases, even more favorable, results (Merry et al., 2012). Apps for smartphones and tablet devices can perform a range of tasks, including using physiological and other data automatically gathered by their device (Beiwinkel et al., 2016). Researchers are also looking at how using incidental mobile phone data, such as phone calls made and amount of time the screen is "on", can help predict an individual's mental health diagnosis (Faurholt-Jepson et al., 2019). There may yet be further ways that smartphones can reveal important information about our mental health, and thus potentially lead to more effective treatments (Hidalgo-Mazzei & Young, 2019). Similarly, factors such as usability / ease of use, aesthetics, and level of gamification may also impact on the success of a mental health app (Rickard et al., 2016). Although the content of an app may be similar to an older web-based program designed for use on a desktop computer, it is still unknown if the delivery of the program by smartphone changes the experience of that content. Intuitively, the content should have similar effects, but research has yet to determine this. For example, does the size of the smartphone screen as opposed to the size of a desktop computer screen have any impact on effectiveness? Does the ability for a smartphone app to be accessed in a park differ to the effectiveness of a web-based program on a desktop computer in an office or a bedroom? There is still much to study in this area. The shortcomings of mental health apps include suboptimal use of the available technological features of smartphones (Frank

et al., 2018; Hendrikoff et al., 2019; Shah et al., 2018). There is now an interest in how stand-alone apps for smartphones (ie, programs that can be downloaded and used without an internet connection) based on CBT or other evidence-based theoretical frameworks may further enhance a clinician's digital toolbox (Torous et al., 2017).

CBT proposes that emotions can be more effectively managed by adjusting thinking and behavior and recognizing physiological responses (Beck et al., 1979). CBT has an established evidence base and history of successfully treating anxiety and depression, both in individual therapy (Butler et al., 2006) and in group settings (McDermut et al., 2001). The CBT framework offers a broad range of efficacious interventions (Bennett-Levy et al., 2010), including the following: the use of *thought diaries* to challenge negative thinking; coming up with personally meaningful affirmations; different types of physical exercise, such as walking and dancing; increasing social connectedness and having personally meaningful social interaction; increasing the amount of time doing pleasurable, or once pleasurable, activities; structured problem solving; and others. Many of these CBT interventions have been incorporated into mental health apps.

Other therapies that have been developed from traditional CBT have also been used successfully to treat anxiety and depression. Positive psychology, for instance, has become a more recent framework that is, under a strict definition, a type of CBT but has a different focus. Rather than simply *fixing* psychopathology, positive psychology aims to help individuals reach their full well-being potential by increasing optimism and happiness (Seligman & Csikszentmihalyi, 2000). Studies on various positive psychology interventions have revealed efficacious empirical support for reducing symptoms of anxiety and depression as well as increasing well-being and happiness. For example, Freedman and Enright (1996) showed that forgiveness significantly decreased depression, Froh et al. (2008) implemented a gratitude

intervention that resulted in significantly greater optimism and life satisfaction as well as decreased negative affect, Lyubomirsky et al (2005) reported an increase in happiness and well-being for participants who completed acts of kindness, and Seligman et al (2005) successfully used three strategies to increase happiness and decrease depressive symptoms: (1) using one's signature strengths in new ways, (2) writing down three positive things each day, and (3) writing a letter of gratitude. All of these interventions could conceivably be incorporated into mental health apps and most already have.

Similarly, other frameworks, often referred to as third-wave therapies (Hayes & Hofmann, 2017), have offered up interventions that are suitable for use in mental health apps. Various mindfulness interventions have proved efficacious in the past (Hofmann et al., 2010), and these meditation and breathing activities are easily incorporated into an app. Dialectical behavior therapy (DBT) is an intersection of various interventions, including mindfulness, distress tolerance, emotional regulation, and improving social relatedness activities. These may involve dialectical interaction with a therapist, including recording distress ratings, among other interventions that have proven effective in reducing symptoms of anxiety and/or depression, with particular focus on individuals with a diagnosis of borderline personality disorder (Lynch et al., 2006). Interpersonal therapy (IPT) involves interventions designed to improve current interpersonal relationships and social dysfunction, rather than focusing on personality, as an efficacious way of improving symptoms of depression (de Mello et al., 2005). Acceptance and commitment therapy (ACT) focuses on an individual accepting distressing thoughts and performing committed actions guided by one's values; ACT has proven efficacious for both anxiety (Vollestad et al., 2012) and depression (A-Tjak et al., 2015). All the frameworks listed here potentially comprise interventions that may conceivably be incorporated into an app.

Objectives of This Study

This study involved a systematic search of app stores and concentrated on apps that offered a comprehensive therapeutic treatment for anxiety and/or depression, as opposed to apps that may offer singular or novel interventions (see *inclusion criteria* in the Methods section for more details about this definition). The research questions were as follows: (1) What proportion of publicly available apps offering a therapeutic treatment for anxiety and/or depression have used an evidence-based framework in their development, and (2) In an effort to understand which specific frameworks are having the most influence on mental health app developers, what are the proportions of specific frameworks? No previous study or review could be located that examined this issue by focusing on publicly available apps listed in the app stores.

Methods

This systematic review used the AMSTAR (A MeaSurement Tool to Assess systematic Reviews) 2 (Shea et al., 2017) and PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Moher et al., 2009) protocols for guidance in conducting the app store search. Although these protocols were developed specifically for literature searches, they offer appropriate direction that can be applied to searching app stores in the absence of such a specific guide. However, there are limitations in their use for guidance in app store searches. We adapted both of these protocols into a unique protocol and checklist specifically for app store searches: the Protocol for App Store Systematic Reviews (PASSR) checklist is a combination and reworking of the items from AMSTAR 2 and PRISMA that can be applied to the systematic search of app stores for any category of apps. The same wording as used in both AMSTAR 2 and PRISMA has been retained wherever possible to enable ease of comparison (see Document B.1 in Appendix B).

Four researchers, including the lead author (JMM), systematically searched the Apple App Store and Google Play store in December 2018, and then again in July 2019, to ensure newly available apps were included for review. These two marketplaces attract more than three times as much revenue from downloads as their next competitor, the Windows Store (Statista, 2019a), and over 90% of available apps for depression can be found in either of these stores (Shen et al., 2015).

Searches in each marketplace were made with the following 19 keywords across all categories: mental health, depression, anxiety, wellbeing, happiness, psychological distress, positive psychology, suicide, mental illness, CBT, cognitive behaviour therapy, cognitive behavior therapy, ACT, acceptance and commitment therapy, DBT, dialectical behaviour therapy, dialectical behavior therapy, IPT, and interpersonal therapy.

Apps were shortlisted based on their app store descriptions, and apps that were available in both stores were only counted once.

Inclusion criteria were as follows:

1. The app language and store description are in English.
2. The app offers a therapeutic treatment for anxiety and/or depression, not just singular elements of a therapy, such as monitoring symptoms, recording thoughts, or diagnosing the disorder, although apps could have any of these as part of a therapeutic treatment. In this way, apps were excluded if they used only singular elements. This approach is in contrast to other similar reviews (Torous et al., 2017) that have identified apps with single tools rather than comprehensive treatments. Therapeutic treatment was defined for this study's purpose as "offering focused treatment suggestions specifically for reducing

symptoms of anxiety and/or depression in a manner comprehensive enough to be considered a type of therapy."

Results

A search of the Apple App Store and Google Play store uncovered a shortlist of 293 apps whose app store descriptions inferred that they offered a therapeutic treatment for anxiety and/or depression; the full list of these apps is available from the corresponding author (JMM). Of these 293 apps, a total of 162 (55.3%) claimed to have an evidence-based theoretical framework informing the app's development. Differences between the Apple App Store (112/197; 56.85%) and the Google Play store (50/96; 52.08%) were negligible.

The evidence-based frameworks found in the apps represented the following proportions: CBT, 30.0% (88/293); mindfulness, 15.7% (46/293); positive psychology, 9.2% (27/293); DBT, 3.4% (10/293); ACT, 1.7% (5/293); and others, 6.8% (20/293). Note, some apps claimed to use multiple frameworks—each time a framework was mentioned, it was counted.

When including only the 162 apps with a theoretical framework in the analysis, the breakdown was as follows: CBT, 54.3% (88/162); mindfulness, 28.4% (46/162); positive psychology, 16.7% (27/162); DBT, 6.2% (10/162); ACT, 3.1% (5/162); and others, 12.3% (20/162).

Of those 162 apps with evidence-based frameworks, 10 (6.2%) were found to have published evidence for their effectiveness (see Table 5.3). The nine research articles that examined the 10 apps were made up of seven randomized controlled trials (RCTs)—the Flett et al. (2019) article contained one RCT with 2 of the listed apps: *Headspace* and *Smiling Mind*—and two feasibility or pilot studies: Kinderman et al. (2016) and Carey et al. (2016). Total participant numbers were comprised of 1017 in intervention conditions and 447 in control

groups. The mean age across those studies that provided enough information was 32.7 years (SD=9.3). Intervention phases fluctuated between 2 and 12 weeks. Three studies reported long-term follow-up, the longest period being 3 months (see Table 5.4 for a summary of research characteristics).

Table 5.3

Coding Used for App Store Search, and the Search Results

Code	Description (ie, framework)	Apps with this framework (N=293), n (%)	Apps with published research, n	Name of the app (or apps) with published research
CBT	Cognitive behavioral therapy	88 (30.0)	5	Agoraphobia Free Catch It PTSD Coach MoodMission Thought Challenger
MIND	Mindfulness	46 (15.7)	3	Headspace Smiling Mind Destressify
POS	Positive psychology	27 (9.2)	1	SuperBetter
DBT	Dialectical behavior therapy	10 (3.4)	0	N/A[a]
ACT	Acceptance and commitment therapy	5 (1.7)	0	N/A
OTH	Other recognized framework	20 (6.8)	1	MindSurf
NONE	No theoretical framework	131 (44.7)	0	N/A

[a]N/A: not applicable.

Table 5.4

Summary of Published Research for Shortlisted Apps

App name	Reference	Sample characteristics	Intervention period; long-term follow-up period	Statistically significant improvements	Outcome measure used
Agoraphobia Free	Christoforou et al. (2017)	Intervention group: n=73 Control group: n=69 Mean age: 39.7 years (SD=11.3)	12 weeks; no follow-up	Anxiety	PAS[a]
Catch It	Kinderman et al. (2016)	Intervention group: n=285 Control group: none Mean age: 48.2 years (SD not given)	6 weeks (but varied among participants); no follow-up	*Positive* and *negative* mood	None[b]
PTSD Coach	Kuhn et al. (2017)	Intervention group: n=62 Control group: n=58 Mean age: 39.3 years (SD=14.6)	3 months; 3 months	Anxiety and depression	PCL-C[c] and PHQ-9[d]
MoodMission	Bakker et al. (2018a)	Intervention group (a): n=56 Intervention group (b): n=56 Intervention group (c): n=50 Control group: n=64 Mean age: 34.2 years (SD=12.1)	30 days; no follow-up	Depression	PHQ-9
Thought Challenger	Stiles-Shields et al. (2019)	Intervention group (a): n=10 Intervention group (b): n=10 Control group: n=10	6 weeks; 4 weeks	Depression	PHQ-9

		Mean age: not given (SD not given)			
Headspace	Flett et al. (2019)	Intervention group: n=67 Control group: n=67 Mean age: 20.1 years (SD=2.8)	40 days; no follow-up	Depression	CES-D[e]
Smiling Mind	Flett et al. (2019)	Intervention group: n=58 Control group: n=67 Mean age: 20.1 years (SD=2.8)	40 days; no follow-up	Depression	CES-D
Destressify	Lee and Jung (2018)	Intervention group: n=77 Control group: n=86 Mean age: 20.6 years (SD not given)	4 weeks; no follow-up	Trait anxiety	STAI[f]
SuperBetter	Roepke et al. (2015)	Intervention group (a): n=93 Intervention group (b): n=97 Control group: n=93 Mean age: 40.2 years (SD=12.4)	4 weeks; 6 weeks	Depression	CES-D
MindSurf	Carey et al. (2016)	Intervention group: n=23 Control group: none Mean age: not given (SD not given)	2 weeks; no follow-up	Anxiety and depression (not statistically significant)	DASS-21[g]

[a]PAS: Panic and Agoraphobia Scale.
[b]Used direct data input from apps based on words used by users to describe mood.
[c]PCL-C: PTSD (posttraumatic stress disorder) Checklist – Civilian.
[d]PHQ-9: Patient Health Questionnaire – Depression Scale.
[e]CES-D: Center for Epidemiologic Studies – Depression Scale.
[f]STAI: State Trait Anxiety Inventory.
[g]DASS-21: Depression Anxiety Stress Scales – 21-Item Version.

Discussion

Principal Findings

The purpose of this study was to locate mobile mental health apps for treating anxiety and/or depression that contained recognized theoretical frameworks underpinning their development. A search of the Apple App Store and Google Play store revealed a total of 293 apps claiming to offer a therapeutic treatment for anxiety and/or depression, with just over half claiming to have been developed using an evidence-based theoretical framework. Of these, CBT was the most quoted framework in app store descriptions. The method used to identify these apps mimicked that of how a consumer in the general population would ordinarily locate an app to treat anxiety and/or depression; that is, by using the search function of an app store, then reading the description of each app.

CBT has over 60 years of research, a longer research history than any of the other theoretical frameworks contained in the shortlisted apps (Thoma et al., 2015). It is widely known as being effective for a range of conditions, including cessation of smoking (Webb et al., 2010) and other drugs (Magill & Ray, 2009), managing pain (Ehde et al., 2014), and relieving symptoms of mental ill-health across a range of psychological disorders (Thoma et al., 2015). CBT is also a term widely known across health-related settings and professions and is often used in information directed at the general community for public campaigns of ways to manage mental ill-health and stress (World Health Organization, 2017). It perhaps comes as no surprise, then, that CBT is the most widely used theoretical framework in apps for anxiety and depression. However, new research is accelerating in many of the other theoretical frameworks listed in this review (Dimidjian et al., 2016); it may be that the proportion of CBT-based apps lessens over time as a result.

The popularity of the term CBT, as well as the growing popularity of the terms used to identify the other evidence-based frameworks in this review, may be contributing to apps

being developed by nonexperts who incorrectly quote these terms in app store descriptions as a means of attempting to gain legitimacy. There are no known safeguards in place anywhere in the world to stop this from happening. While government agencies have started to regulate the health app space, this regulation has thus far focused on apps that only pose a risk of harm to users (Marshall et al., 2020b). While this is important and a welcome addition to the oversight of mental health apps, it does not provide checks on accuracy of information in app store descriptions of apps that may not fall into the category of posing a risk of harm. Such apps may provide fake or incorrect information that may, at worst, be ineffective at reducing anxiety or depression, but they may continue to be available in app marketplaces because they do not meet criteria that would identify them as posing a risk of harm to users. If an app claims to wrongly categorize its interventions as CBT, DBT, ACT, mindfulness, or positive psychology, one of the dangers is that such misuse of these terms may lead users to believe that such theoretical frameworks do not work if the user does not get any benefit from the app.

A detailed analysis of the research that accompanied 10 of the shortlisted apps is outside the scope of this paper, as this review is focused on the theoretical frameworks that underpin mental health apps. The quality and quantity of research into mental health apps has been detailed elsewhere (Alyami et al., 2017; Bakker et al., 2016; Marshall et al., 2020b; Orman & O'Dea, 2018). However, the research found to accompany the apps listed here does appear to vary greatly in methodology; this reaffirms the claims of heterogeneity made in those literature reviews. These reviews all call for more research and ongoing evaluation of research methodology into studying apps, as the current methodologies may not be the most appropriate (Broussard & Teng, 2019; Clough & Casey, 2015b; Marshall et al., 2020e).

By examining evidence-based frameworks in mental health apps, this review has highlighted the high proportion (131/293, 44.7%) of apps that claim to offer a therapeutic

treatment for anxiety and/or depression that do not rely on validated techniques. It is useful to think of an evidence-based theoretical framework as being like a map that guides clinicians in their therapeutic practices. This is at the heart of a clinician being effective in their treatment and, at the very least, not doing harm to their client or patient. While many mental health apps claim to be using an evidence-based theoretical framework, as many as 44.7% may not be; this leaves open the possibility, therefore, that a large proportion of these apps may be ineffective and possibly run the risk of doing the user harm. It is acknowledged that harm to patients and clients can be caused by various factors in the context of mental health, including delays to accessing treatment. Apps offer the potential to assist in alleviating such a cause of harm. However, if an app provides ill-conceived and non-evidence-based treatment suggestions, such potential benefits of app usage are squandered.

Limitations

This review of app stores suffers from the same limitation as other reviews of app stores: the restricted way that searches are conducted. The Apple App Store and Google Play store search results can be challenging because important information may be absent. How developers have completed online questionnaires prior to registering their app for public download, as well as the interaction of these with algorithms developed by the app store, determine the outcomes of a search. Consequently, there are differences in the order in which apps are presented when specific search terms are employed, and there are limited search functions compared to those available when performing a literature search. The results of an app store search are not necessarily presented in a logical order to users because they are unable to choose to display results according to multiple criteria, such as being able to filter from most recent to oldest, as one can do in a literature search. The outcome can be considerable ambiguity about the order of displayed results.

Another difference between a search of app stores and a search of the literature is that research may be found on a particular app, but that app may not be available for download to the general public. For example, Torous et al. (2017) discovered peer-reviewed publications on the efficacy of four CBT-based apps—excluding apps that were based on DBT or ACT—but when they searched the app stores for these apps, none could be found. While research is welcomed and encouraged in the development of mental health apps, in the end what matters most is how many of these efficacious apps are available to the general population.

Another limitation of this research is that results were based on the contents of descriptions in the app stores. None of the shortlisted apps were trialed to confirm that their description corresponded to actual content. For example, a non-expert may have developed an app and claimed in the description that it used a CBT framework, but the developer may not have incorporated any genuine CBT interventions into the app functions. If such pseudo-CBT apps exist, it is likely that they will fail to assist users to manage anxiety and/or depression and may, therefore, lead consumers to believe that CBT is ineffective (Torous et al., 2017). Furthermore, to confirm adherence to a theoretical framework, an app would have to be scrutinised closely to ensure the underlying principles of that framework were represented.

Conclusions

This review has highlighted difficulties faced by clinicians and consumers when searching the app stores for an app that offers a therapeutic treatment for anxiety and/or depression. The limited search capabilities of the app stores make it difficult to find the most appropriate app for one's needs. If an individual wants to find a mental health app based on an evidence-based framework, it is difficult to sort through the many other apps that do not have that guiding framework.

Just as mental health clinicians are trained to follow evidence-based frameworks in their practice, it is reasonable to assume that mental health apps should do the same in their functioning. This review found that little more than half do, according to their app store descriptions. Just as successful therapeutic outcomes of face-to-face therapy can be attributed to more than the theoretical framework—factors such as rapport with the therapist, therapist skills, and an individual's motivation to change—so too are there other elements of a mental health app that may contribute to a successful therapeutic outcome, such as usability and ease of use, aesthetics, level of gamification, etc (Bakker et al., 2016; Jessen et al., 2018). There would appear to be much research that still needs to be done on all these factors and how they interrelate with theoretical frameworks as well as whether certain factors mediate or moderate others.

Another suggested area of future research is to compare apps developed with evidence-based theoretical frameworks with their face-to-face equivalents. Clinicians and consumers need to know about the effectiveness and limitations of apps and where they sit alongside traditional evidence-based approaches. If clinicians and consumers become more confident in understanding how mental health apps can assist in reducing symptoms of anxiety and/or depression, it may increase the take-up of this new treatment modality and turn the potential advantages of using mental health apps into a reality.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 6

In Chapter 5 – Study 2A and Study 2B, app store searches were performed to gain an estimate of the proportion of mental health apps that have research evidence, the nature of that evidence, the proportion of apps that were developed with expert and/or government input, and the proportion and nature of mental health apps that have been developed using an evidence-based framework. When combining the results of both Study 1 and Study 2, it was clear that the quantity of research was inadequate, and that the methodology of randomised controlled trials was a deterrent to such research due to the time and cost it takes to conduct such studies. In realising that perhaps single-case designs may offer a solution, this led to the development and planning for Study 3 and Study 4. Ethics approval was obtained earlier than expected (in November 2019) and it was decided that a pilot study (Chapter 7 – Study 3) be conducted to test the planned methodology (in particular, the feasibility of having participants send daily text messages for as long as 16 weeks) ahead of commencing the main study (Chapter 10 – Study 4) in January 2020. This opportunity to conduct a pilot study occurred after the research protocol was published. The research protocol is presented in Chapter 6.

Chapter 6: Effectiveness of Using Mental Health Mobile Apps as Digital Antidepressants for Reducing Anxiety and Depression: Protocol for a Multiple Baseline Across-Individuals Design

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Effectiveness of using mental health mobile apps as digital antidepressants for reducing anxiety and depression: Protocol for a multiple baseline across-individuals design. *JMIR Research Protocols, 9*(7), Article e17159. https://doi.org/10.2196/17159

Abstract

Background: The use of mental health mobile apps to treat anxiety and depression is widespread and growing. Several reviews have found that most of these apps do not have published evidence for their effectiveness, and existing research has primarily been undertaken by individuals and institutions that have an association with the app being tested. Another reason for the lack of research is that the execution of the traditional randomized controlled trial (RCT) is time prohibitive in this profit-driven industry. Consequently, there have been calls for different methodologies to be considered. One such methodology is the single-case design, of which, to the best of our knowledge, no peer-reviewed published example with mental health apps for anxiety and/or depression could be located.

Objective: The aim of this study is to examine the effectiveness of 5 apps (ie, Destressify, MoodMission, Smiling Mind, MindShift, and SuperBetter) in reducing symptoms of anxiety and/or depression. These apps were selected because they are publicly available, free to download, and have published evidence of efficacy.

Methods: A multiple baseline across-individuals design will be employed. A total of 50 participants will be recruited (10 for each app) who will provide baseline data for 20 days. The sequential introduction of an intervention phase will commence once baseline readings have indicated stability in the measures of participants' mental health and will proceed for 10 weeks. Postintervention measurements will continue for a further 20 days. Participants will be required to provide daily subjective units of distress (SUDS) ratings via SMS text messages and will complete other measures at 5 different time points, including at 6-month follow-up. SUDS data will be examined via a time series analysis across the experimental phases. Individual analyses of outcome measures will be conducted to detect clinically significant changes in symptoms using the statistical approach proposed by Jacobson and Truax. Participants will rate their app on several domains at the end of the intervention.

Results: Participant recruitment commenced in January 2020. The postintervention phase will be completed by June 2020. Data analysis will commence after this. A write-up for publication is expected to be completed after the follow-up phase is finalized in January 2021.

Conclusions: If the apps prove to be effective as hypothesized, this will provide collateral evidence of their efficacy. It could also provide the benefits of (1) improved access to mental health services for people in rural areas, lower socioeconomic groups, and children and adolescents and (2) improved capacity to enhance face-to-face therapy through digital homework tasks that can be shared instantly with a therapist. It is also anticipated that this methodology could be used for other mental health apps to bolster the independent evidence base for this mode of treatment.

Effectiveness of Using Mental Health Mobile Apps as Digital Antidepressants for Reducing Anxiety and Depression: Protocol for a Multiple Baseline Across-Individuals Design

Introduction

Background

Mobile health apps for smartphones and tablet devices have become a lucrative business, with worldwide expenditure estimated to be over US $92 billion (newzoo, 2018). Apps are increasingly being used to monitor, assess, and improve mental health. There are now more than 10,000 publicly available mental health–specific apps (Torous, Firth, et al., 2018). Most of these apps lack published evidence for their effectiveness, making it difficult for clinicians and consumers to know which app is the most appropriate (Marshall et al., 2019). Currently, choices are made using reviews and ratings available in app stores (Huang & Bashir, 2017), but these can produce unreliable results (Xie & Zhu, 2015).

Although effective treatments for anxiety and depression exist, many people do not access these for various reasons (Gunter & Whittal, 2010). However, with ownership of smartphones being at 70% of the global population and rising (Barboutov et al., 2017), mental health apps potentially offer a partial solution to limitations in service availability and acceptability.

Previous Research

Published reviews have found that mental health apps can be effective for reducing anxiety (Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017) and depression (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017) with an overall effect size of small to moderate (Lai & Jury, 2018). Within this research, there are some notable shortcomings,

including substantial heterogeneity across studies. For example, there have been differences in dosage (Firth et al., 2018; Fleming et al., 2018) duration of interventions (Boisseau et al., 2017; Howells et al., 2016) and the absence of long-term follow-up data (Paul & Fleming, 2019).

Another limitation of previous research is that most of it has been carried out by individuals who have developed the app, who have stood to gain financially from its sales, and/or who were otherwise associated with it (Marshall et al., 2019). For instance, a recent review of app stores found that only 1.02% of mental health apps offering therapeutic treatment for anxiety and/or depression had been evaluated using *independent* research (Marshall et al., 2019). Furthermore, in a meta-analysis of 9 studies on apps targeting anxiety (Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017) and in another meta-analysis of 18 studies on apps targeting depression (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017), none involved independent research or replication (note that some studies were included in both meta-analyses).

A possible reason for the lack of research on apps is the time factor for large-scale experimental designs. Specifically, randomized controlled trials (RCTs) that demonstrate the efficacy of an intervention by measuring and comparing the outcomes of matched treatment and control groups are often lengthy to conduct. This is a barrier to achieving results in a time frame that is acceptable to the profit-driven app market. In the time it takes to complete an RCT, the app being studied may have been updated or disappeared from the market altogether as newer apps with enhanced features emerge in its place. Furthermore, RCTs are not necessarily the most appropriate study design for every situation, with another limitation of RCTs being lower ecological validity (Watson et al., 2004).

Single-Case Designs

Single-case research designs address the issue of ecological validity by testing the effectiveness of an intervention for individuals (ie, performance under real-world conditions). However, single-case designs can also control for threats to internal validity and thus test for the efficacy of a treatment. Such designs go beyond a study with a sample of 1 participant and involve continuous and repeated measurements, random assignment, sequential introduction of the treatment, and specific data analysis and statistics (Krasny-Pacini & Evans, 2018). Robust results in clinical psychology and behavioral science can be demonstrated when benefits are shown in 3 to 5 cases (Barlow et al., 2009; Horner et al., 2005; Kazdin, 2017). A single-case design is also safer than an RCT for vulnerable participants because their well-being is monitored by gathering and analyzing data more frequently during the study, and treatment can be altered if there is a clinically significant decline in status (Bentley et al., 2019; Machalicek & Horner, 2018). For mental health apps, single-case designs are a viable alternative for accelerating the evidence base (Clough & Casey, 2015a; Mehrotra & Tripathi, 2018).

Objectives and Aims

The main objective of this study is to use a single-case design to examine the effectiveness of 5 mental health apps that purport to have efficacy for reducing symptoms of anxiety and/or depression.

This study seeks to answer the following research questions: (1) Do the apps in this study provide clinically significant improvements in symptoms of anxiety and/or depression? (2) What individual characteristics of participants influence the results in this regard? and (3) What individual characteristics of the apps influence the results?

The only hypothesis to be tested in this study is based on question 1:

- Hypothesis 1: the use of apps will produce improvements in mental health and well-being in line with the 3-phase model of psychotherapy outcomes (Howard et al., 1993).

The Howard et al. (1993) model proposed that the outcome of any psychotherapeutic intervention will involve progressive reductions in subjective distress, then symptomatology, and, finally, an increase in overall life functioning.

Methods

Study Design

This study was registered with the Australian and New Zealand Clinical Trials Registry (ANZCTR), which is a primary registry in the World Health Organization Registry Network (registration number: ACTRN12619001302145p).

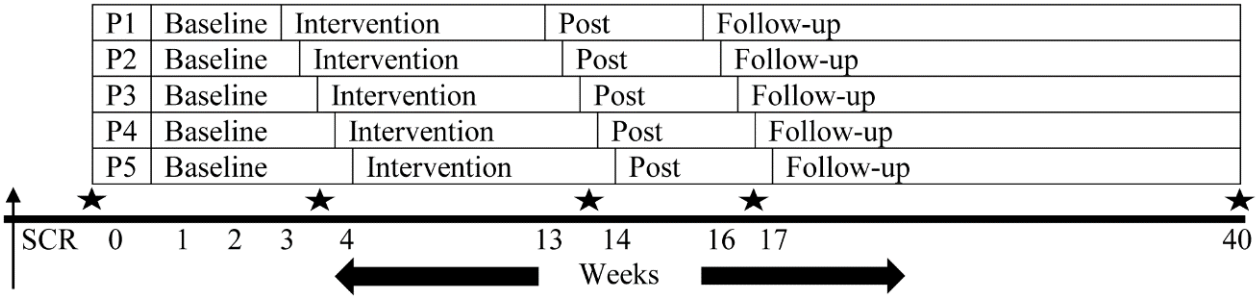
A multiple baseline across-individuals design will be employed. Multiple baseline designs in mental health intervention studies are those in which a baseline period of stability of symptoms is established before the intervention is introduced. In this way, each participant acts as their own control, and internal validity is demonstrated when there is no change in symptoms until after the treatment is introduced (Barlow et al., 2009). In the design to be used in this study, all participants commence the baseline period at the same time but start the treatment at different times after a minimum number of daily data readings (at least 20) have been received. This sequential commencement approach further strengthens the internal validity by reducing the likelihood of history, maturation, or other extraneous factors explaining any observed emotional or behavioral change that occurs simultaneously with the introduction of the treatment. The multiple data recordings allow for the use of analytical techniques such as a time series analysis (Jebb et al., 2015; Kazdin, 2017) and will involve participants reporting ratings of subjective units of distress (SUDS) via SMS text messaging

using a 10-point scale. In this design, 4 or more baselines are recommended (Barlow et al., 2009; Kazdin, 2017), and these will follow the pattern shown in Figure 1.

This research will use a *prescribed dosage* approach, as if the app was a *digital antidepressant*: one 10-min *dose* of app use per day for 5 days per week. The 10-week intervention period creates equivalence with one 50-min session per week for 10 weeks, which is the annual maximum number of psychology sessions rebated under Australia's Medicare system (Australian Government Department of Health, 2019b). The rationale for a minimum 3-week postintervention period (to demonstrate the stability of the treatment effect) is similar to that described earlier, namely, that 20 daily SUDS ratings are needed for a valid data analysis. The justification for using the 5 chosen apps are as follows: (1) each has some evidence of efficacy published in a peer-reviewed journal, (2) all have publicly available free versions, and (3) all can fit the *prescribed dosage* approach. The chosen apps use 3 popular evidence-based frameworks employed across mental health settings for treating anxiety and/or depression: cognitive behavioral therapy (CBT), mindfulness, and positive psychology. Finally, the rationale for using 10 participants per app is that, even accounting for a 50% to 60% attrition rate (Hochheimer et al., 2016), approximately 3 to 5 participants per app will provide enough data for a valid statistical analysis using time series conventions. As previously noted, this number of replications in a single-case design with similarly presenting individuals can produce robust and generalizable findings if the results are comparable in each case (Barlow et al., 2009; Horner et al., 2005; Kazdin, 2017).

Figure 6.1

Overview of the Proposed Study Design



SCR=Initial screening before intervention.
P1-P5=Participants 1 to 5.
Baseline=Minimum of 3 weeks (P1) and maximum of 4 weeks (P5) for daily SUDS ratings, no prescribed app use.
Intervention=10 weeks of prescribed app use and daily SUDS ratings.
Post=3 weeks of daily SUDS ratings, no prescribed app use.
Follow-up=6 months, no SUDS ratings, no prescribed app use.
★=Mental health and well-being outcome measures administered.

Recruitment

Commencing in January 2020, participants will be recruited throughout Australia by advertising the proposed study to nongovernment organizations that run programs for clients with mental illness (eg, the Benevolent Society), contacting associations of mental health professionals that may alert their members to the proposed study (eg, Australian Psychological Society), and contacting support groups and other organizations in the mental health sector (eg, Mental Health Victoria), requesting they advertise the proposed study on their various social media platforms. The advertisement is shown in Figure B.1 in Appendix B. Recruitment will cease once 50 participants are recruited. Owing to the nature of the proposed study design, new participants cannot commence after the study has started because the multiple baseline design requires participants to begin at the same time and then have specifically staggered phase commencements after that. Figure 6.1 demonstrates this process.

All 50 participants (10 for each app) will be randomized to an app and their position in the single-case design (ie, P1 to P10) using the web-based random number generator,

Research Randomizer (Urbaniak & Plous, 2019). The full inclusion and exclusion criteria are presented in Textboxes 1 and 2, respectively. A financial reimbursement will be offered to participants of Aus $0.50 (US $0.33) per daily SUDS rating sent via a text message. The researchers acknowledge that financial payments have the potential to interfere with ecological validity, because a person in the community would not normally be paid for using a mental health app, and intrinsic motivation, because people could potentially use the app for the benefit of financial remuneration rather than for the value of improving their mental health. However, the low amount of remuneration being offered of approximately Aus $45.00 (US $29.41) on average is not considered payment for participation in the proposed study but rather reimbursement of personal expenses incurred while taking part in the proposed study. Given that the effectiveness of mental health apps has the potential to benefit those from low socioeconomic groups, being reimbursed for providing in excess of 80 text messages will alleviate reasons that a potential participant of lower socioeconomic background could provide for being out of pocket for the cost of sending text messages from their mobile phone. Therefore, it is envisaged that prospective participants will more likely have motivation to improve their mental health beyond receiving financial remuneration. In addition, this financial incentive will not be advertised, and participants will only learn about this when they provide consent when completing the demographics questionnaire.

Textbox 6.1

Study Eligibility: Inclusion Criteria

Inclusion criteria:

- 18 years of age or older
- Ability to read English
- Have access to a smartphone or tablet device capable of connecting to the internet and downloading the required app and sending and receiving SMS text messages
- Agreeable to providing daily subjective units of distress ratings via SMS text messages and to completing self-report measures at 5 different time points (including 6-month follow-up)
- Mild-to-moderate anxiety and/or depression, diagnosed by a qualified health professional and confirmed by the researchers (all of whom are clinical psychologists) after screening. Screening involves analyzing the participants' scores on the first completed set of outcome measures: the Depression Anxiety Stress Scale-21 short-form version and the Outcome Questionnaire-45 second edition version. For more information on these, see the *Mental Health and Well-Being* subsection.

Textbox 6.2

Study Eligibility: Exclusion and Removal Criteria

Exclusion criteria:

- Severe anxiety and/or depression, as indicated by the initial outcome measures and in any responses to specific questions in the demographics questionnaire
- History of psychosis or other complex mental health presentation as deemed by the researchers to be unsuitable for participation in this research. There will be a question in the demographics questionnaire that asks participants for their complete mental health diagnoses
- Current suicidal ideation, as indicated by a participant's responses on the initial outcome measures

Removal criteria

- Not providing any subjective units of distress rating for a 2-week period
- Not providing a minimum of 20 subjective units of distress ratings in the baseline and postintervention phases or a minimum of 40 subjective units of distress ratings in the intervention phase
- Not completing outcome measures either preintervention or postintervention
- Clinically significant/unsafe decline in mental health as indicated by subjective units of distress ratings or outcome measures or in the judgment of researchers
- Suicidal ideation

Materials

Participants will supply their own smartphones and/or tablet devices. In total, 5 different apps will be used: (1) *Destressify* (Lee & Jung, 2018; Stress Refuge Inc., 2015), (2) *MoodMission* (Bakker et al., 2018a, 2018b; Bakker & Rickard, 2019; MoodMission, 2019), (3) *Smiling Mind* (Flett et al., 2019; Smiling Mind, 2019), (4) *MindShift* (Anxiety Canada, 2019; Paul & Fleming, 2019), and (5) *SuperBetter* (Roepke et al., 2015; SuperBetter LLC, 2019.

All the apps are supported by published research demonstrating statistically significant efficacy for the treatment of anxiety and/or depression. Each app has an accompanying website with further information and an accessible privacy policy. Detailed information about each app and its accompanying research is provided in Appendix B.

Measures

A number of measures of participants' experiences and outcomes will be used, as described in the following sections.

Biographic and Demographic Features

The demographics questionnaire has been developed by the researchers to obtain information that will be examined to ascertain if any patterns in the outcome data are related to aspects of an individual's demographic profile. Areas covered include mental health literacy (Jorm, 2012), motivation to change (Addis & Jacobson, 2000), chronicity of anxiety and/or depression (Hamilton & Dobson, 2002), and technology proficiency, all of which may influence results. Appendix B contains the complete demographics questionnaire and all other measures used in this research.

Mental Health and Well-Being

A 3-phase model of psychotherapy outcomes (Howard et al., 1993) is applied.

1. Subjective well-being: SUDS ratings—participants rating their well-being in response to the question, “How do you feel today?,” with 0 indicating no distress and 10 indicating worst distress (Wolpe & Lazarus, 1966).
2. Symptoms: the Depression Anxiety Stress Scale-21 short-form version (DASS-21) (Henry & Crawford, 2005). Participants rate their experience of symptoms of depression, anxiety, and stress over the previous week on a 4-point scale, ranging from 0 (did not apply to me at all) to 3 (applied to me very much or most of the time). Items in each subscale are summed to provide scores for symptoms of depression and anxiety, with higher scores indicating greater severity of symptomatology. The total scores for depression and anxiety subscales are multiplied by 2 in order to interpret the norms (Antony et al., 1998; Lovibond & Lovibond, 1995). Severity ratings for the depression subscale are 0-9 (normal), 10-13 (mild), 14-20 (moderate), 21-27 (severe), and ≥28 (extremely severe). Norms for the anxiety subscale were set as 0-7 (normal), 8-9 (mild), 10-14 (moderate), 15-19 (severe), and 20+ (extremely severe). The DASS-21 has been shown to demonstrate sound psychometric properties and validity (Osman et al., 2012).
3. Life functioning: the Outcome Questionnaire-45 second edition version (OQ-45.2) (Boswell et al., 2013) is a 45-item self-report scale that measures overall interpersonal relationships and social role functioning in adults aged 18 years and older (Beckstead et al., 2003). An index for overall life functioning is calculated (Lambert & Finch, 1999). Participants rate their feelings over the previous week on a 5-point scale, ranging from 0 (never) to 4 (always). The scale consists of both positive and negative items that are reverse-scored; higher scores indicate greater symptoms of distress and difficulties in interpersonal relations. A total score of ≥63 is indicative of clinically significant symptoms, with the subscale cutoffs for clinical significance being 35 for

symptom distress, 14 for interpersonal relations, and 11 for social role (Lambert & Finch, 1999). The OQ-45.2 has demonstrated high internal consistency (α=.90) and test-retest reliability of *r*=.84 over a minimum 3-week period (Lambert, 2012). The OQ-45.2 has also shown good construct and concurrent validity in a community sample when using the total score as opposed to interpreting the 3 individual subscales (Mueller et al., 1998).

Experience of App Usage

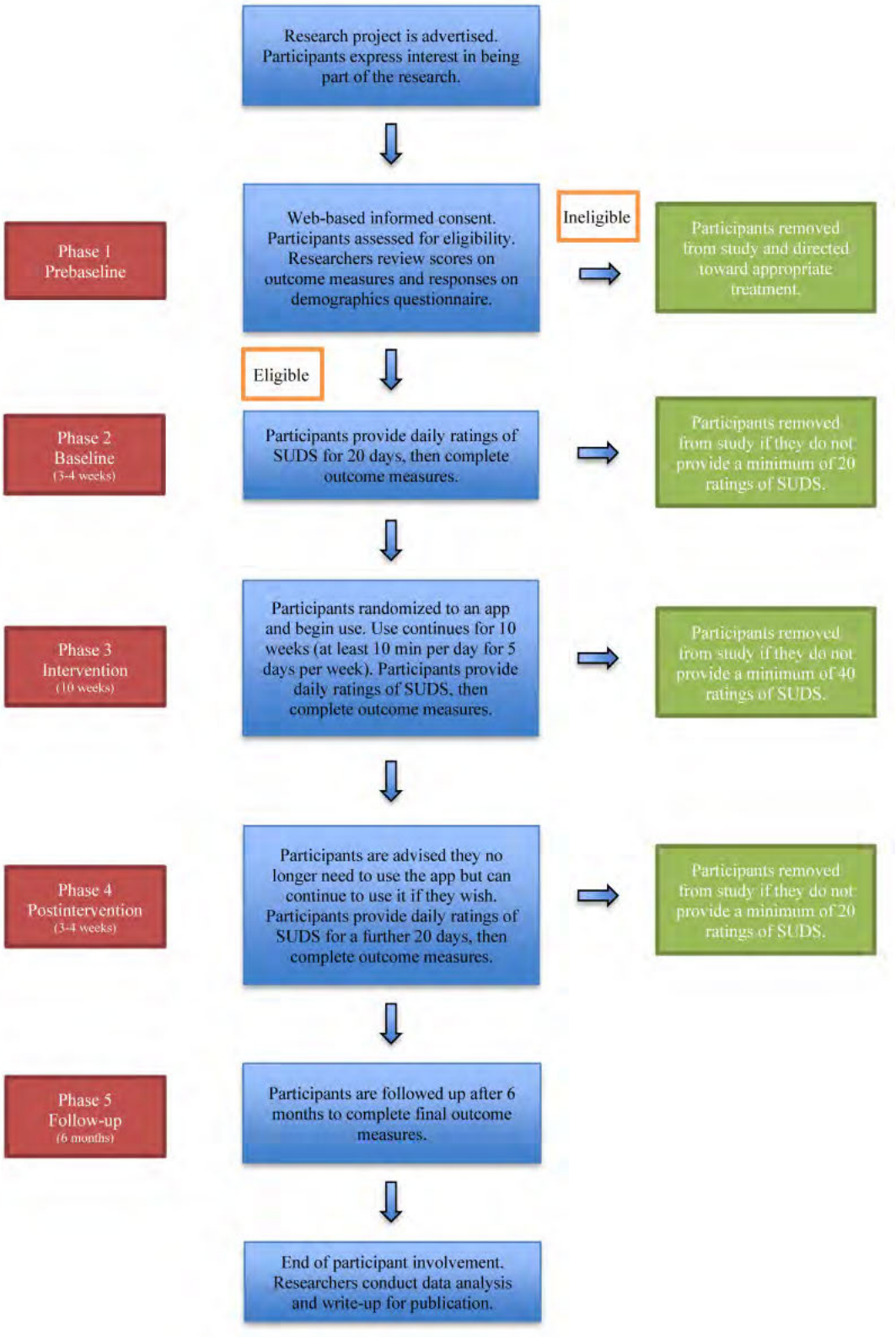
The Mobile Application Rating Scale-user version (uMARS) (Stoyanov et al., 2016) is a 20-item questionnaire that records an individual's rating on the quality of a mobile app. It contains multiple-choice and Likert-type responses and a free text field allowing users to provide a qualitative description of any aspect of the app or their experience of using the app that they wish to comment on. The uMARS contains 5 subscales: engagement, functionality, aesthetics, information quality, and a subjective quality appraisal. It has been found to have excellent internal consistency (***α***=.90) and good test-retest reliability (Stoyanov et al., 2016).

Procedure

Figures 6.1 and 6.2 illustrate the phases of this research. Recruitment commenced in January 2020.

Figure 6.2

Flowchart of Study and Participant Involvement



Phase 1 (Prebaseline)

Web-based links to the information sheet for participants and consent form, the demographics questionnaire, and the mental health and well-being outcome measures (DASS-21 and OQ-45.2) are sent by email and SMS text messages and completed digitally by participants using the Qualtrics survey platform (Qualtrics, 2019) and the OQ-Analyst platform (OQ Measures, 2019). Participants are screened for suitability to be in the proposed study by having their outcome measure and demographics questionnaire responses analyzed for evidence of severe anxiety and/or depression, suicidal ideation, or the presence of other severe mental illnesses such as psychosis. If researchers require further information from any participant, the participant will be contacted to clarify any queries or concerns. If a participant is deemed inappropriate for the proposed study, she/he will be directed by the researchers toward more appropriate forms of care.

Phase 2 (Baseline)

Accepted participants provide daily SUDS ratings for a minimum of 20 days or until a stable baseline profile of current psychological distress is achieved. At the end of this phase, the mental health and well-being outcome measures (DASS-21 and OQ-45.2) are completed.

Phase 3 (Intervention)

Participants are provided with generic instructions for all apps, links to both the Apple App Store and Google Play Store for their app, and specific instructions on how to use their app once it is downloaded (see Appendix B). In addition, website links to information on the type of evidence-based framework their app uses and emergency contact information in the event of a mental health crisis are provided. Participants continue to supply daily SUDS ratings for the minimum 10-week intervention. Data analysis will be ongoing throughout this phase and will be used to assist in determining whether any participant's mental health is

significantly deteriorating. If a participant provides a SUDS rating of 10 for 2 consecutive days, they will be contacted for a check on their welfare. Similarly, if a participant's SUDS ratings are above 8 for 5 consecutive days, they will also be contacted for a check on their welfare. We have chosen these cutoff values because the information provided to participants about the SUDS indicates that 8 is equal to their perception of feeling *very distressed* and 10 is equal to their perception of feeling the *worst distress.* The SUDS does not have a universal categorization label for each point on the scale, in addition to the number. Instead, it was designed to allow flexibility in an individual's self-assessment (Wolpe, 1969) and labeling can vary from study to study. The mental health and well-being outcome measures (DASS-21 and OQ-45.2) and uMARS are completed at the end of this phase. If a participant's responses on the outcome measures reveal a clinically significant decline in mental health compared with their responses at the beginning of the intervention, which places them in a severe category of mental illness, she/he will be contacted for a check on their welfare. In all cases, if a participant is categorized as being inappropriate for continuation in the proposed study, she/he will be directed toward more appropriate forms of care.

Phase 4 (Postintervention)

Participants provide SUDS ratings for at least 20 days following the completion of their official intervention period. Once a minimum of 20 SUDS ratings have been received, they will complete another DASS-21 and OQ-45.2. Participants are given information on all the apps so that they may explore the others if they wish.

Phase 5 (Follow-Up)

Participants are followed up at 6 months, where they will be asked to complete the mental health and well-being outcome measures (DASS-21 and OQ-45.2).

Expected Time Frames

The daily SUDS text messages will take a few seconds to reply to; app use will be a minimum of 10 min per day, 5 days per week for 10 weeks; the mental health and well-being outcome measures (DASS-21 and OQ-45.2) are expected to take less than 10 min each to complete on 5 different occasions; and the demographics questionnaire completed in phase 1 and the uMARS questionnaire completed in phase 3 are expected to take 15 min each.

The proposed study will run for approximately 40 weeks. Data analysis will be completed by approximately December 2020. A write-up for publication is expected to be completed by January 2021.

Data Analysis

Descriptive Statistics and Qualitative Accounts

Descriptive statistics will be used to compare individuals and augment other analytical techniques. The data obtained from the uMARS will be used to assist in gaining an enhanced understanding of participant attitudes toward their app. Depending on the amount of qualitative information provided by participants, it will be converted via a *content analysis* (Krippendorff, 2013) and will be coded into networks that hierarchically classify, identify, and summarize key themes. Data obtained from the uMARS will be plotted, as explained in the Visual Inspection subsection.

Time Series Analysis

A process for conducting time series analyses for psychological research was described by Borckardt et al. (2008) using the *R* statistical software package. In the proposed study, the commencement of the intervention will be the predictor in a regression model that uses data before and after this point to determine if there has been a statistically significant impact on subjective distress, as measured by SUDS ratings. A minimum of 20 data points

are required in each phase (Jebb et al., 2015; Kazdin, 2017). The *R* statistical package will use conventions of autoregressive integrative moving average (ARIMA) modeling to account for autocorrelated data (Houle, 2009) when building the model.

The time series analyses (Borckardt et al., 2008) will evaluate statistically significant changes across the phases of the proposed study. Overall level and trends across time will be considered, and if necessary, adjustments will be made for irregular variation effects. An *irregular factor* is similar to the error terms used in many statistical models, such as generalized linear modeling. The methods of making such adjustments differ depending on the nature of the collected data but may include the augmented Dickey-Fuller test, Durbin-Watson test and/or the Ljung-Box test as part of an ARIMA model.

Clinical Significance and Statistical Reliability

Meaningful or clinically significant changes occur when an individual is in the dysfunctional (clinical) range at the commencement of treatment and in the functional (nonclinical) range at the end of treatment (Evans et al., 1998; Jacobson & Truax, 1991). The clinical significance index (CSI) indicates whether individuals have made meaningful improvements to their emotional health and moved from being clinically dysfunctional to functional (Jacobson et al., 1999). The reliable change index (RCI) verifies the statistical significance of any change in an individual's score from pre- to postintervention (Jacobson & Truax, 1991). This approach is particularly useful for single-case designs because it allows researchers to focus on individual functioning (Kazdin, 1999) and to adjust treatment if necessary. Jacobson and Truax (1991) developed a classification system to describe the change in a participant's mental health in a study's conclusion: *recovered* = clinically significant and statistically reliable; *improved* = not clinically significant, but statistically reliable; *unchanged* = not clinically significant or statistically reliable; and *deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

To determine the CSI, a cutoff point between the scores obtained by the functional and dysfunctional populations on a particular measure is identified (Evans et al., 1998; Jacobson & Truax, 1991). Scores on either side of this point are statistically more likely to indicate whether an individual is functional or dysfunctional (Evans et al., 1998; Jacobson & Truax, 1991). Normative data are required for both functional and dysfunctional populations for the measures being used. CSI is based on the following formula (Jacobson & Truax, 1991):

$$\text{CSI cut-off} = \frac{(SD_1 \times M_2) + (SD_2 \times M_1)}{SD_1 + SD_2}$$

Note: 1 = Non-clinical population
2 = Clinical population

The RCI is a function of a measure's standard deviation and reliability (Evans et al., 1998). It measures an individual's change in self-reported score from pretreatment to follow-up for statistical reliability. If an individual's change exceeds 1.96 times the SE, the change is statistically reliable at $P<.05$ because it is unlikely to occur more than 5% of the time as a result of measure discrepancy or chance (Evans et al., 1998). RCI is calculated as follows (Jacobson & Truax, 1991):

$$RCI = 1.96 \times S_{diff}$$

Note: $S_{diff} = \sqrt{2}(S_E)^2$
Where S_E = *SD* of both groups x ($\sqrt{1}$ – test-retest reliability)

Clinical significance will be calculated based on participants' scores on the mental health and well-being outcome measures (DASS-21 and OQ-45.2) across the various phases using the framework suggested by Jacobson and Truax (1991). Using the OQ-Analyst

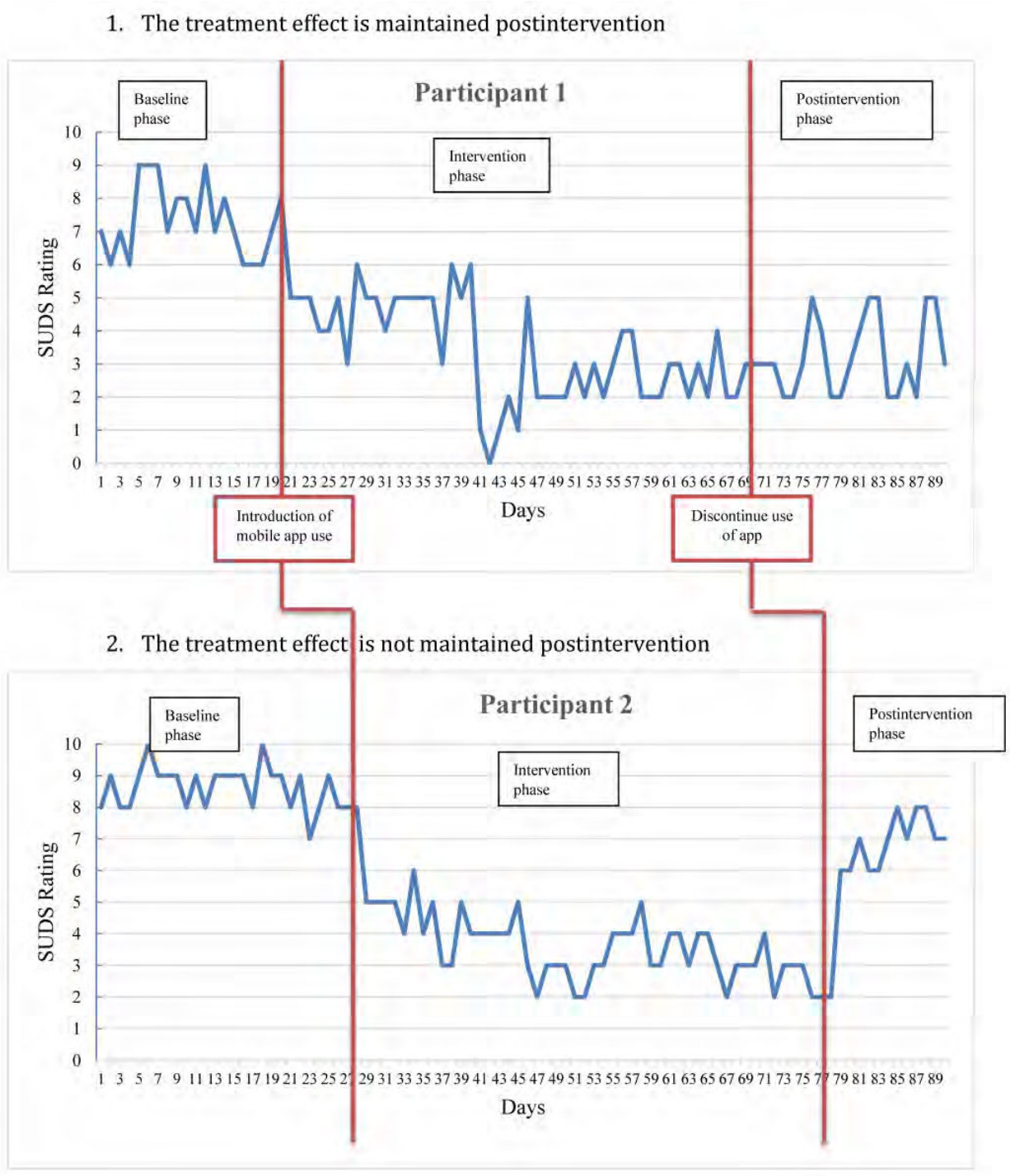
platform, clinical significance will be compared with statistical significance and visual inspection.

Visual Inspection

Visual inspection of plotted data allows for a personal judgment about the effect of an intervention and can often produce more meaningful information than approaches involving the calculation of statistical significance (Kazdin, 2017). In this study, visual inspection will be possible using up to 112 data points of SUDS ratings, as illustrated in Figure 3. Data obtained from the uMARS will be plotted against participant ratings from the mental health and well-being outcome measures (DASS-21 and OQ-45.2) and SUDS data are inspected for any observed relationships.

Figure 6.3

Example of how Participant Data may be Affected and Graphed Using a Multiple Across-Individuals Design



Data Management

Management and storage of data will occur in line with the Management and Storage of Research Data and Materials Policy of the University of New England (The University of New England, 2019). Specifically, all nondigital materials will be scanned and digitally stored indefinitely with all other digital information pertaining to this research at the University of New England research data cloud storage facility. Digital information will be password protected and accessible only to appropriate research staff.

Data Exclusion

Data will be excluded from time series and visual analyses if a participant fails to provide a minimum of 20 SUDS ratings in phase 1 (baseline) and/or phase 4 (postintervention) and 40 SUDS ratings in phase 3 (intervention). Data will be excluded from CSI and RCI analyses if a participant fails to complete baseline and/or postintervention mental health and well-being outcome measures (DASS-21 and OQ-45.2). Owing to the nature of the proposed study design, participants who dropout because they did not provide the required minimum data cannot be replaced by new participants once the baseline phase has commenced.

Ethics Approval

Ethics approval was granted by the University of New England Human Research Ethics Committee on November 1, 2019 (approval number: HE19-186). This research will be conducted under the guidelines of the National Statement on Ethical Conduct in Human Research by the Australian National Health and Research Council (National Health and Medical Research Council, 2018). Any changes to procedures outlined in this protocol will be forwarded to the University of New England Human Research Ethics Committee for approval before implementation. Such changes will also be acknowledged on the trial registration at the ANZCTR.

Results

Reporting of results will follow the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth (Eysenbach & CONSORT-EHEALTH Group, 2011) guidelines. The Procedure section provides the estimated timelines.

Discussion

Principal Findings

Information about and descriptions of the previous research on each app are provided in Appendix B. Previous research on mental health apps is lacking, and there are other issues impacting this research. The methodologies employed in the studies based on the apps used in the proposed study are heterogeneous, and this is in keeping with other previous research on other mental health apps (Marshall et al., 2019). The studies here had varying attrition rates (the *Destressify* research (Lee & Jung, 2018) reported 19.9%; *MoodMission* (Bakker et al., 2018b), 54.8%; *Smiling Mind* (Flett et al., 2019), 17.7%; *MindShift* (Paul & Fleming, 2019), 46.7%; and *SuperBetter* (Roepke et al., 2015), 73.9%), but the varying intervention times and methodologies may have contributed to these attrition rates. Some studies were conducted by researchers who either developed the app or had an otherwise pre-existing association with the app being tested (in the case of *MoodMission* and *SuperBetter*). Other studies had participants with varying degrees of severity of anxiety and depression at commencement that were measured with varying outcome instruments (*Destressify*, *MoodMission,* and *Smiling Mind* did not specify the participants' mental health status in their inclusion criteria; *MindShift* had participants with moderate-to-high levels of anxiety, as indicated by their ratings on at least one scale of the Patient Health Questionnaire; and *SuperBetter* had participants who scored higher than 16 on the Center for Epidemiological Studies Depression Scale). Nevertheless, all studies were published in peer-reviewed journals and authored by individuals who have associations with legitimate academic institutions and

mental health organizations. Therefore, we hypothesize that our results will reveal the apps to be effective for reducing symptoms of anxiety and depression, as reflected in the previous research.

Potential Added Value for Clinicians and Consumers

This study is unique with respect to any published study of mental health apps that could be located. This independent research, using a single-case methodology, allows for an in-depth examination of personal factors that may impact the effectiveness of these apps. It will therefore add value to existing studies on these specific apps. However, it is also anticipated that this research will be automated to the point where the design can be used to examine larger samples with the rigor of an RCT experimental design, thereby having a positive impact on increasing future research at a faster rate. It also provides an opportunity for evidence-based mental health treatment to reach those who are not already receiving it.

If mental health apps have demonstrated effectiveness, they could be incorporated by clinicians into face-to-face therapy to enhance the experience of consumers. For example, some apps allow users to complete homework tasks set by their therapist or to make thought diary entries that can be shared digitally with their therapist. Other information such as physiological readings and user-entered information such as SUDS ratings can also be sent digitally to clinicians to gain a more accurate reading of their client's emotional health between sessions (Beiwinkel et al., 2016).

There are potential benefits for health systems and mental health consumers if apps can gain increased legitimacy for their ability to effectively manage anxiety and depression. These include the following: more economical for low socioeconomic groups to obtain mental health treatment compared with face-to-face services (Jones et al., 2014), improved access for those in rural areas where there may be limited treatment options (Sakai et al., 2014), reduced stigma (Bowers et al., 2013) because of anonymous assistance, access for

children and adolescents who are already large consumers of smartphones and the internet (Holloway et al., 2013), and it is simply a preferred way to receive mental health information for some (Wang et al., 2018). Therefore, it is important to increase research on the efficacy and effectiveness of mental health apps using appropriate and scientifically validated methodologies in addition to RCTs, as the widely considered gold standard of RCTs may not be the most appropriate for analyzing mental health apps (Clough & Casey, 2015a; Marshall et al., 2020e).

This study will examine a number of mental health apps that differ in several ways: (1) having different theoretical frameworks (*MoodMission* and *MindShift* use CBT, *Destressify* and *Smiling Mind* use mindfulness, and *SuperBetter* uses positive psychology), (2) being developed by different teams in different countries (*Destressify* was developed in the United States by individuals with an interest in mindfulness meditation, *MoodMission* was developed in Australia by a team of psychologists and researchers at Monash University, *Smiling Mind* was set up as a not-for-profit organization in Australia by mental health and meditation experts, *MindShift* was developed in Canada by a not-for-profit mental health organization, and *SuperBetter* was developed in the United States by a game designer and mental health researchers from Stanford University and the University of Pennsylvania), and (3) containing different aesthetic qualities and types of activities with different aims (*Destressify* focuses on reducing stress, *MoodMission* focuses on providing short activities designed to help an individual in response to how they are feeling at that time, *Smiling Mind* focuses on teaching mindfulness skills in a structured format using guided meditation, *MindShift* focuses on reducing anxiety by using a number of different interventions such as graded exposure and using a thought journal, and *SuperBetter* is very colorful with a playful tone that may appeal to individuals who like video games). Appendix B provides further information about each app. Having a diversity of apps is important because there may be

differences in the way consumers react to different aspects of an app. It is known that face-to-face therapy outcomes can be influenced by client-therapist rapport (Tang & DeRubeis, 1999), client motivation (Addis & Jacobson, 2000), and chronicity/history of mental illness (Hamilton & Dobson, 2002). Therefore, there may be different aspects of a mental health app that contribute to its effectiveness, such as gamification, aesthetics, usability/interface (Rickard et al., 2016), and evidence-based framework.

In sum, the reasons mentioned earlier support the need for a vigorous research agenda on the effectiveness of mental health apps, and this study methodology can assist in realizing this.

Limitations and Strengths

This study has some limitations. First, it may not be possible to generalize the findings if the outcomes for participants with the same condition differ in significant ways. Second, there is no certainty that participants will provide daily SUDS ratings for 16 weeks, despite the minimal effort involved. Third, with many brands of smartphones using different versions of software, there is a risk that the technology between phone and app may not be compatible for some participants. Fourth, the researchers will not be able to access data directly from the app itself, and therefore will be unable to know exactly how much time participants spend using the app and thus comment on treatment fidelity. Although the technology exists to do this, it would require the resources of a study with greater funding, and would require the involvement and cooperation of app developers. This means that judgments about the extent that the app was responsible for any change in symptomatology or life functioning are limited. Furthermore, if access to app data was available, those participants that drop out of the study would also be able to contribute data that may impact on the findings. Having access to app data is a consideration for future research in this area.

This study also has several strengths. By using questionnaires that consider subject distress, symptoms, and life functioning at different time points, the design allows for a comprehensive approach toward the impact of the apps. The use of the DASS-21 and OQ-45.2 questionnaires at multiple time points will allow an examination of issues such as suicidality, dysfunctional coping activities (such as excessive alcohol consumption), physical health, and sleep disturbance. The single-case design will also provide in-depth information about individual responses and offers a way in which clinicians may be able to contribute to the evaluation process (see the Conclusions section). Finally, there is the ambitious goal of offering a future methodology that could be applied to larger RCTs via a highly digitized procedure.

Conclusions

The evidence base for mental health apps that offer treatments for anxiety and depression is currently low. This study may assist in improving this situation in several ways. First, it may allow more clinicians to participate in the research process. Marshall et al. (2020e) have outlined a proposal to establish a centralized database where clinicians and researchers contribute information and data on the effectiveness of mental health apps by using a standardized protocol that forms the basis of this research. Such a repository of information on mental health apps would mean an ever-increasing knowledge base that clinicians, researchers, government authorities, and academic institutions could refer to. Although there are existing websites that offer professional reviews with useful insights into mental health apps (eg, *PsyberGuide* (One Mind, 2019), *Head To Health* (Australian Government Department of Health, 2019a), *Reachout Australia* (Reachout.com, 2019), *Health Navigator* (The Royal New Zealand College of General Practitioners, 2019), and the *NHS App Library* (National Health Service, 2019)), these are based on professionals' perspectives and not systematic and scientific observations. The results of increased research,

such as that outlined in the proposed study, have the potential to add valuable empirical data to such websites to reinforce the reviews posted there.

If a collaborative scientific methodology was used by clinicians and researchers to rate the effectiveness of mental health apps, this would also potentially allow more transparent categorization of mental health apps in the various app stores (Marshall et al., 2020e). Currently, reliance on app stores leads to potential confusion for consumers as ratings and reviews may be unreliable or even fake (Xie & Zhu, 2015). Using the methodology in this study is one way that, if willing, the app stores could *certify* mental health apps as having reached an acceptable level of independently verified effectiveness (Marshall et al., 2019). This would allow consumers to more clearly identify apps validated by scientific research.

Finally, given the large number of consumers who own a smartphone globally (Statista, 2019b), if more people are able to use efficacious mental health apps on their phones, it could potentially free up scarce face-to-face services in communities struggling to meet the demand for interventions to address mild-to-moderate mental health problems.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 7 – Study 3

In Chapter 6, the research protocol outlined the planned methodology for Study 3 and Study 4, which followed on from the rationale and overviews provided in earlier chapters. In Chapter 7 – Study 3, a description of the pilot study is given, culminating in initial verification that a mental health app can be examined for evidence by using a digitally-enhanced single-case design. It also showed that the app, *SuperBetter*, was effective in alleviating symptoms of anxiety and depression, replicating findings from the limited research that has been conducted on this app in the past.

Chapter 7 – Study 3: Can a Smartphone App Make You Feel Super Better? A Pilot Study Utilizing a Multiple Single-Case Design

Marshall, J. M., Dunstan, D. A., & Bartik, W. (Submitted). Can a smartphone app make you feel super better? A pilot study utilizing a multiple single-case design. *Technology, Mind, and Behavior*.

Abstract

Consumers are increasingly turning to mobile mental health apps to manage a wide range of presenting concerns. Using a reinvigorated single-case study approach, this pilot study evaluated the effectiveness of the mental health app, *SuperBetter*, which has been developed for people with anxiety and depression. The study incorporated a multiple baseline across-individuals design with four participants (one male and three females, aged 20, 25, 27 and 49). Outcome measures of symptoms and functioning were taken at four different time points and SMS text messages were used to gather continuous ratings of subjective distress (SUDS) from baseline, intervention and post-intervention phases. The app was appraised using the user version of the Mobile Application Rating Scale. The intervention commenced after confirming a stable level of baseline responding. Distress data were analyzed using visual inspection and a time-series analysis with an autoregressive integrative moving average model. Statistically significant reductions in subjective distress were found for three out of four participants and these were maintained post-intervention. All participants showed beneficial changes on measures of symptoms and life functioning with variability being explained by individual differences. Participants rated the app favorably overall, indicating that they would recommend it to others. The findings of this research suggest that *SuperBetter* is effective at improving ratings of distress, anxiety, depression, and life

functioning across a diverse range of presentations and that a single-case research design has high utility for research in this field.

Can a Smartphone App Make You Feel Super Better? A Pilot Study Utilizing a Multiple Single-Case Design

Consumers are downloading mental health applications (apps) in increasing numbers. Currently, there are over 10,000 mental health apps to choose from (Torous, Firth, et al., 2018) but most of these have not been developed by recognized mental health experts (Alyami et al., 2017; Shen et al., 2015). There is wide agreement that greater regulatory oversight is required to ensure the safety of users (Marshall et al., 2020b), and that research is needed to bolster the fledgling evidence base of the effectiveness of these new digital tools (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Lui et al., 2017; Marshall et al., 2020e).

This study builds on the growing body of work on the use of technology in the treatment of mental disorders, recently summarized by Wilhelm et al. (2020). Following a brief overview of the research difficulties associated with examining mental health apps, this paper outlines a reinvigorated methodology incorporating a single-case research design, and reports on the findings of a pilot study utilizing this approach.

Research Difficulties for Mental Health Apps

Much of the limited research on mental health apps is heterogeneous, making it difficult to draw reliable generalizations from the findings. While it is acknowledged that this is partly due to smartphone apps being a relatively recent phenomenon, it is surprising given the extent of worldwide smartphone ownership and the widespread usage of apps. Heterogeneity has occurred as a result of a number of factors, including: studies employing different interpretations of particular methodologies, such as having and not having control groups (e.g. Flett et al., 2019; Paul & Fleming, 2019); variability in the length of the

intervention period that can range from 10 days (e.g. Howells et al., 2016) to 12 weeks (e.g. Boisseau et al., 2017); differing outcome measures (e.g. Lai & Jury, 2018); and, dosage instructions (i.e., frequency of use) that range from non-specific (Kuhn et al., 2017) to precise (Roy et al., 2017). Therefore, caution needs to be taken when interpreting the limited research completed to date, which has shown that mental health apps are efficacious in treating symptoms of anxiety and/or depression with a small to moderate effect size (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Lai & Jury, 2018; Mehrotra & Tripathi, 2018; Menon et al., 2017).

Another significant reason for the lack of research into mental health apps is the time required to conduct traditional experimental studies. The gold standard randomized controlled trial (RCT) can often take too long to run in the context of rapidly changing consumer demand and technology. By the time an app has gone through a lengthy RCT process, including participant recruitment, an intervention phase, a follow-up period and peer-reviewed publication, the app may have been replaced by updated versions or become obsolete as newer, more technologically superior apps have come onto the market (Clough & Casey, 2015a).

Criticism of the lack of use of RCTs is tempered in this context by the reliance on group means data, which makes micro aspects of a participant's response invisible to researchers (Barlow et al., 2009). It is important to acknowledge this shortcoming because research has shown that in-person, face-to-face therapy outcomes can be influenced by multiple factors, including client-therapist rapport (Tang & DeRubeis, 1999), client motivation (Addis & Jacobson, 2000), and chronicity/history of mental illness (Hamilton & Dobson, 2002). Consequently, there may be different aspects of a mental health app that contribute to its effectiveness, such as gamification, aesthetics, usability / interface (Rickard et al., 2016), as well as its evidence-based framework. Therefore, while acknowledging that

RCTs will remain a key approach for demonstrating the efficacy of an underlying intervention, to examine the effectiveness of mental health apps, a more nimble research approach may be required (Clough & Casey, 2015a).

Single-Case Designs for Mental Health App Research

One means of improving research into mental health apps is the use of single-case research designs (Clough & Casey, 2015b; Marshall et al., 2020e; Mehrotra & Tripathi, 2018). It is argued that when participants have similar presentations, receive the same intervention, and demonstrate similar results, conclusions from single-case design research are as robust and valid as those from larger experimental group designs, even when as few as three participants are involved (Barlow et al., 2009; Horner et al., 2005; Kazdin, 2017). Furthermore, the additional information available from single-case designs has the potential to enhance the findings of larger group designs (Buckley et al., 2014; Sheridan, 2014).

Single-case research designs can offer benefits to mental health app research that RCTs struggle to match. For example, participants can be monitored closely and this is crucial when suicide may be a risk. Daily ratings of Subjective Units of Distress (SUDS) is one method of monitoring that has several benefits, including: the rapid identification of changes in functioning and the impact of an intervention (Machalicek & Horner, 2018); the tailoring of the intervention if necessary (Bentley et al., 2019); and, the testing of nuanced hypotheses and analysis of results (Barlow et al., 2009).

Use of an across-individuals single-case design means that each participant becomes their own control, as opposed to needing a separate control group (such as placebo, waitlist, or treatment as usual). Specifically, each participant starts a baseline phase simultaneously, and must demonstrate a period of relative stability prior to the implementation of the intervention. Thereafter, the effects of the intervention can be judged reliably. Relatedly, each participant begins the intervention in a staggered way, thus allowing for easy identification of

any external / environmental events that may have impacted an individual's response to treatment. In these ways, internal validity is achieved.

Single-case research designs also afford the opportunity for practicing clinicians to become more involved in the research process by gathering "real world" information from data collected in their everyday workflow (Clough & Casey, 2015a; Marshall et al., 2020e). One idea that has been outlined elsewhere (Marshall et al., 2020e) describes how clinicians could contribute to the evaluation of mental health apps using the single-case design methodology described in this paper. This goes beyond what is offered on current review forums, such as *PsyberGuide* (https://psyberguide.org/); *Head To Health* (https://headtohealth.gov.au/); and the *NHS Apps Library* (https://www.nhs.uk/apps-library/).

The *SuperBetter* App and its Previous Research

The app used in this study was *SuperBetter*. Two published studies on the efficacy of *SuperBetter* could be located in the peer-reviewed literature.

Roepke et al. (2015) conducted an RCT with 283 participants (200 females) with a mean age of 40. The study focused on symptoms of depression. They were recruited from the Penn Authentic Happiness website and the Craigslist.org community bulletin board. One hundred and twenty four participants were taking psychotropic medication and 101 were engaged in psychological therapy, with no indication about the crossover of participants who fell under both of these categories. Statistically significant improvements in levels of anxiety and depression were reported for users of *SuperBetter* compared to the wait list control group. Other than the high attrition rate of 81% at the conclusion of the study, the other key limitation was the low treatment fidelity in that high numbers of participants did not download all of the information that formed part of the treatment.

Worthen-Chaudhari et al. (2017) conducted a non-randomised study with 42 teenage participants (31 females) with a mean age of 15. The study focused on unresolved concussion

symptoms, but used a depression scale and an optimism scale in its data collection. Participants were recruited through a concussion clinic. Statistically significant improvements in optimism were recorded for the app group compared to the treatment as usual group, but did not reach statistical significance for improvements in depression. The limitations of this study included a lack of random assignment, participants were recruited from a single source, and there was no distinguishing between type or severity of concussion injury.

The Present Study

The TREND Statement (Des Jarlais et al., 2004) was developed to improve the reporting quality of nonrandomized investigations into behavioral and public health interventions and mirrors the objectives of the CONSORT Statement (Schulz et al., 2010) in relation to RCTs. The Trend Statement has been used to inform the conduct of this study – the checklist is in Appendix B. The main objective of the study was to research the effectiveness of a mental health app in reducing symptoms of anxiety and/or depression, and to do so using a single-case design. The *SuperBetter* app was randomly chosen from a pool of five apps (see Procedure section below for further details on the rationale for app selection). The study also served the dual purpose of being a pilot study examining the overall feasibility of the research methodology prior to implementing it with a greater number of apps and participants.

The present study sought to answer the following research questions:

1. Is a mental health app (in this case, *SuperBetter*) effective at reducing subjective distress and clinically significant symptoms of anxiety and/or depression, and improving life functioning?
2. Are there specific factors about the participants that impact on the results?
3. Can a single-case research methodology using digital technologies adequately and comprehensively assess the effectiveness of a mental health app in an applied,

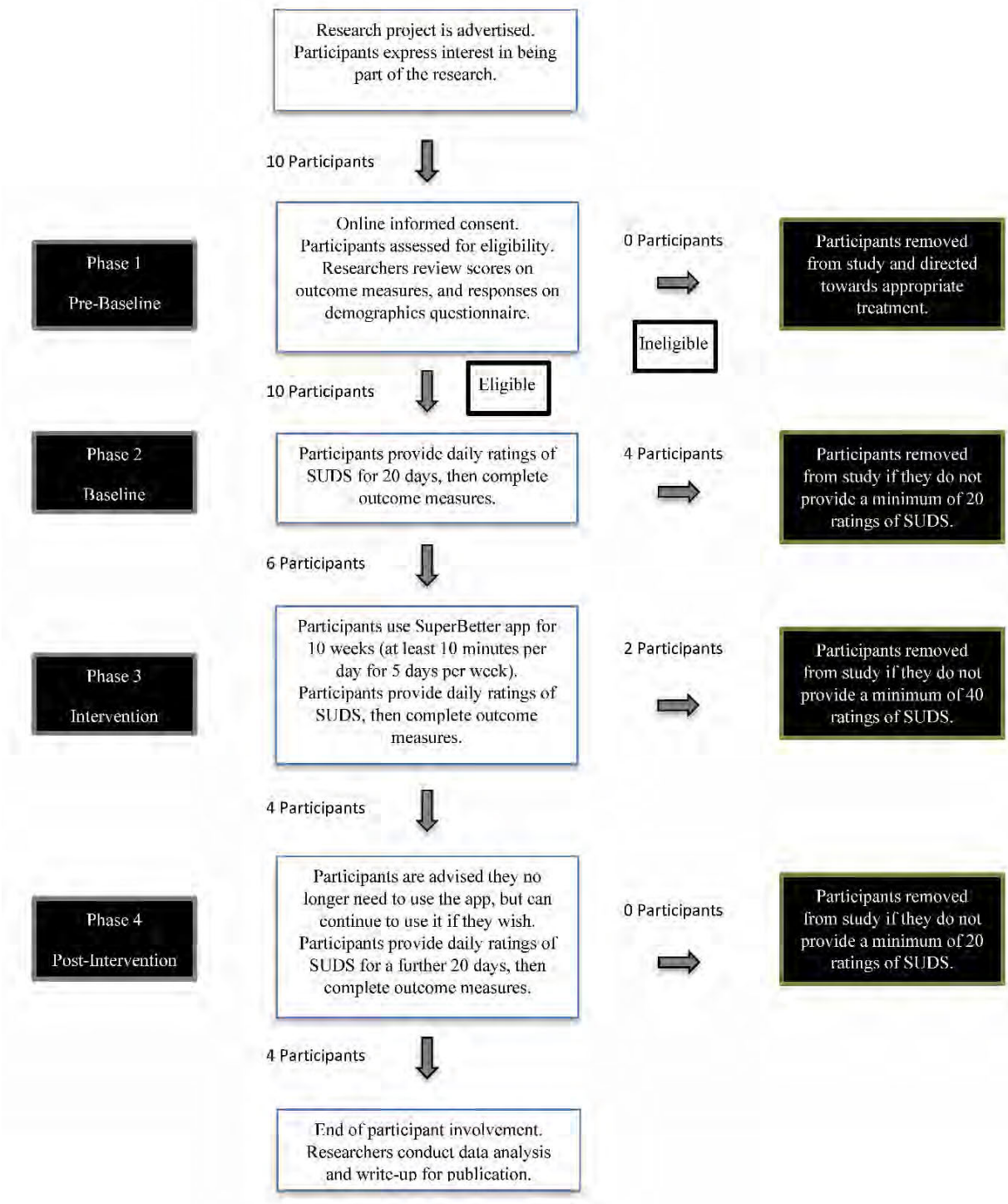
real-world setting? Can this methodology be scaled up to assess larger numbers of participants?

4. What was the participants' experience of the app?

Based on previous research (Roepke et al., 2015; Worthen-Chaudhari et al., 2017), it was hypothesized that *SuperBetter* would be effective at reducing symptoms of anxiety and depression, but that results may highlight micro aspects of participants that have been influential in achieving these results.

Figure 7.1

Flowchart of Study and Participant Involvement



Method

Participants

Inclusion criteria:

1. 18 years of age or older;
2. Ability to read English;
3. Have access to a smartphone or tablet device capable of connecting to the Internet and downloading the required app, and sending and receiving SMS text messages;
4. Agreeable to providing daily SUDS ratings via SMS text message, and to completing self-report measures at four different time points; and
5. Mild to moderate anxiety and/or depression, as previously diagnosed by a qualified health professional prior to applying to be in this research, and confirmed by the researchers (all of whom are Clinical Psychologists) after screening of the initial self-report measures to indicate the presence of symptoms of anxiety and / or depression using the cutoff points identified by previous research.

Exclusion criteria:

1. Severe anxiety and/or depression, as indicated by the initial outcome measures and in any responses to specific questions in the Demographics Questionnaire;
2. History of psychosis, or other complex mental health presentation as deemed by the researchers to be unsuitable for participation in this research; and,
3. Current suicidal ideation, as indicated by a participant's responses on the initial outcome measures.

Removal criteria:

1. Not providing any SUDS rating for a two-week period;

2. Not providing a minimum of 20 SUDS ratings in the baseline and post-intervention phases, or a minimum of 40 SUDS ratings in the intervention phase;
3. Not completing outcome measures either pre-intervention, or post-intervention;
4. Clinically significant/unsafe decline in mental health as indicated by SUDS ratings or outcome measures, or in the judgment of researchers; and
5. Suicidal ideation that has developed during the participants' involvement in the study.

Materials and Measures

The App

The app used in this study was *SuperBetter* (Roepke et al., 2015; Worthen-Chaudhari et al., 2017; https://www.superbetter.com). It was developed by a game designer in conjunction with researchers at Stanford University and the University of Pennsylvania. It uses concepts from positive psychology and neuroscience (specifically neuroplasticity) to inform the interventions it delivers via its game-like interface.

Demographic and Biographic Features

A demographics questionnaire was developed by the researchers to elicit demographic and biographic information. Areas covered included mental health literacy (Jorm, 2012), motivation to change (Addis & Jacobson, 2000), chronicity of anxiety and/or depression (Hamilton & Dobson, 2002), and proficiency with technology, all of which may influence results. Information gained from this questionnaire is used to interpret patterns in the outcomes data that may be related to aspects of an individual's personal profile.

Mental Health and Wellbeing

The three-phase model of psychotherapy outcomes (Howard et al., 1993) was used as the framework for examining participant outcomes relating to subjective distress, symptomatology, and life functioning as follows:

1. Subjective distress: SUDS ratings – participants rated their level of distress in response to the question: "How do you feel today?", with '0' indicating no distress, '10' indicating worst possible distress, and a score of '3' or more indicating a mild but noticeable level of upset (Wolpe & Lazarus, 1966). SUDS ratings have been shown to be a valid measure of emotional discomfort when compared with other measures of distress ($r = 0.351$, $p < .05$; Tanner, 2012).
2. Symptoms: The Depression Anxiety Stress Scale – 21 short-form version (DASS-21; Henry & Crawford, 2005). Participants rated their experience of symptoms of depression, anxiety and stress over the previous week on a 4-point scale ranging from '0' (*did not apply to me at all*) to '3' (*applied to me very much, or most of the time*). Items in each subscale are summed to provide scores for depression, anxiety/panic, and stress/generalized anxiety, with higher scores indicating greater severity of symptomatology. The total scores for the subscales are multiplied by two in order to interpret the severity ratings according to the longer 42-item scale (Antony et al., 1998; Lovibond & Lovibond, 1995). In this study, only the depression and anxiety subscales were used. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*). The DASS-21 has been shown to demonstrate sound psychometric properties (Osman et al., 2012): Depression scale $\alpha = .85$ and average inter-item correlation (AIC) of .47; anxiety scale $\alpha = .81$ and AIC of .40. The DASS-21 has also displayed good construct and concurrent validity in a sample of university students using multiple measures of distress in a study by Osman et al. (2012).
3. Life functioning: The Outcome Questionnaire – 45 2nd Edition version (OQ-45.2; Boswell et al., 2013) is a 45-item self-report scale that measures distress,

interpersonal relationships and social role functioning in adults 18 years and older (Beckstead et al., 2003). An index for overall life functioning is calculated (Lambert & Finch, 1999). Participants rate their feelings over the previous week on a 5-point scale ranging from '0' (*never*) to '4' (*always*). The scale consists of both positive and negative items that are reverse-scored; higher scores are indicative of greater negative symptoms and difficulties. Possible scores range from 0 to 180 with a total score of 63 or more being indicative of clinically significant symptoms (Lambert & Finch, 1999). Lambert et al. (2004) have suggested the following interpretive labels: > 105 is *High*, 83 – 104 is *Moderately High*, 63 – 82 is *Moderate*, and < 63 is *Normal*. The OQ-45.2 has demonstrated high internal consistency (α = .90) and test-retest reliability of *r* = .84 over a minimum 3-week period (Lambert, 2012). The OQ-45.2 has also shown good construct and concurrent validity in a community sample when using the total score as opposed to interpreting the three individual subscales (Mueller et al., 1998).

App Appraisal

The Mobile Application Rating Scale – User Version (uMARS; Stoyanov et al., 2016) is a 20-item questionnaire recording an individual's rating on the quality of a mobile app. It contains multiple choice and Likert-type responses, and offers a free text field allowing users to provide a qualitative description of any aspect of the app, or their experience of using the app. The uMARS contains five subscales measuring: engagement, functionality, aesthetics, information quality, and a subjective quality appraisal (satisfaction). It has been found to have excellent internal consistency for the total score (α = .90), and the individual scales (engagement α = .80; functionality α = .70; aesthetics α = .71; information α = .78; and satisfaction α = .78; Stoyanov et al., 2016). Also demonstrated is good test-retest reliability over a 1-month follow-up (*r* = .66) and at 6-month follow-up (*r* = .70; Stoyanov et al., 2016).

Data Analysis Plan

Descriptive Statistics and Qualitative Accounts

Descriptive statistics were used to compare individuals and augment other analytical techniques.

Visual Inspection

Visual inspection was used to assess the impact of the intervention on subjective distress (SUDS). Plotted data allows for a personal judgement about the effect of an intervention and can often produce more meaningful information than approaches involving the calculation of statistical significance (Kazdin, 2017). In this study, visual inspection was possible using up to 122 data points of SUDS ratings (this was the highest number of individual SUDS ratings, by Participant 1 – see Figure 7.2).

Time Series Analysis

A time-series analysis was used to assess the statistical significance of changes in participants' plotted data across each phase of the study. Scores at the commencement of the intervention were used as the predictor in a regression model. A minimum of 20 data points was required in each phase (Jebb et al., 2015; Kazdin, 2017). The *R* statistical package used conventions of autoregressive integrative moving average (ARIMA) modeling to account for autocorrelated data (Houle, 2009) when building the model. Overall level and trends across time were considered, and adjustments were made for autocorrelated data using the augmented Dickey-Fuller test and the Ljung-Box test as part of an ARIMA model.

Clinical Significance and Statistical Reliability

Symptoms of life functioning were assessed for clinical significance and change using the clinical significance index (CSI) (Jacobson et al., 1999) and the reliable change index (RCI) (Jacobson & Truax, 1991). Normative data for a scale is used to calculate the CSI,

which identifies the cut-off point between the scores obtained by functional (non-clinical) and dysfunctional (clinical) populations (Evans et al., 1998; Jacobson & Truax, 1991). The CSI was used to note each participant's clinical status pre- and post-intervention (Evans et al., 1998; Jacobson & Truax, 1991). The CSI cut-off point for the OQ-45.2 measure used in this study was: *Total Score* 63. Participants who scored at or below this figure were identified as non-clinical, and scores above this figure were considered clinical.

The reliable change index (RCI) was used to assess and classify the statistical significance of any change in participants' score from pre- to post-intervention: *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction. The RCI cut-off point for the OQ-45.2 measure used in this study was: *Total Score* 14.

Procedure

The present research was a pilot study for a larger research project that is registered with the Australian and New Zealand Clinical Trials Registry (ANZCTR), registration number ACTRN12619001302145p (http://www.ANZCTR.org.au/ACTRN12619001302145p.aspx). The research protocol for the larger study has been published (Marshall et al., 2020c), and the present pilot study followed the same protocol.[1] Data from this project is publicly available through the University of New England's *Research UNE* website (https://doi.org/10.25952%2F5f19135c5fbe5). Ethics approval was obtained by the University of New England Human Research Ethics Committee on the 1st November 2019, Approval Number HE19-186.

[1] Interested readers are invited to look at these two online documents for further information regarding the Method.

In November 2019, participants were recruited throughout Australia from non-government mental health services, mental health associations (both consumer and professional), and support groups and other organizations in the mental health sector. Ten participants were recruited by early December 2019. After informed consent was obtained, participants commenced the baseline phase simultaneously, and were randomly selected to begin the intervention phase in staggered order. Randomization was achieved using the online random number generator, *Research Randomizer* (Urbaniak & Plous, 2019; https://www.randomizer.org). See Figure 7.1 for a flowchart of the study's phases and participant involvement. No participant was removed from this study after commencement.

Participants received instructions on usage and background information relating to their particular app. Participants were asked to use their app for at least 10 minutes per day for five days per week for 10 weeks. The 10-week intervention period creates equivalence with one 50-min session per week for 10 weeks, which is the annual maximum number of psychology sessions rebated under Australia's Medicare system. It is acknowledged that these usage instructions do not necessarily equate to the ideal usage instructions across all mental health apps. However, in the interests of streamlining the process for all participants, it was believed that this instruction of at least 10 minutes per day for five days per week would encourage participants to engage with their app in a meaningful way.

The app used in this study was *SuperBetter* (Roepke et al., 2015; Worthen-Chaudhari et al., 2017; https://www.superbetter.com). It was randomly chosen, using the *Research Randomizer* website (Urbaniak & Plous, 2019; https://www.randomizer.org), from a pool of five mental health apps that had been selected for use in the main study. The other apps were *Destressify* (Lee & Jung, 2018; https://www.destressify.com), *MoodMission* (Bakker et al., 2018a, 2018b; Bakker & Rickard, 2019; https://www.moodmission.com), *Smiling Mind* (Flett

et al., 2019; https://www.smilingmind.com.au), and *MindShift* (Paul & Fleming, 2019; https://www.anxietycanada/resources/mindshift-cbt/). For more information about all of these apps, including why they were selected, see Chapter 6.

Phase 1 (Pre-Baseline): Online links to the Information Sheet for Participants and Consent Form, the Demographics Questionnaire, and the mental health and wellbeing outcome measures (DASS-21 and OQ-45.2) were sent by e-mail and SMS text message to participants. These were completed digitally using the Qualtrics (https://www.qualtrics.com/au/); NovoPsych (https://novopsych.com.au/); and OQ-Analyst (https://www.oqmeasures.com/oq-analyst-3/) platforms. Participants were screened for suitability to be in the proposed study and none were deemed to be inappropriate.

Phase 2 (Baseline): Participants provided daily SUDS ratings for a minimum of 20 days and until a stable level of recording was shown. At the end of this phase, the mental health and wellbeing outcome measures (DASS-21 and OQ-45.2) were completed again.

Phase 3 (Intervention): Participants were provided with information about the app, links to both the Apple App Store and Google Play, and beginning instructions on what to do once their app was downloaded. Also provided were website links to information on the evidence-based framework of the app and emergency contact information in the event of a mental health crisis. Participants continued to supply daily SUDS ratings for the ten-week intervention period so that a comparison could be made with the SUDS ratings supplied during the baseline and post-intervention phases. Data analysis was ongoing throughout this phase and was used to identify any significant deterioration in a participant's mental health. The mental health and wellbeing outcome measures (DASS-21 and OQ-45.2), and uMARS were completed at the end of this phase.

Phase 4 (Post-Intervention): Participants provided SUDS ratings for at least 20 days following the completion of the intervention period. Once a minimum of 20 SUDS ratings were received, they again completed the DASS-21 and OQ-45.2 scales.

Results

Participant Characteristics and Descriptive Statistics

Ten participants started the study. Four withdrew after accessing the first online questionnaire which also contained information about the study. A further two participants withdrew after commencing the intervention period. Four participants remained for the entire study and provided adequate data to be included in the final analyses. The age range of the four finishing participants was 20 to 49 (M = 35.25, SD = 14.9); three were female; three reported co-morbid anxiety and depression, and one reported anxiety disorders only; two had very chronic illness of >11 years and two were receiving concurrent treatment (either psychotherapy or psychotropic medication); all were ambivalent in their motivation to comply with the app. The age range of the six participants that dropped out was 31 to 55 (M = 44.3, SD = 10.5), which was not significantly different to the mean age of the finishing participants ($t(9) = -1.14$, $p = .29$). Four were male; two reported co-morbid anxiety and depression, and four reported depression only; five had very chronic illness of >11 years and five were receiving concurrent treatment (either psychotherapy or psychotropic medication); one was strongly motivated and five were ambivalent in their motivation to comply with the app.

All participants reported average or better ability with technology. See Tables 7.1 to 7.3 for further information regarding the demographic and biographic features of all participants, including those who dropped out of the study.

Table 7.1

Demographic and Biographic Information for Participants that Finished - Participants 1 to 4

Variable	Participant 1	Participant 2	Participant 3	Participant 4
Age	49	47	25	20
Sex	M	F	F	F
Highest education	Uni	Uni	Secondary – Years 11-12	Secondary – Years 11-12
Mental health status	Depression, PTSD	Depression, panic attacks	Social anxiety, agoraphobia, generalized anxiety	Depression, social anxiety, panic attacks
Years with mental illness	11+	11+	6-10	1-5
Currently receiving counselling?	No	No	No	Yes - Psychiatrist
Currently on medication?	Yes - Antidepressant	No	No	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Somewhat disagree	Neither agree nor disagree	Neither agree nor disagree	Somewhat agree
Ability with technology generally	Average	Average	Good	Good

Table 7.2

Demographic and Biographic Information for Participants That Failed to Finish – Participants 5 to 7

Variable	Participant 5	Participant 6	Participant 7
Age	31	50	33
Sex	M	M	M
Highest education	Secondary – Years 11-12	Secondary – Years 11-12	Uni
Mental health status	Depression	Depression	Depression
Years with mental illness	1-5	11+	11+
Currently receiving counselling?	No	No	Psychologist
Currently on medication?	Yes - Antidepressant	No	No
I am motivated to do what the mobile app suggests?	Somewhat agree	Strongly agree	Neither agree nor disagree
Ability with technology generally	Average	Good	Average

Table 7.3

Demographic and Biographic Information for Participants That Failed to Finish – Participants 8 to 10

Variable	Participant 8	Participant 9	Participant 10
Age	43	54	55
Sex	F	F	M
Highest education	Uni	Secondary – Years 7-10	Secondary – Years 11-12
Mental health status	Depression	Depression, panic attacks, obsessive compulsive disorder	Depression, anxiety
Years with mental illness	11+	11+	11+
Currently receiving counselling?	No	Psychologist	Psychologist
Currently on medication?	Yes - Antidepressant	Yes - Antidepressant	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Somewhat agree	Somewhat agree	Neither agree nor disagree
Ability with technology generally	Excellent	Average	Good

Effectiveness of the App in Reducing Subjective Distress

Visual inspection of the plotted SUDS data for the four participants who completed the study revealed that all were experiencing noticeable feelings of distress at baseline. By post-intervention, three participants had achieved a reduction in subjective distress; two to a non-noticeable level (i.e., a rating of < 3). Table 7.4 shows the mean SUDS ratings per participant by phase; Figure 7.2 displays the continuous data. These data indicate that Participants 1, 3, and 4 had a relatively stable block of ratings during their baseline phase, a clear reduction in levels of distress during the intervention phase, and that this was maintained during the post-intervention phase. The graph for Participant 2 does not evidence improvement, with subjective distress remaining relatively stable from baseline to post-intervention. The finding is reflected in the mean SUDS scores shown in Table 7.4.

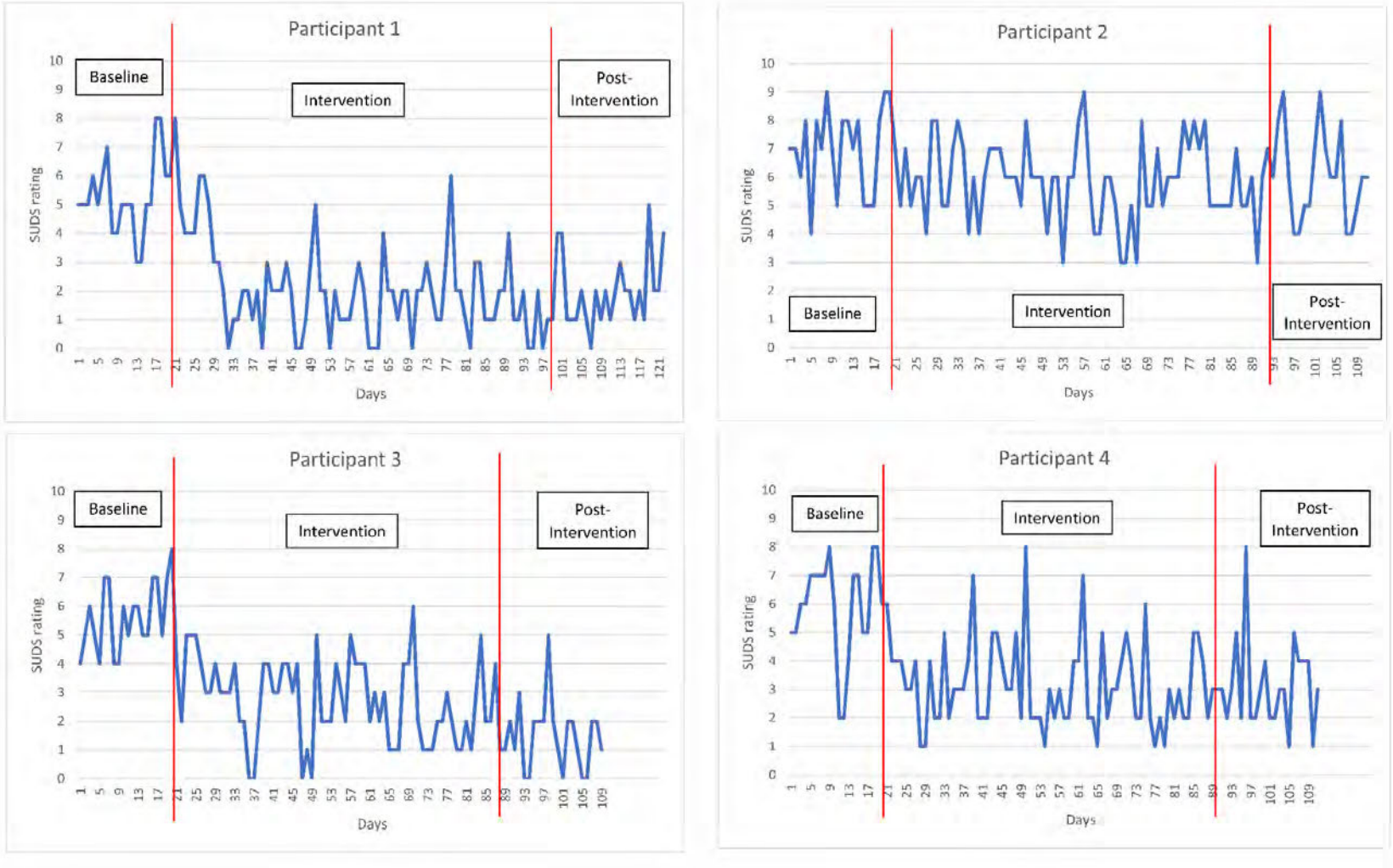
Table 7.4

SUDS Data Summary

Participant		Baseline SUDS	Intervention SUDS	Post-intervention SUDS
1	*M*	5.3	2.1	1.9
	SD	1.3	1.7	1.2
	Frequency (*N*)	20	77	25
2	*M*	7.0	5.8	6.1
	SD	1.5	1.4	1.6
	Frequency (*N*)	20	71	20
3	*M*	5.5	2.7	1.5
	SD	1.2	1.4	1.2
	Frequency (*N*)	20	67	22
4	*M*	5.9	3.1	3.1
	SD	1.7	1.5	1.6
	Frequency (*N*)	21	68	22

Figure 7.2

Participant SUDS Ratings Across Each Phase



Time-series analyses confirmed the findings observed through visual inspection of the plotted SUDS data. Using the statistical package, *R*, version 1.2.5033, an interrupted time series analysis (ITSA) used autoregressive integrated moving average (ARIMA) models to evaluate intervention effects on each participant's data. Autocorrelation effects were addressed using the augmented Dickey-Fuller Test (Mushtaq, 2011) and Ljung-box Q (Burns, 2002). The residuals in the models exhibited independence and normality.

The SUDS data are presented in Table 7.5 and can be matched to the relevant participant in Figure 7.2. Statistically significant decreases in SUDS ratings from baseline to post-intervention were observed for Participants 1, 3 and 4; however, Participant 2 did not show a statistically significant decrease in SUDS from baseline to post-intervention (see Table 7.5), which confirms the conclusion from visual inspection of Participant 2's graph in Figure 7.2. The change in SUDS ratings within the baseline and post-intervention phases for all participants was not significant compared to the overall change from baseline to post-intervention. This indicates that all four participants had a stable baseline phase prior to the intervention. It also indicates that all four participants had statistically stable post-intervention phases, suggesting that improvements in SUDS ratings during the intervention phase for Participants 1, 3, and 4 were due to the intervention and were maintained post-intervention.

Table 7.5

SUDS Time Series Analysis Summary

Participant	Baseline	Intervention	Post-intervention	Overall (baseline to post-intervention)
1	$t = 0.98$, $p = .34$	$t = -3.54$, $p = .001$	$t = 1.29$, $p = .21$	$t = -6.62$, $p < .001$
2	$t = 0.63$, $p = .54$	$t = -1.07$, $p = .29$	$t = -1.08$, $p = .29$	$t = -0.26$, $p = .80$
3	$t = 1.55$, $p = .14$	$t = -2.29$, $p = .03$	$t = -0.65$, $p = .53$	$t = -6.72$, $p < .001$
4	$t = -0.30$, $p = .77$	$t = -0.80$, $p = .43$	$t = 0.40$, $p = .97$	$t = -3.22$, $p = .003$

Effectiveness of the App for Reducing Anxiety

The severity of each participant's symptoms of anxiety, measured by the DASS-21 Anxiety subscale, are shown in Table 7.6. Participant 1 experienced normal levels of symptoms of anxiety throughout the study. Participant 2 commenced the study in the extremely severe range, but had improved to the moderate range at post-intervention. Participant 3 commenced the study in the mild range, but had worsened to be at the top of the moderate range at post-intervention. Participant 4 commenced the study in the extremely severe range, but had improved to be at the lower end of the moderate range by post-intervention. See Figure 7.3 below for a summary of the anxiety outcomes.

Table 7.6

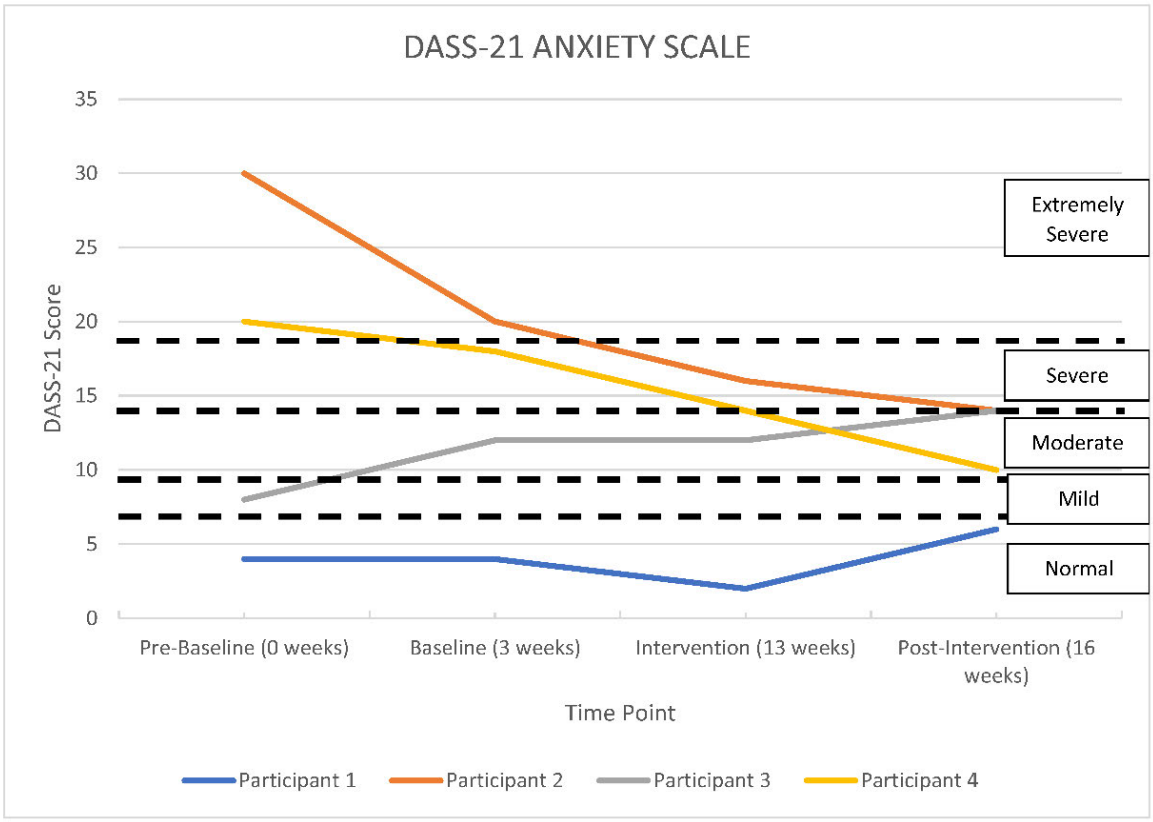
DASS-21 Scores for Participants 1 – 4

Participant	Scale	DASS Score at Beginning of Baseline	DASS Score at Beginning of Intervention	DASS Score at Beginning of Post-intervention	DASS Score at end of Post-intervention
1	Depression	34	26	18	16
	Anxiety	4	4	2	6
2	Depression	12	10	4	4
	Anxiety	30	20	16	14
3	Depression	4	6	8	10
	Anxiety	8	12	12	14
4	Depression	18	16	14	12
	Anxiety	20	18	14	10

Note. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).

Figure 7.3

DASS-21 Anxiety Scale Scores for Participants 1 – 4



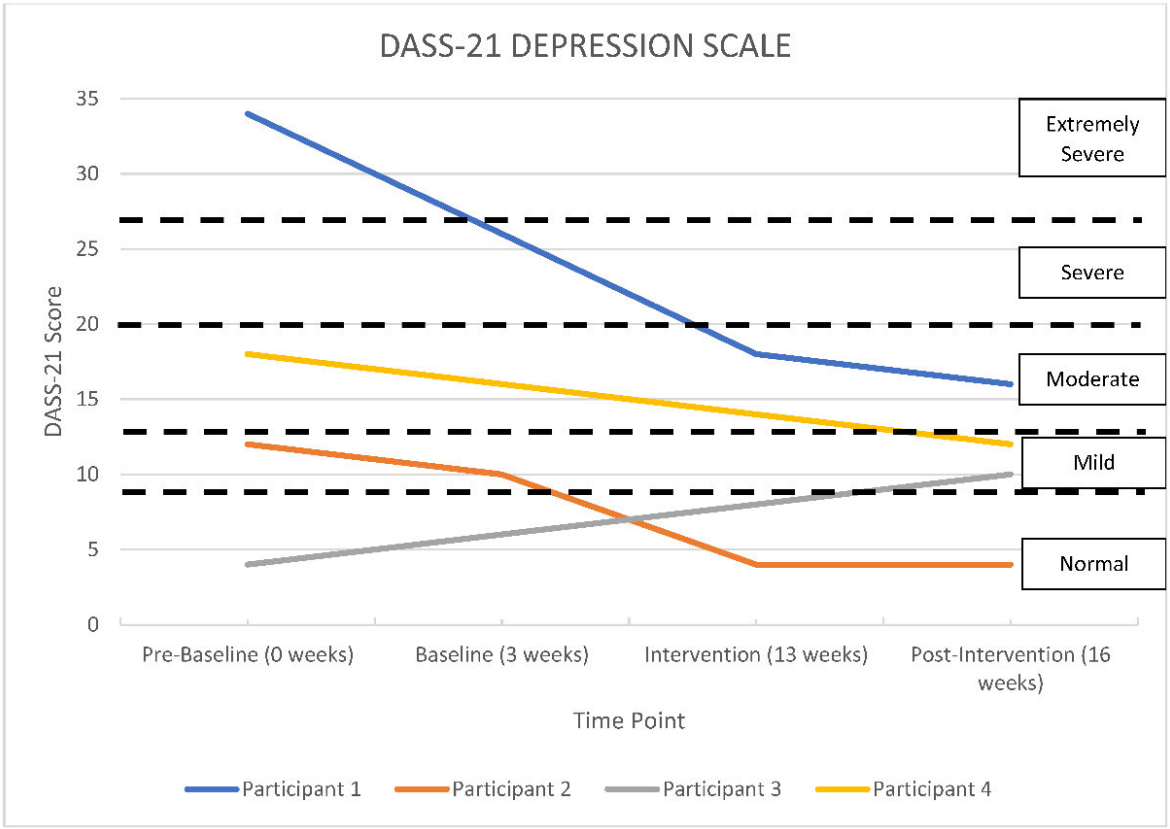
Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Effectiveness of the App for Reducing Depression

The severity of each participant's symptoms of depression, measured by the DASS-21 Depression subscale, are shown in Table 7.6. Participant 1 commenced in the extremely severe range, but had improved to be in the moderate range at post-intervention. Participant 2 commenced in the mild range, and had improved to be in the normal range at post-intervention. Participant 3 commenced in the normal range, but had worsened to be in the mild range at post-intervention. Participant 4 commenced in the moderate range, and had improved to be in the mild range at post-intervention. See Figure 7.4 for a summary of depression outcomes.

Figure 7.4

DASS-21 Depression Scale Scores for Participants 1 – 4



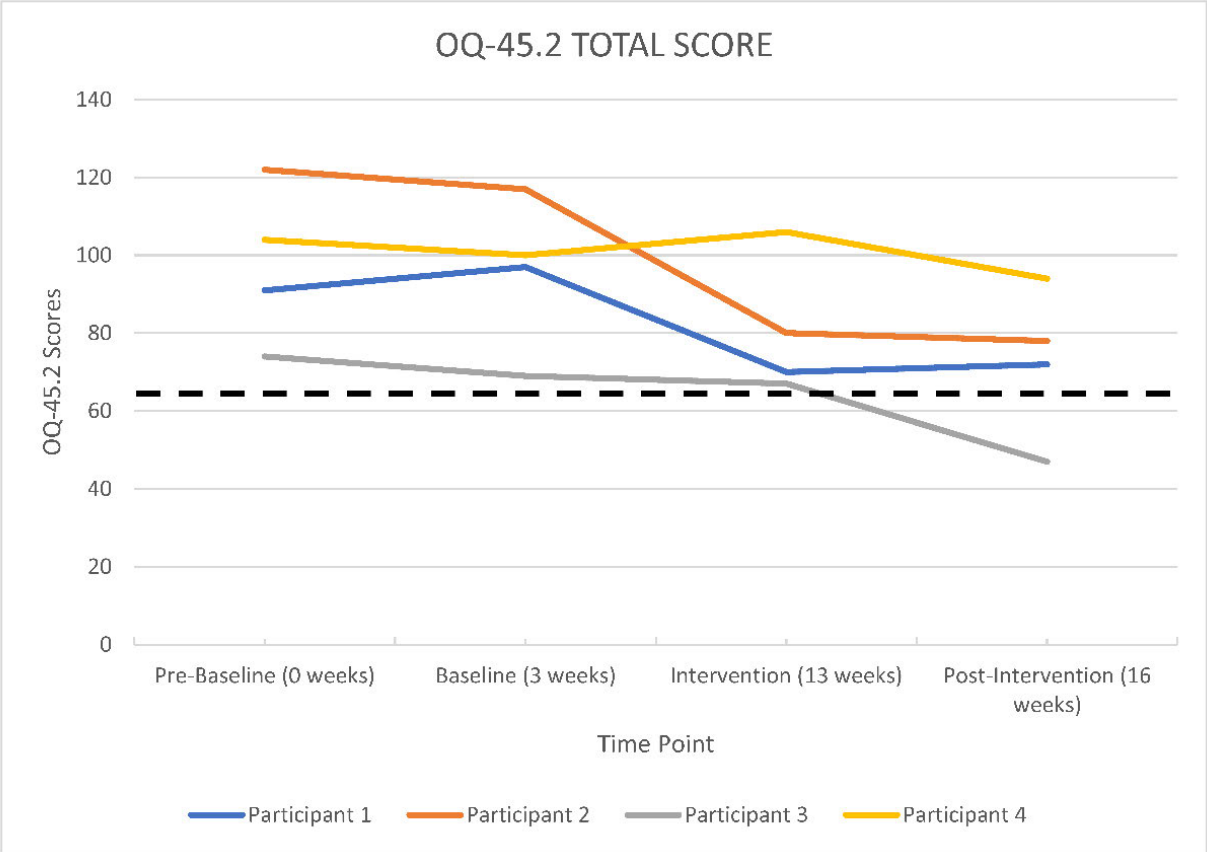
Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Effectiveness of the App in Improving Life Functioning

Each participant's overall functioning, measured by the OQ-45 Total Score, is shown in Table 7.9. As can be seen in Figure 7.7, at pre-baseline (Phase 1) all participants recorded a clinically significant impairment in life functioning. At 3-week follow-up (Phase 4), all participants were improved. The improvements shown by Participants 1, 2, and 3 were clinically significant, with Participant 3's scores dropping to the non-clinical range. See Table 7.9 and Figure 7.7 for a summary of life functioning outcomes.

Figure 7.5

OQ-45.2 Total Scores for Participants 1 – 4



Note. Clinical cut-off of 63 from the OQ-45.2 technical manual (Lambert et al., 2004).

Table 7.7

OQ-45.2 Total Scores for Participants 1 – 4

Participant	Pre-baseline to post-baseline	Post-baseline to post-intervention	Post-intervention to post-3-week follow up	Overall (pre-baseline to post-3-week follow up)	Classification (pre-baseline to post-3-week follow up)	Improvement (pre-baseline to post-3-week follow up)
1	(91-97) ^	97-70 *	(70-72) ^	91-72 *	Improved	20.88%
2	122-117 ^	117-80 *	80-78 ^	122-78 *	Improved	36.07%
3	74-69 ^	69-67 ^	67-47 *	74-47 *	Improved	36.49%
4	104-100 ^	(100-106) ^	106-94 ^	104-94 ^	Unchanged	9.62%

Note. Based on RCI = 14 and CSI = 63 for Total Score. ^ denotes not statistically reliable, not clinically significant; * denotes statistically reliable ($p < .05$); + denotes clinically significant; *+ denotes statistically reliable and clinically significant. Parentheses indicate change in a worsening direction. *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

Participant Factors that Impacted the Results

All participants who finished the study showed some level of improvement in one or more areas examined by the self-report measures. As reported above, the two factors that were associated with discontinuation in the study were a diagnosis of depression alone, and a longer duration of mental illness. Four of the six non-completers had depression only, and five of the six non-completers had their mental illness for longer than 11 years. There were no distinguishing characteristics between the four participants that dropped out during the baseline phase and the two participants that dropped out during the intervention phase.

App Ratings

Overall, participants who finished the study rated the app 3.25 stars out of 5. Participants gave the highest rating to the app's aesthetics ($M = 4.08$, $SD = 0.57$), followed by information ($M = 3.88$, $SD = 0.60$) and engagement ($M = 3.85$, $SD = 0.41$). The lowest average rating was for functionality ($M = 3.75$, $SD = 0.61$). All participants were more positive than negative about the app and its potential to assist, and all stated they would recommend it to others.

As a mental health intervention, the participants reported that the app helped in their awareness ($M = 3.75$, $SD = 0.95$), knowledge ($M = 3.50$, $SD = 0.58$), attitudes ($M = 4.00$, $SD = 1.16$), intention to change ($M = 4.25$, $SD = 0.96$), and help-seeking surrounding health behaviors ($M = 3.50$, $SD = 1.00$). Participants provided the highest overall ratings to the app's ability to increase their intention or motivation to address their mental health. (See Table B.5 in Appendix B for individual participant app ratings.)

Discussion

This study has evaluated the effectiveness of the mental health app, *SuperBetter*. Participants who completed the study were diverse in terms of age, level of subjective distress, symptoms of anxiety and depression, and impairment in life functioning. Even so, all

recorded a significant improvement in some aspect of their pre-intervention presentation indicating that using this mobile app may aid in improving the mental health of people with symptoms of anxiety and/or depression. However, in interpreting the effectiveness of the app, several potential limitations need to be considered (discussed later in this section), including that actual app usage by participants was unable to be recorded.

Answering the Research Questions

The present study sought to answer four research questions:

Is a mental health app (in this case,* SuperBetter*) effective at reducing subjective distress and clinically significant symptoms of anxiety and/or depression, and improving life functioning?

Given the results outlined above, it is evident that *SuperBetter* is effective at improving multiple dimensions of mental health and wellbeing. These results enhance the previous evidence obtained by Roepke et al. (2015) using a RCT design, and Worthen-Chaudhari et al. (2017) using a non-randomized controlled design.

Are there specific factors about the participants that impact on the results?

The *SuperBetter* app seems to have potential to offer a range of positive results over a broad cross-section of individuals. Interestingly, clues about potentially relevant participant characteristics emerged most clearly from an examination of the features of the six participants that dropped out of the study. These include gender, and the nature and chronicity of the mental health condition.

Three-quarters of finishers were female, whereas two-thirds of non-completers were male. This suggests that gender needs to be examined in closer detail to determine if being female is a predictor of engagement with the app, and hence greater effectiveness. The next factor is diagnosis.

No finishing participant had a sole diagnosis of depression. Rather, all finishers had anxiety, with three of them having comorbid depression. In contrast, four of the six non-completers had a sole diagnosis of depression. This suggests that *SuperBetter* may be more effective for people with anxiety and/or people with comorbid anxiety and depression rather than depression alone.

Finally, while 50% of finishers had a chronic condition, five of the six non-completers (83.33%) had a diagnosis of greater than 11 years duration, suggesting that like many interventions (Kessler et al., 2017; Lorenzo-Luaces et al., 2020), the mental health app could be more effective for those with conditions of shorter duration.

Can a single-case research methodology using digital technologies adequately and comprehensively assess the effectiveness of a mental health app in an applied, real-world setting? Can this methodology be scaled up to assess larger numbers of participants?

The findings of this study indicate that a single-case research methodology is able to provide nuanced information about the effectiveness of an app that is not available in an RCT. For instance, not only was it established that the app is effective in facilitating improvements in a number of domains (subjective distress, psychopathology symptoms, and life functioning) across a wide age range (20 – 49 years of age), this also occurred in a variety of contexts. These include: with individuals receiving concurrent psychotherapy (Participant 4); taking antidepressant medication (Participants 1 and 4); having co-morbid anxiety and depression (Participants 1, 2, and 4); and having relatively low levels of motivation (all participants). Therefore, not only can a single-case research design adequately and comprehensively assess the effectiveness of a mental health app in a real-world setting, but given the automated nature of the methodology, it could easily be scaled up to accommodate more participants with the same level of detailed information available.

What was the participants' experience of the app?

Although it is difficult to directly compare the results of app ratings by participants of this study to ratings from studies with a different experimental design, it would appear that the finishing participants rated the app at least moderately positively with an average rating of 3.25 stars out of 5. Participants gave the highest overall average rating to the app's aesthetics, suggesting high satisfaction with the look of the app. The lowest average rating was for functionality, suggesting less satisfaction with the practical useability of the app. All participants stated they would recommend the app to others. See Table B.5 in Appendix B for individual participant app ratings.

Strengths of the Present Study

This study was able to closely examine the outcomes of four diverse participants in response to using an app to improve their mental health. The multiple baseline across-individuals design allowed detailed information on participants to be considered when examining how effective an app had been at improving their mental health and wellbeing. In this case, the *SuperBetter* app was effective on a number of variables across a range of participants, providing evidence in support of existing research findings.

Limitations

The findings of this study should be read with the following limitations in mind. Firstly, there was a relatively large attrition rate of 60%. The study started with 10 participants and finished with four; however, this rate is lower than the 81% reported in the Roepke et al. (2015) study on *SuperBetter*. There is not enough information in the Roepke et al. article to determine if the participants who dropped out of that study shared similar characteristics to the participants that dropped out of this study. However, as Roepke et al. focused on participants with symptoms of depression, they may have shared that characteristic with the participants who dropped out of this study, four of which had

standalone depression and the other two having comorbid depression and anxiety. Participants who dropped out were not individually followed up. Therefore, in an effort to reduce attrition rates in the future, the researchers could be more assertive in following up participants who disengage from the study.

Secondly, the small number of participants means that generalizations need to be made in the context of the demographics of these participants. The evidence presented in this paper does not mean that the app will be effective for all consumers. However, it does suggest that the app can be effective for people aged between 20 and 49, even if they have some level of doubt about an app's ability to improve mental health and are not concurrently receiving psychotherapy.

Thirdly, it is unknown if the concurrent pharmacological and/or psychological therapy being received by participants remained stable over the intervention period. Therefore, it is unknown to what extent that improvements made during the intervention period were due to app usage, or due to changes in a participant's concurrent treatment.

Fourthly, the researchers were unable to use data directly from the app itself, and therefore were unable to know exactly how much time participants spent using the app which meant that assumptions about treatment fidelity cannot be made. Although the technology exists to do this, it would require the resources of a study with greater funding, and would have required the involvement and cooperation of app developers. This means that judgments about the extent that the app was responsible for any change in symptomatology or life functioning are limited. Furthermore, if access to app data was available, those participants that dropped out of the study would also have been able to contribute data that may have impacted on these findings. Having access to app data is a consideration for future research in this area.

Fifthly, by not incorporating SUDS data from participants who dropped out of the study, there is a risk that the sample is biased, and therefore that the time-series analysis results are skewed. Interpreting the time-series analysis results must be made with caution in light of this.

Sixthly, participants who dropped out of the study were not followed up in relation to reasons why they did so. If such participants were given the opportunity to describe why they dropped out, it may have shed further light on particular characteristics of the apps themselves and how these may detrimentally interact with participants who have specific characteristics.

Seventhly, while the applicability of the results may stretch across a wide variety of participant characteristics, the level of applicability may be limited when it comes to specific minority groups, such as First Nations peoples and culturally and linguistically diverse (CALD) communities. Specific research into such minority groups is needed to confirm the level of this acceptability towards those groups.

Eighthly, the survey instrument used to obtain demographic information from participants did not seek details relating to ethnicity, language, country of birth, and Indigenous status, thereby not allowing any investigation into participant characteristics related to First Nations peoples or CALD communities. Furthermore, the survey instrument contained a label with the heading “other mental health professional” which may have alienated professionals who identify as mental health nurses, peer workers, social workers, occupational therapists, and Indigenous health workers. Such omissions may have skewed the research methodology, the interpretation of data, and the reporting of results.

Ninthly, participants were only sought from Australia. By not seeking and including participants from other countries, the utility of the findings may be narrower as a consequence.

Finally, a validated brief symptom measure such as the Patient Health Questionnaire (PHQ-2 or PHQ-9) could have been used more regularly to monitor changes in symptomatology over time in the same fashion as the SUDS ratings. This would have provided additional data to the larger measures used less regularly and provided a useful comparison with the daily SUDS ratings.

Future Directions

The present research was a pilot study for a larger project using a multiple single-case design. It has shown that this methodology is useful in taking a more nuanced approach to examining the effectiveness of a mobile app. The methodology provides a number of possible avenues of future research through the detailed information obtained about participants. Future research also needs to establish how the principles of human computer interaction and design interact with participant characteristics to contribute to treatment outcomes.

Conclusion

Currently, searching the app stores for a mental health app that is efficacious and safe is problematic. There are a limited number of mental health apps with research evidence, and those that do are difficult to differentiate from the many more that do not. The methodology of the present study demonstrates that a single-case research design is able to provide useful evidence to compliment the evidence of efficacy obtained in larger RCTs. The evidence base for the *SuperBetter* app is further enhanced with the results obtained here, which have shown it to be effective to varying degrees across several domains, including subjective distress, symptoms of psychopathology, and life functioning. It is possible to scale this methodology so that it can be used for many other apps and increased participant numbers, with one of the ultimate aims being to give consumers greater choice of evidence-based mental health apps when searching the app stores. In much the same way that different psychological and pharmacological treatments for anxiety and depression are available to cater for differing

levels of efficacy and effectiveness across populations, a larger selection of mental health apps with demonstrated evidence are required to cater for the different needs of a cross-section of community members.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Interlude – The COVID-19 Pandemic: Research Progression to Chapter 8

Soon after the pilot study concluded, and shortly after the main study commenced, the world was struck by the COVID-19 pandemic. Many aspects of health were impacted and continued to be impacted at time of writing. Mental health was one such aspect that received global attention. In Chapter 8, the Australian response to COVID-19 from a mental health perspective, and the role played by digital mental health resources, is elucidated. This brings further context to the main intervention study of this thesis that is detailed in Chapter 10.

Chapter 8: The Role of Digital Mental Health Resources to Treat Trauma Symptoms in Australia During COVID-19

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). The Role of digital mental health resources to treat trauma symptoms in Australia during COVID-19. *Psychological Trauma: Theory, Research, Practice, and Policy, 12*(S1), S269-S271. http://dx.doi.org/10.1037/tra0000627

The Role of Digital Mental Health Resources to Treat Trauma Symptoms in Australia During COVID-19

How is the Situation in Australia Regarding the COVID-19 Pandemic?

Australia's health system is underpinned by Medicare which offers rebates for treatment of mental illness in the community, achieving more effective outcomes in comparison to many other countries (Hewlett & Moran, 2014). Preceding coronavirus disease 2019 (COVID-19), rebates for accessing psychological therapy were only available for in-person sessions, except for people in designated rural regions (Department of Health, 2019; Zhou et al., 2020). On the 30th of March 2020, in response to COVID-19, the Australian Federal Government announced that mental health professionals who were authorized to deliver Medicare-rebated therapy could also deliver sessions via telehealth for every Australian no matter where they lived (Department of Health, 2020).

How is the Pandemic Affecting the Population from a Mental Health Perspective?

Demand for online mental health services has increased on a global scale since the spread of COVID-19 (Liu et al., 2020) and there may be further increases in trauma presentations in its aftermath. In Australia, which has a large number of refugees and immigrants who have fled war-ravaged countries, some mental health clinicians have experienced an increase in people with previous trauma presenting with new symptoms in reaction to COVID-19, thus causing re-traumatization (Dunn, 2020). As telehealth sessions are not necessarily accessible for all people due to the increased demand, many are turning to other digitized, automated options such as mobile applications (apps). Mental health apps are potentially one way for people to access trauma-related treatment to lessen the burden on primary health care (Azarang et al., 2019) during the COVID-19 crisis.

How do People Respond to the Situation in Australia?

Downloads of mental health apps have risen exponentially since the proliferation of COVID-19 (Basu, 2020; Heilweil, 2020). *Calm* and *Headspace* ranked two and three for worldwide revenue achieved in March 2020 for health and fitness apps in the Google Play store, with over US$1,149,000 and US$838,000 respectively (Statista, 2020). Add to this that 89% of the Australian population own a smartphone (Deloitte, 2018), and it is not surprising that mental health apps are popular during the COVID-19 pandemic.

This has seen Australians accessing digital information on mental health apps, online programs, online counselling options, and other digital resources. A number of Australian government and non-government websites contain this information, all of which have added content related to COVID-19. Examples include: Beyond Blue; Black Dog Institute; eMHprac; Headspace; Head To Health; Kids Helpline; Lifeline; Mensline Australia; Mental Health Australia; Reachout; Sane Australia; Suicide Call Back Service; and others (see Table 8.1 for URLs). Although it is still too early, at time of writing, to quantify the mental health of Australians during the COVID-19 pandemic with accurate statistics, the Lifeline crisis support service reported in the week beginning 15th March that 23% of callers discussed COVID-19 (Lifeline, 2020), and a week later this had jumped to 39% (Medhora, 2020). Furthermore, online forums and resources relating to COVID-19 offered by many of the above-mentioned websites had thousands of visitors, such as the Beyond Blue COVID-19 forum which received 26,000 views in its first week (Medhora, 2020). All of which points to Australians turning to digital resources to deal with increased mental health-related concerns as a result of COVID-19.

What is Helpful and What is Less Helpful in Dealing With the Situation?

The utilization of these digital mental health resources may have positive consequences for Australia during the COVID-19 pandemic. By turning to digital options for less severe mental health presentations, it may free up primary care services to deal with

more acute health situations without clogging up waiting rooms and risking the further spread of infection between people. Telehealth therapy, which has an established and still-growing evidence base for treating trauma-related symptoms (Sloan et al., 2011), may provide some protection for both therapists and clients from becoming infected with COVID-19. Furthermore, the wide availability of digital mental health resources to the Australian population may promote resilience and wellbeing on a wider community level as mental health information is disseminated widely and potentially destigmatizes mental illness, thus allowing more people to accept treatment offered through digital sources. What is less helpful is that others may develop so-called digital mental health resources that are not necessarily developed using an evidence-based theoretical framework (Marshall et al., 2020a) or with expert input (Alyami et al., 2017; Shen et al., 2015), and this is where government intervention is needed.

How is Health Care Currently Organized?

The Australian Federal Government has considered digital mental health services to be a priority area for future prevention and treatment of mental health issues for at least the last decade (Department of Health and Ageing, 2012). This has resulted in the Federal Government proactively supporting the use of apps and other digital options to assist in managing mental health. The Government is also attempting to provide more effective oversight in this area (Australian Commission on Safety and Quality in Health Care, 2020) during a time when many governments are struggling to develop suitable regulations around digital mental health resources (Wang et al., 2018). In Australia, this has meant the development of a number of mental health apps, some of which have had government funding, that have published evidence of efficacy e.g. *MoodMission* (Bakker et al., 2018a, 2018b), *Smiling Mind* (Flett et al., 2019), and *myCompass* (Proudfoot et al., 2013). This has occurred at a time when the vast majority of mental health apps currently publicly available

around the world do not have any published evidence of efficacy or effectiveness (Marshall et al., 2019). By being a leader in the development and promotion of digital mental health resources, Australia is well placed to deliver a model that potentially offers effective digital mental health solutions that may ease the burden on primary health care services during the current crisis. As part of their national television information campaign about COVID-19, the Australian Federal Government is promoting the importance of mental health and wellbeing (e.g. https://www.youtube.com/watch?v=rEzv1d0OpiE). When the crisis is over, it will be interesting to see the statistics on the use of digital mental health resources in Australia in an effort to understand if they were effective in managing the population's mental health in comparison with other countries whose governments and culture do not recognize the importance of mental health in the same proactive way.

Table 8.1

Australian Websites Providing Information and Online Treatment Related to Trauma and General Mental Health

Resource	URL
Beyond Blue	https://www.beyondblue.org.au/home
Black Dog Institute	https://www.blackdoginstitute.org.au/research/digital-dog/what-is-digital-dog
eMHprac	https://www.emhprac.org.au/
Headspace	https://headspace.org.au/
Head To Health	https://headtohealth.gov.au/
Kids Helpline	https://kidshelpline.com.au/
Lifeline	https://www.lifeline.org.au/
Mensline Australia	https://mensline.org.au
Mental Health Australia	https://mhaustralia.org/
Reachout	https://au.reachout.com/tools-and-apps
Sane Australia	https://www.sane.org/
Suicide Call Back Service	https://www.suicidecallbackservice.org.au/

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Interlude – The COVID-19 Pandemic: Research Progression to Chapter 9

In Chapter 8, the Australian response to COVID-19 from a mental health perspective, and the role played by digital mental health resources, was described. In Chapter 9, the global use of smartphone mental health apps during the COVID-19 pandemic is illustrated, focusing on the widespread shortage of in-person mental health services to deal with the increased demand due to heightened community distress. The COVID-19 situation is also a major test of the ability of mental health apps to fulfil their potential of providing effective treatment for mental health issues.

Chapter 9: Treating Psychological Trauma in the Midst of COVID-19: The Role of Smartphone Apps

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020). Treating psychological trauma in the midst of COVID-19: The role of smartphone apps. *Frontiers in Public Health, 8*, Article 402. https://doi.org/10.3389/fpubh.2020.00402

Abstract

With the COVID-19 pandemic confronting health systems worldwide, medical practitioners are treating a myriad of physical symptoms that have, sadly, killed many thousands of people. There are signs that the public is also experiencing psychological trauma as they attempt to navigate their way through the COVID-19 restrictions impinging on many aspects of society. With unprecedented demand for health professionals' time, people who are unable to access face-to-face assistance are turning to smartphone apps to help them deal with symptoms of trauma. However, the evidence for smartphone apps to treat trauma is limited, and clinicians need to be aware of the limitations and unresolved issues involved in using mental health apps.

Treating Psychological Trauma in the Midst of COVID-19: The Role of Smartphone Apps

Introduction

Although many medical and allied health professionals are conducting telehealth sessions with patients and clients during the COVID-19 pandemic, the increased level of demand means that some people may not be able to access services in an adequate timeframe (Liu et al., 2020). In response, increasing numbers of sufferers are turning to digitized, automated options such as mobile applications (apps) (Basu, 2020). For people with symptoms of acute stress disorder, which if present for longer than a month is reclassified as posttraumatic stress disorder (PTSD; American Psychiatric Association, 2013), mental health apps are potentially one way to access treatment and lessen the burden on primary health care.

It is no surprise that individuals are turning to digital options – over 5.2 billion people worldwide own a smartphone (Barboutov et al., 2017). When the COVID-19 situation rapidly worsened, downloads of mental health apps accelerated (Basu, 2020; Heilweil, 2020). For example, the apps *Calm* and *Headspace* ranked two and three for worldwide revenue achieved in March 2020 for health and fitness apps in the Google Play store, achieving sales worth over US$1,149,000 and US$838,000 respectively (Statista, 2020). Around the world, there is other evidence that both authorities and people are turning to apps and other digital options in large numbers to cope with the trauma of COVID-19: in Australia (Koh, 2020), China (Feng, 2020), India (Mukherjee, 2020), New Zealand (Mindfood, 2020), U.K. (Chowdhury, 2020), U.S. (Heilweil, 2020), and others. However, we do not yet have accurate

data on how many people are experiencing symptoms of PTSD as a result of COVID-19, nor how many are relying on smartphone apps to cope with these symptoms.

The Evidence for Treating Symptoms of Trauma with an App

Best practice involves initially treating symptoms of trauma with specially administered psychological therapy, such as trauma-focused cognitive-behavioral therapy (CBT), eye movement desensitization and reprocessing (EMDR), cognitive processing therapy, narrative exposure therapy, or prolonged exposure therapy (Charney et al., 2018). If there is no or little improvement in symptoms as a result of these treatments, clinicians may explore pharmacological options, either as an adjunct to therapy or as a front line treatment if psychotherapy has been ineffective (Psychotropic Expert Groups, 2013). If psychotropic medication is prescribed for PTSD, it will usually initially be an antidepressant, and mostly it will be a selective serotonin reuptake inhibitor (SSRI) or serotonin-norepinephrine reuptake inhibitor (SNRI) (Charney et al., 2018). However, the evidence for efficacy of pharmacological treatments of PTSD is less than that for psychological treatments (Psychotropic Expert Groups, 2013), so further treatment options would be welcomed, especially with the anticipated worldwide surge in trauma-related presentations that may arise during and after the COVID-19 situation with both health workers and the general public (Australian Government Department of Health, 2020; Carthaus, 2020; Kallivayalil, 2020; Mougin, 2020).

The description of "antidepressant" for a class of drug is confusing because many antidepressants not only treat depression, but also anxiety (Jennings, 2018). When pharmacological treatment is prescribed for anxiety or depression in the form of an antidepressant, it has been done so for a medication that has been approved by government regulators in that country that have identified the medication as safe and efficacious. A digital antidepressant, therefore, can be thought of as any app, website or other digital tool that is

specifically designed to treat symptoms of anxiety or depression. Therefore, if this digital tool is a form of non-pharmaceutical medication, there is a need for these digital tools to have research behind them. Government regulators would not allow a pharmaceutical antidepressant to become available without research of efficacy, and using the term “digital antidepressant” is a reminder of how important it is to consider apps that claim to treat anxiety and depression symptoms as things that need to have research for their efficacy as well. The idea that a health professional could “prescribe” a mental health app (Byambasuren et al., 2018) is attractive at this time because of the difficulties accessing face-to-face treatment, whether by telehealth or in-person. Medical practitioners are concerned about the welfare of their patients, and if their patients cannot access “human” help, they may seek out other options in the form of digital technology. However, clinicians should be aware of the evidence for mental health apps, and understand that a digital antidepressant may not produce the desired clinical outcomes.

Similarly, the above-mentioned evidence-based psychotherapies for treating symptoms of trauma require the specialized skills of highly trained practitioners. To be effective, such practitioners have to be experts in recognizing the signs and symptoms of trauma, adapting their psychotherapy in response to changes in client presentation, and acting appropriately if their client's condition deteriorates, especially in response to a risk of self-harm or suicide. If a general practitioner or other health professional does not have such specialized training in the area of psychologically treating symptoms of trauma, they may have the mistaken belief that a mechanized version of psychotherapy is possible in an app without being aware that evidence may not exist for the app's effectiveness in treating trauma presentations.

Evidence for the efficacy and effectiveness of mental health apps is limited (Marshall et al., 2020b; Orman & O'Dea, 2018). The research suffers from methodological deterrents (e.g., quality randomized controlled trials [RCTs] can be expensive and take years to run

which is an impediment for the profit-driven app sector), heterogeneity across studies, no published replication studies to speak of, and a lack of independence (i.e., studies completed by researchers who have not had any association with the app) (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). This last point is important and is illustrated through a comparison with chemical antidepressants.

The early evidence base for chemical antidepressants, which was largely established by pharmaceutical companies who developed these medications, showed far higher effect sizes and greater levels of statistical significance than more recent studies (Cipriani et al., 2018). The later studies have included a much greater proportion of independent trials by researchers who have no association with the medications being tested. Their results have demonstrated significantly less efficacious and effective outcomes (Cipriani et al., 2018). Using the Cochrane Bias Tool, it has been estimated that 82% of all previous published studies on antidepressant medications are at moderate or high risk of bias due to the involvement of pharmaceutical companies (Cipriani et al., 2018).

If the efficacy of mental health apps is to be free of the limitations affecting many drug trials, unbiased research is required while their development is still relatively young. In a recent review of the two major app stores, only 1% of apps that claimed to offer a therapeutic treatment for anxiety and depression had independent research to back-up claims of efficacy (Marshall et al., 2019). That is, research that was conducted by individuals, institutions or organizations who were not involved in the development of the app, and who would not stand to gain financially or otherwise from it. This is not to discourage research by app developers – quite the contrary. More app developers also need to conduct research on their product because a huge proportion of publicly available apps have no research support whatsoever (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). While much of the research conducted by those who have an

association with the app may be of acceptable quality, it is important that independent research in this area is increased to further legitimize the evidence base and minimize concerns regarding bias.

There is less evidence for apps specifically treating symptoms of trauma. The principal research on this front has come from the U.S. Department of Veterans' Affairs (VA), developers of *PTSD Coach* (Kuhn et al., 2017) and a suite of other apps, most of which are specifically aimed at veterans or their families. At time of writing, no published evidence could be found for the efficacy or effectiveness of any individual app specifically designed to treat symptoms of trauma, other than those produced by VA (Owen et al., 2018; Sander et al., 2020). This is despite hundreds of publicly available apps purporting to do this (Rodriguez-Paras et al., 2017; Sander et al., 2020), and the availability of a standardized framework for developing PTSD-focused apps that stipulates the importance of demonstrated efficacy (Schellong et al., 2019). Without a more diverse evidence base, it is unknown to what extent people benefit when they attempt to manage trauma by using apps. Worse still, we do not know how much damage is being caused by misinformed and poorly developed apps.

Government Regulation

Examples of potential harm from an app include breaches of privacy, misuse of personal data, providing inappropriate advice (Hussain et al., 2015), and poor app functionality leading to possible app failure at a critical emotional point for the user. It has also been found that for over 20% of publicly available apps claiming to treat symptoms of PTSD, their app store descriptions did not contain any specific PTSD evidence-based content (Sander et al., 2020). This is where government intervention is urgently required, but many governments are struggling to develop suitable regulations around mental health apps (Armontrout et al., 2018). Authorities such as the U.S. Food and Drug Administration are focused on removing apps that may cause harm (U.S. Food and Drug Administration, 2019).

A recent report by the Australian Commission on Safety and Quality in Health Care (2020) attempts to encompass a wider regulatory view of digital mental health resources founded on a "model of care" with "best available evidence and best practice" (p.20). While regulation of mental health apps may also have some disadvantages, such as the potential to stifle the development of new apps if developers are reluctant or unable to meet imposed standards, and conflicting regulations across territories meaning difficulties with apps being available across country borders, the potential advantages of considered regulation outweigh the disadvantages, especially in terms of providing greater numbers of quality, evidence-based apps.

However, these new standards in digital mental health care, which include mental health apps, are yet to be implemented. In the meantime, medical practitioners who are expecting an influx of trauma-related presentations in the wake of COVID-19, must use caution in directing patients towards apps to assist in managing symptoms of trauma.

Discussion

There is uncertainty about the effectiveness of many mental health apps (i.e., their ability to deliver beneficial treatments in a real-world setting) and the most appropriate methods of examining this (Lui et al., 2017). The COVID-19 crisis brings to the fore the need for a centralized database of information for use by medical professionals, governments, therapists, researchers and consumers (Marshall et al., 2020e). Clearly, there is a need to move beyond reliance on app store ratings and reviews, which may be inaccurate, ill-informed, or fake (Xie & Zhu, 2015), to the requirement for app researchers to provide accessible and timely research evidence. While the time demands of the gold standard RCT can be an impediment to research, other methodologies that do not sacrifice scientific rigor and integrity can potentially be conducted on apps in a more timely manner (Clough & Casey, 2015a). These include scalable single-case designs involving practicing clinicians

working with researchers (Marshall et al., 2020e). In such a model, clinicians could contribute their findings to a centralized database that may continually be updated with results that occur from individuals using their app in real-world settings. The single-case methodology has the added advantage of being able to provide potentially more information on the characteristics of the individual than might otherwise be identified in larger RCT designs. This may in turn lead to more informed hypotheses about how individual characteristics may impact on the effectiveness of a mental health app.

Other than the need for further research on efficacy and effectiveness, future development of apps for treating symptoms of trauma (and indeed for all apps treating mental illness generally) needs to take into account a number of factors, and there are many existing development blueprints that app developers can refer to (Bakker et al., 2016; Moshi et al., 2018; Zelmer et al., 2018). Building a mental health app on the foundations of an evidence-based framework is vital (Marshall et al., 2020a). There are several evidence-based frameworks that inform PTSD psychological treatments (as mentioned above), and it would seem plausible that such treatment interventions could be incorporated into an app. There needs to be expert input from qualified clinicians and/or researchers into the development of a mental health app – many apps claiming to treat symptoms of mental illness do not have such input (Alyami et al., 2017; Marshall et al., 2019; Shen et al., 2015). Given that the development of mobile mental health apps is still in its infancy, we are still not certain about the mechanisms of action of such apps, and therefore the level of importance of characteristics such as app design and usability is still being investigated. However, it would seem plausible to assume that a mental health app has to be easy to use, engaging, and aesthetically pleasing to be efficacious and effective (Bakker et al., 2016), and therefore having the input from experienced app designers would seem to be a necessity. This is only a brief summary of necessary aspects of a successful mental health app, but even here it can be

seen that many different aspects of development need to be considered before an app is able to make claims that it can successfully treat symptoms of trauma, and of any other mental illness.

Another concerning feature of around 40% of mental health apps is that they lack a publicly accessible privacy policy (Parker et al., 2019). Given the sensitive nature of people's mental health information, app developers need to pay more attention to this. One reason why app developers have been able to get away with this for so long is that there is no government regulation about the need for privacy policies for digital mental health resources. If government authorities can broaden and strengthen their oversight of this sector, protecting people's privacy will be as necessary a factor as ensuring that mental health apps do no harm to their users.

There are websites that clinicians (and consumers) can access for reviews and further information about choosing mental health apps, listed in Table 9.1. The rationale for providing the information in Table 9.1 is so that clinicians can become more informed about recommending appropriate apps for their patients and clients. We did not recommend specific apps, as that is not the purpose of this paper. Although lacking information on how to measure effectiveness, these websites are nevertheless useful resources for clinicians who need assistance identifying potentially suitable apps. While the current evidence base is lacking, it is hoped the COVID-19 crisis is a potential catalyst for ensuring that mental health apps have demonstrated effectiveness for treating specified mental health disorders, including PTSD-related trauma, into the future.

Table 9.1

Website Resources for Clinicians that Provide Reviews and Further Information on Mental Health Apps for Symptoms of Trauma

Resource	Summary
American Psychiatric Association's App Evaluation Model (https://www.psychiatry.org/psychiatrists/practice/mental-health-apps/app-evaluation-model)	U.S.-based framework for clinicians and researchers on how to conduct their own evaluation of apps
Anxiety and Depression Association of America (https://adaa.org/finding-help/mobile-apps)	U.S.-based non-government, non-profit organization providing reviews of mental health apps by volunteers with recognized mental health qualifications who do not have any association with the apps being rated
beacon (https://beacon.anu.edu.au/)	Australian university website providing research summaries of digital health resources, including mental health apps
Head To Health (https://headtohealth.gov.au/)	Australian government website providing information regarding digital mental health resources, including apps

Health Navigator (https://www.healthnavigator.org.nz/)	New Zealand government-supported website, with input from professional health-related bodies, providing information regarding digital health and medical resources, including mental health apps
mHabitat (https://wearemhabitat.com/)	U.K. government-supported website, with the involvement of various departments of the National Health Service, dedicated to developing partnerships with developers of digital health solutions, including mental health apps
MindApps (https://mindapps.dk)	A mental health app review website by The Centre for Telepsychiatry, Psychiatry in the Region of Southern Denmark. It includes reviews by therapists, academics, and consumers
NHS Apps Library (https://www.nhs.uk/apps-library/)	Coordinated by the U.K.'s National Health Service, this tool allows users to search for all types of health apps, including mental health apps, with summaries about what the app does, links to the app's website, and links to the App Store and/or Google Play for download.
The Organisation for the Review of Care and Health Apps (ORCHA) (https://www.orcha.co.uk/who-we-help/health-and-care/)	Private organization based in the U.K. offering a number of tech-related health services, including reviews, accreditation, curation and prescriptions services for health and mental health apps

reachout.com (https://au.reachout.com/tools-and-apps)	Australian non-government organization providing expert and consumer reviews on mental health apps
PsyberGuide (https://psyberguide.org/)	U.S. non-government organization providing expert reviews on mental health apps

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Research Progression to Chapter 10 – Study 4

In Chapter 7 – Study 3, the pilot study demonstrated that the methodology was appropriate and effectual at examining the effectiveness of a single mental health app. Chapters 8 and 9 illustrated the widespread impact of COVID-19 and how digital resources, particularly smartphone apps, were playing a part in managing global mental health. In Chapter 10 – Study 4, the single-case research methodology is scaled up to examine five mental health apps simultaneously, and is conducted with the serendipitous opportunity to examine each app's effectiveness in the context of COVID-19 and the resultant global increase in mental ill-health. Furthermore, the scaled-up study implemented the strategy of assertive follow-up of participants at risk of dropping out of the study, which was one of the lessons learned from the pilot study. The methodology was able to identify how participants' daily subjective distress was affected by COVID-19. It also provided evidence on how effective the apps were at alleviating symptoms of mental illness during this heightened state of community distress, thus adding further value to this research and the thesis as a whole.

Chapter 10 – Study 4: Smartphone Psychological Therapy During COVID-19: A Study on the Effectiveness of Five Mental Health Apps Using a Multiple Baseline Across-Individuals Design

A manuscript for submission to a peer-reviewed journal is in preparation.

Abstract

The aims of this study were to examine the effectiveness of a range of smartphone apps with demonstrated efficacy for managing symptoms of anxiety and depression; and, to assess the utility of a single-case research design for enhancing the evidence base for this mode of treatment delivery. The study was serendipitously impacted by the COVID-19 pandemic, which allowed for effectiveness to be additionally observed in the context of significant community and individual distress. Thirty-nine participants started the study (27 females and 12 males, M_{Age} = 34.04 years, *SD* = 12.20); with 29 finishing the intervention phase and completing post-intervention measures. This study used a digitally enhanced, multiple baseline across-individuals single-case research design to examine the effectiveness of five mental health apps, all with evidence of efficacy in treating anxiety and/or depression, but using a diverse range of theoretical approaches. Participants were randomly assigned to the following apps: *SuperBetter* (*n* = 8), *Smiling Mind* (*n* = 7), *MoodMission* (*n* = 8), *MindShift* (*n* = 8), and *Destressify* (*n* = 8). Mental health symptoms and life functioning were measured at five different time points corresponding to the commencement of the phases of the study: pre-baseline/screening, baseline, intervention, and three-week post-intervention. In accordance with multiple baseline designs, all participants commenced simultaneously and functioned as their own control. Individuals entered the intervention phase in a staggered manner after a baseline period of providing ratings in subjective distress, which was

measured daily across the phases via SMS text messages. The user version of the Mobile Application Rating Scale was used to analyse individual perceptions of the apps, including a final overall 5-point rating similar to the method used by the major app stores. Data were analysed using visual inspection, time-series analysis, and methods of statistical and clinical significance pioneered by Jacobson and Truax (1991). Positive results were observed for all apps irrespective of the theoretical approach employed. Within each app group, some participants gained statistically significant reductions in subjective distress and symptoms of anxiety and/or depression, as well as improved life functioning. Overall, more favourable outcomes were achieved by younger participants, those concurrently undertaking psychotherapy and/or psychotropic medication, and with a shorter history of mental illness. It was concluded that a diverse range of evidence-based therapies offered via apps can be effective in managing mental health and improving life functioning even during times of significant global unrest and uncertainty and, like all psychotherapies, are influenced by client features. Additionally, a single-case research design is a low cost / high value means of assessing the effectiveness of mental health apps and has the potential to provide efficacy data as well.

Keywords: mHealth; mental health apps; single-case designs; smartphones; COVID-19.

Smartphone Psychological Therapy During COVID-19: A Study on the Effectiveness of Five Mental Health Apps Using a Multiple Baseline Across-Individuals Design

Currently, there are over 10,000 mental health apps publicly available (Torous, Firth, et al., 2018), but most of these have not been developed using established theoretical frameworks (Marshall et al., 2020a), or by recognised mental health experts (Alyami et al., 2017; Shen et al., 2015). Additionally, the vast majority of the comparatively few apps with research evidence of efficacy are not further supported by additional studies by researchers unaffiliated with the app or in diverse samples of participants (Marshall et al., 2019). In the interests of public safety and greater understanding of the usage of individual apps, more research needs to be carried out (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Lui et al., 2017). To achieve this, a low cost/high value research design is required (Clough & Casey, 2015a; Marshall et al., 2020c).

Mental Health Apps as Mechanised Psychotherapy

Best practice for treating symptoms of anxiety and depression depends on the individual's unique presentation and will involve evidence-based psychotherapy and/or antidepressant medication (American Psychiatric Association, 2009, 2010; Cuijpers et al., 2013). Widely used evidence-based psychotherapies include: cognitive-behavioural therapy (CBT) (Beck et al., 1979; Bennett-Levy et al., 2009; Bennett-Levy et al., 2010; Butler et al., 2006), interpersonal therapy (IPT) (de Mello et al., 2005), dialectical behaviour therapy (DBT) (Lynch et al., 2006), acceptance and commitment therapy (ACT) (A-Tjak et al., 2015; Vollestad et al., 2012), and positive psychology interventions (Seligman & Csikszentmihalyi, 2000; Seligman et al., 2005).

Research has shown that a number of factors influence the prognosis and outcomes of psychotherapy. These include client-therapist rapport (Tang & DeRubeis, 1999), client motivation (Addis & Jacobson, 2000), chronicity/history of mental illness (Hamilton & Dobson, 2002), functional impairment, social support, coping style, level of client resistance, subjective distress, and readiness to change (Beutler & Clarkin, 1990). Such factors combined with appropriate treatment may account for over 90% of the variance in successful outcomes (Beutler et al., 1999). In terms of appropriate treatment, while CBT is effective in treating depression, IPT may be more useful in circumstances where the precipitating factor is an interpersonal relationship issue (Zhou et al., 2017). Similarly, positive psychology approaches may be more applicable when highly motivated, older individuals wish to focus on strengths and positive interventions to maximise their psychological wellbeing (Sin & Lyubomirsky, 2009). In this way, positive psychology strategies may complement rather than replace traditional CBT approaches (Harvard Medical School, 2008). Overall, it is reasonable to assume that individual participant characteristics, the treatment approach, and the participant's perceived engagement with a mental health app will influence clinical outcomes. It is these types of influences that are examined in effectiveness research (Singal et al., 2014).

The Importance of Effectiveness Research

The current evidence base supporting the use of mental health apps for anxiety and depression includes a small number of studies of efficacy and even fewer of effectiveness. In clinical psychology, efficacy studies occur under controlled conditions where participants are screened for their suitability to improve the homogeneity of the experimental group, whereas effectiveness studies are designed to measure interventions in "real world" clinical settings with more heterogeneous populations (Kazdin, 2017). An intervention that has been found to be efficacious also needs to demonstrate effectiveness in clinical practice (Singal et al., 2014). An efficacy trial may inflate an intervention's clinical impact (effect size) in practice,

therefore it is important for treatments to have demonstrated effectiveness in this context. Although proven efficacy increases the chances of observing an intervention effect if one exists, effectiveness research accounts for individual clinician, client, and process characteristics that may moderate an intervention's effect (Singal et al., 2014).

If the research on mental health apps is to be free of the limitations of inflated effect sizes found in efficacy studies, effectiveness studies are required. In a recent review of the two major app stores, only 1% of apps that claimed to offer a therapeutic treatment for anxiety and depression had independent research to back up their assertions (Marshall et al., 2019); thus, the majority of apps do not have the research evidence needed to inform individuals or clinicians (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). If the proportion of research of both efficacy and effectiveness was increased, it could boost the chances of mental health apps achieving widespread acceptance and validation by consumers and clinicians alike.

Previous Research on the Apps used in this Study

All five apps used in the present study have published evidence of efficacy in peer-reviewed journals.

Roepke et al. (2015) conducted an RCT with 283 participants with a mean age of 40 on the positive psychology-based app, *SuperBetter*. The study focused on symptoms of depression. They were recruited from the Penn Authentic Happiness website and the Craigslist.org community bulletin board. Participants were instructed to use the app for 10 minutes per day for one month. Statistically significant improvements in levels of anxiety and depression were reported for users of *SuperBetter* compared to the waitlist control group.

Flett et al. (2019) had one control group (n = 67) and two intervention groups of New Zealand undergraduate students with a mean age of 20.1 years. The intervention groups used either the *Smiling Mind* (n = 58) app or another app to provide a mindfulness-based

treatment, with instructions to use the app for 10 minutes per day for the first 10 days, then use it at the participant's discretion for the next 30 (for a total of 40 days). *Smiling Mind* produced a statistically significant reduction in anxiety, but did not produce a statistically significant reduction in depression.

Bakker et al. (2018b) conducted a RCT with 226 participants recruited online with the assistance of various mental health organisations with a mean age of 34 years using one of three different apps or being in the waitlist control group, with 50 participants using the *MoodMission* app. Participants were instructed to use the CBT-based app for 30 days, with no further instructions than this. A statistically significant improvement in symptoms of depression was observed for *MoodMission*, but there was no such statistically significant improvement in symptoms of anxiety.

Paul and Fleming (2019) had a single intervention group that obtained pre to post data for 16 participants, all of whom were U.S. undergraduate students with a mean age of 19.8 years. The group used the *MindShift* app to provide anxiety-focused CBT therapy, with instructions to use the app for at least 15 minutes per day, 5 days per week, for 3 weeks (21 days). Statistically significant reductions were found in somatic and generalized anxiety, and depression, from pre- to post-treatment, but not for panic.

Lee and Jung (2018) had a control group of 86 participants and an intervention group of 77 participants, all Canadian undergraduate students with a mean age of 20.6 years. The intervention group used the *Destressify* app to provide mindfulness-based therapy, with instructions to use the app for 5 days per week for 4 weeks (28 days). Statistically significant reductions in anxiety were achieved from pre- to post-treatment, but there was no significant change in depression.

Challenges of Researching Mental Health Apps

There are several challenges to conducting traditional efficacy studies on mental health apps, and these are mainly due to the rapid pace of app development, the high cost of running large research trials, and the obsolescence of digital products – some mental health apps that go through a research process never become publicly available (Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017). The app industry is populated by young start-up companies with large investment funds to produce the "next big thing" in health-related apps (Medical Startups, 2020) and bring it to market as soon as possible. As such, efficacy research using the traditional gold standard randomised controlled trial (RCT) may be an impediment for mental health app research, given the long periods (sometimes years) to organise, run, and analyse a trial. During this time, other apps aimed at the same market may be listed for download, making the app going through this research process obsolete even before it has been publicly released (Clough & Casey, 2015a). Such costs on top of already large financial sums that have gone into the development of a product up to that point can be difficult for investors to accept. Thus, while RCTs are the "gold standard" for demonstrating efficacy, a different research approach may be required in the area of mental health apps (Clough & Casey, 2015a).

Single-Case Designs

Single-case research designs are a viable alternative to large group designs, such as RCTs, and have the capacity to evaluate both the efficacy and effectiveness of mental health apps (Clough & Casey, 2015b; Marshall et al., 2020e; Mehrotra & Tripathi, 2018). This is because single-case research designs can assess the causal relationship between an intervention and outcomes (i.e., its efficacy), while also having external validity to demonstrate effectiveness in heterogeneous samples (Lobo et al., 2017).

Single-case designs control for threats in internal validity by having continuous and repeated measurement of outcomes (dependent variables), random assignment, the potential

for multiple participants, the sequential introduction of varying levels of an intervention (the independent variable) across "phases" of a study, replication, and specific data analysis and statistics (Krasny-Pacini & Evans, 2018). With a baseline phase of "no treatment", a participant acts as their own control. In a design involving multiple participants, random assignment to the staggered introduction of an intervention addresses threats to internal validity from history, maturation and testing. In circumstances where three or more participants share similar presentations, receive identical treatment, and show strong outcomes, the results are considered to be a legitimate demonstration of efficacy (Barlow et al., 2009; Horner et al., 2005; Kazdin, 2017). By taking into account the features of individual participants, such designs also provide data on effectiveness (Buckley et al., 2014; Sheridan, 2014). See the Procedure section below for details of the *multiple baseline across-individuals* design of the present study.

There have been calls for practicing clinicians to be more involved in the process of researching mental health interventions, especially those that are well-suited to being incorporated into real-world therapeutic settings, such as smartphone apps (Clough & Casey, 2015a). The use of single-case designs could facilitate the recruitment of practicing clinicians to research the efficacy and effectiveness of mental health apps by focusing on a limited number of participants from the clinician's usual client load (Clough & Casey, 2015a; Marshall et al., 2020e). Marshall et al. (2020e) summarises a model of how clinicians may contribute to a centralised database of efficacy and effectiveness information on mental health apps by following the design of the present study. Such a database would offer an ever-increasing knowledge hub that complements app review websites such as *PsyberGuide* (https://psyberguide.org/); *Head To Health* (https://headtohealth.gov.au/); and the *NHS Apps Library* (https://www.nhs.uk/apps-library/).

The Impact and Consequences of COVID-19

Soon after the present study commenced in early 2020, the COVID-19 pandemic began to have widespread negative impacts around the globe. It became clear that mental health was one such negative impact across countries, including: Australia (Koh, 2020), China (Feng, 2020), India (Mukherjee, 2020), New Zealand (Mindfood, 2020), U.K. (Chowdhury, 2020), U.S. (Heilweil, 2020), and others. Due to the increased demand for mental health professionals, many people struggled to access in-person services and this led to increased demand for online / telehealth services (Lifeline, 2020; Liu et al., 2020; Marshall et al., 2020d; Medhora, 2020). This included a 50% increase in the number of Australian young people aged 18 – 25 who were accessing online mental health help (Reachout.com, 2020), and increased downloads of mental health apps (Basu, 2020; Heilweil, 2020; Marshall et al., 2020d; Statista, 2020).

Mental health apps may seem like a good option to manage mental illness during a pandemic. After all, over 5.2 billion people worldwide own a smartphone (Barboutov et al., 2017), including over 89% of the Australian population (Deloitte, 2018), and these figures are growing. With such potential for wide access to mental health apps, and with ongoing difficulties accessing in-person treatment for mental health issues (Liu et al., 2020), it was little wonder that people turned to digital options during the pandemic.

Mental health apps are also attractive for general practitioners. Apps have the potential to reduce the burden on primary health care at a time when primary health care is dealing en masse with the acute need to treat COVID-19 (Azarang et al., 2019). It is possible that many general practitioners believed that they could "prescribe" a mental health app for their patients (Byambasuren et al., 2018) due to the shortage of in-person mental health treatment options. However, it is likely many would not have been aware of some of the issues described earlier in this paper, including the lack of evidence for the efficacy and effectiveness of most publicly available mental health apps.

The timing of the pandemic in relation to the present study is both serendipitous and intriguing. The baseline period for all participants commenced on the 30th January 2020, and all participants were using their assigned app by the 28th February 2020. In Australia, where the study was completed, the Federal Government made several key announcements (including lockdown orders and making available additional government payments for people who became unemployed) between the 12th and 23rd March (Klapdor, 2020). This 12-day period saw an increase in stress across communities, including panic buying at grocery stores (Wright, 2020).

In relation to the present study, all participants had been using their app for at least two weeks before the pandemic received increasing attention in the mainstream media in Australia. The methodology was able to detect reliable spikes in Subjective Units of Distress (SUDS) between the crucial period of the 12th to 23rd March, and in the weeks and months afterwards. Therefore, the study has been able to provide data on how well these five mental health apps were able to assist people to manage symptoms of anxiety and/or depression during what has arguably been the most stressful period, on a countrywide scale, in a generation. More broadly, the results of this study provide quality evidence of the effectiveness of these apps to help manage anxiety and/or depression during a period of massive global upheaval.

Lessons from Pilot Work

A pilot study using a randomly chosen app from the five used in this study (*SuperBetter*) was conducted to test the feasibility of the proposed methodology. See Chapter 7 – Study 3 of this thesis for further details. The pilot study informed the current methodology: specifically, that the methodology can be used to answer the research questions; and, that assertive follow-up of participants who prematurely stop providing daily SUDS ratings should be used in an effort to reduce the rate of attrition.

This pilot study also revealed that while the positive psychology orientation of *SuperBetter* delivered benefits to all participants who finished the study, all but one of those who failed to complete had a mental health condition of more than 11 years duration. The pilot results also provided data for comparative comments about the apparent effectiveness of the intervention when delivered in this study and in the context of COVID-19.

The Present Study

The TREND Statement (Des Jarlais et al., 2004) was used to inform the conduct of this study – the checklist is located in Appendix B. The main objective of the study was to examine the effectiveness of five mental health apps, each drawing on one of a range of theoretical orientations for reducing symptoms of anxiety and/or depression. The apps selected were *SuperBetter*, *Smiling Mind*, *MoodMission*, *MindShift*, and *Destressify* (see Materials and Measures section below for further details).

The research protocol for the present study has been published (Marshall et al., 2020c) and the study is registered with the Australian and New Zealand Clinical Trials Registry (ANZCTR), which is a primary registry in the World Health Organization Registry Network, registration number ACTRN12619001302145p (http://www.ANZCTR.org.au/ACTRN12619001302145p.aspx). Readers are encouraged to refer to the published protocol for further information relating to the Methods used in the present study.

The present study sought to answer the following research questions:

1. Can a range of mental health apps, employing diverse theoretical orientations, reduce subjective distress and clinically significant symptoms of anxiety and/or depression, and improve functioning in a sample of heterogeneous participants?
2. Are there specific factors about the participants that impact on the results?
3. What are the participants' experiences of using the apps?

Method

The following Methods section is a summary. Refer to Chapter 6, the published research protocol (Marshall et al., 2020c), for the complete Methods section.

Participants

Inclusion criteria:

1. 18 years of age or older;
2. Ability to read English;
3. Have access to a smartphone or tablet device capable of connecting to the Internet and downloading the required app, and sending and receiving SMS text messages;
4. Agreeable to providing daily SUDS ratings via SMS text message, and to completing self-report measures at five different time points; and
5. Mild to moderate anxiety and/or depression, as previously diagnosed by a qualified health professional prior to applying to be in this research, and confirmed by the researchers (all of whom are clinical psychologists) after screening of the initial self-report measures to indicate the presence of symptoms of anxiety and / or depression using the cutoff points identified by previous research.

Exclusion criteria:

1. Severe anxiety and/or depression, as indicated by the initial outcome measures and in any responses to specific questions in the Demographics Questionnaire;
2. History of psychosis, or other complex mental health presentation as deemed by the researchers to be unsuitable for participation in this research (there was a question in the Demographics Questionnaire that asks participants for their complete mental health diagnoses); and
3. Current suicidal ideation, as indicated by a participant's responses on the initial outcome measures.

Removal criteria:

1. Not providing any SUDS rating for a two-week period;
2. Not providing a minimum of 20 SUDS ratings in the baseline and post-intervention phases, or a minimum of 40 SUDS ratings in the intervention phase;
3. Not completing outcome measures either pre-intervention, or post-intervention;
4. Clinically significant/unsafe decline in mental health as indicated by SUDS ratings or outcome measures, or in the judgment of researchers; and
5. Suicidal ideation that has developed during the participants' involvement in the study.

Materials and Measures

The Apps

The apps used in this study were *SuperBetter* (Roepke et al., 2015; Worthen-Chaudhari et al., 2017; https://www.superbetter.com); *Smiling Mind* (Flett et al., 2019; https://www.smilingmind.com.au/); *MoodMission* (Bakker, Kazantzis, Rickwood, & Rickard, 2018a, 2018b; Bakker & Rickard, 2018; http://moodmission.com/); *MindShift* (Paul & Fleming, 2019; https://www.anxietycanada.com/resources/mindshift-cbt/); and *Destressify* (Lee & Jung, 2018; https://www.destressify.com/). These apps were purposively selected on the basis of using an evidence-based treatment approach; evidence of efficacy in reducing symptoms of anxiety and/or depression; and, had an accompanying website with further information, including privacy statements. In terms of theoretical orientations, *SuperBetter* uses a positive psychology framework and incorporates ideas from neuroscience in the area of neuroplasticity; *Smiling Mind* uses a structured mindfulness-based framework; *MoodMission* uses a CBT framework that emphasises a behavioural approach, but still contains cognitive elements; *MindShift* uses a more cognitively-focused CBT framework; and *Destressify* uses a less structured mindfulness-based framework compared to *Smiling Mind*.

Demographic and Biographic Features

A questionnaire was developed by the researchers to elicit demographic and biographic information.

Mental Health and Wellbeing

The three-phase model of psychotherapy outcomes (Howard et al., 1993) was used as the framework for examining participant outcomes relating to subjective distress, symptomatology, and life functioning as follows:

1. Subjective distress: SUDS ratings – participants rated their level of distress in response to the question: "How do you feel today?", with 0 indicating *no distress* and 10 indicating *worst possible distress*, and a score of '3' or more indicating a mild but noticeable level of upset (Wolpe & Lazarus, 1966). SUDS ratings have been shown to be a valid measure of emotional discomfort when compared with other measures of distress ($r = 0.351$, $p < .05$; Tanner, 2012).
2. Symptoms: The Depression Anxiety Stress Scale – 21 short-form version (DASS-21; Henry & Crawford, 2005). Participants rated their experience of symptoms of depression, anxiety and stress over the previous week on a 4-point scale ranging from 0 (*did not apply to me at all*) to 3 (*applied to me very much, or most of the time*). The total scores for the subscales are multiplied by two in order to interpret the severity ratings according to the longer 42-item scale (Antony et al., 1998; Lovibond & Lovibond, 1995) In this study, only the depression and anxiety subscales were used. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).
3. Life functioning: The Outcome Questionnaire – 45 2nd Edition version (OQ-45.2; Boswell et al., 2013) is a 45-item self-report scale that measures distress,

interpersonal relationships and social role functioning in adults 18 years and older (Beckstead et al., 2003). An index for overall life functioning is calculated (Lambert & Finch, 1999). Participants rate their feelings over the previous week on a 5-point scale ranging from 0 (*never*) to 4 (*always*). Possible scores range from 0 to 180 with a total score of 63 or more being indicative of clinically significant symptoms (Lambert & Finch, 1999). Lambert et al. (2004) have suggested the following interpretive labels: > 105 is *High*, 83 – 104 is *Moderately High*, 63 – 82 is *Moderate*, and < 63 is *Normal*.

App Appraisal

The Mobile Application Rating Scale – User Version (uMARS; Stoyanov, Hides, Kavanagh, & Wilson, 2016) is a 20-item questionnaire recording an individual's rating on the quality of a mobile app. It contains multiple choice and Likert-type responses, and also contains a free text field allowing users to provide a qualitative description of any aspect of the app, or their experience of using the app.

Data Analysis Plan

Data from this project are publicly available through the University of New England's *Research UNE* website, https://rune.une.edu.au/web/.[2]

Descriptive Statistics and Qualitative Accounts

Descriptive statistics were used to describe individual participant features and augment the findings from the other analyses.

Visual Inspection

Visual inspection was used to assess the impact of the intervention on subjective distress (SUDS). Plotted data allows for a personal judgement about the effect of an

[2] At time of writing, we are waiting to receive a DOI for the dataset.

intervention (Kazdin, 2017), and in this study, visual inspection was possible using up to 122 data points of SUDS ratings (this was the highest number of individual SUDS ratings, by Participant B4 – see Table 10.12).

Time Series Analysis

A time-series analysis was used to assess the statistical significance of changes in each participant's plotted data across each phase of the study. Scores at the commencement of the intervention were used as the predictor in a regression model.

Clinical Significance and Statistical Reliability

Clinically significant symptoms of depression and anxiety, and changes in level of severity, were identified according to the published normative data for the DASS-42. Life functioning was assessed for clinical significance and change using the clinical significance index (CSI) (Jacobson et al., 1999) and the reliable change index (RCI) (Jacobson & Truax, 1991). Normative data for a scale is used to calculate the CSI, which is the cut-off point between the scores obtained by functional (non-clinical) and dysfunctional (clinical) populations (Evans et al., 1998; Jacobson & Truax, 1991). In this study, the CSI was used to note each participant's clinical status pre- and post-intervention (Evans et al., 1998; Jacobson & Truax, 1991). The reliable change index (RCI) was used to assess and classify the statistical significance of any change in participants' score from pre- to post-intervention: *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

Procedure

The University of New England Human Research Ethics Committee approved the project on the 1st November 2019, Approval Number HE19-186. Between 1st November 2019

and 30th January 2020, participants were recruited throughout Australia by directly approaching non-government mental health services, mental health associations (both consumer and professional), and support groups and other organisations in the mental health sector. Thirty-nine participants responded to calls for expressions of interest in the study. After informed consent was obtained, participants commenced the baseline phase simultaneously, and were randomly selected to begin the intervention phase in staggered order. Randomization was achieved using the online random number generator, *Research Randomizer* (Urbaniak & Plous, 2019; https://www.randomizer.org). See Figure 10.1 for a flowchart of the study's phases and participant involvement.

Participants received instructions on usage and background information relating to their particular app. Participants were asked to use their app for at least 10 minutes per day for five days per week for 10 weeks. The 10-week intervention period creates equivalence with one 50-min session per week for 10 weeks, which is the annual maximum number of psychology sessions rebated under Australia's Medicare system. It is acknowledged that these usage instructions do not necessarily equate to the ideal usage instructions across all mental health apps. However, in the interests of streamlining the process for all participants, it was believed that this instruction of at least 10 minutes per day for five days per week would encourage participants to engage with their app in a meaningful way.

By using a single-case design, participants were able to be observed closely and in "real-time" allowing for highly responsive treatment. This is an important consideration in mental health research where participants may be experiencing suicidal ideation. Individual wellbeing was monitored by participants providing daily SUDS ratings by sending a SMS text message from their smartphone to a centrally monitored hub. While it is acknowledged that a rating out of 10 is itself limited in its ability to convey the complexities of an individual's mental health, it can allow a mental health researcher / clinician to evaluate the

relatively immediate influence of a treatment (Machalicek & Horner, 2018), adjust the intervention in response to changes in ratings, or halt the intervention and rapidly arrange crisis support if necessary (Bentley et al., 2019). However, halting the intervention or crisis support was not required for any participant.

Ten participants were lost to the study by the time the intervention phase had finished. A total of 29 participants completed the post-intervention phase, producing an attrition rate of 25.60%. This was substantially improved from the pilot study's attrition rate of 60% and is attributed to assertive follow-up by the researcher when participants did not provide SUDS ratings for three consecutive days during the baseline or intervention phases. As noted above, this strategy was introduced following the outcomes of the pilot study, and was the single difference in methodology between the present study and the pilot study.

For a more detailed breakdown of the processes of each phase, refer to Chapter 6, the published research protocol (Marshall et al., 2020c). The phases were identified as: Phase 1 (Pre-Baseline), Phase 2 (Baseline), Phase 3 (Intervention), and Phase 4 (Post-Intervention).

Results

Participant Characteristics and Descriptive Statistics

A total of 39 participants commenced the study. Of the 29 who finished the post-intervention phase, 20 were female. Seven participants (B7, C3, D3, D5, E1, E4, and E6) were assertively followed up during the study when they did not provide SUDS ratings for three consecutive days, and then re-joined the study. The age-range of completers was 18 to 57 (M = 34.0, SD = 12.2); 16 (55.17%) had their diagnosis for 5 years or less, 15 (51.72%) were receiving concurrent counselling, and 14 (48.28%) were taking psychotropic medication. Eleven (37.93%) had an anxiety disorder only, eight (27.59%) had depression only, and 10 (34.48%) had co-morbid anxiety and depression. In terms of motivation to

comply with the intervention, 10 of the completers (34.48%) agreed that their psychological health would improve, 13 (44.83%) were neutral, and six (20.69%) thought their psychological health would not improve.

Of the 10 non-completers, seven dropped out in the baseline phase and three in the intervention phase. All 10 were followed up once after not providing SUDS for three days and encouraged to continue in the study. The age range of the non-completers was 29 to 68 (M = 44.5, SD = 12.3) and was significantly different to the completers ($t(39) = -2.34$; p = .03) who were younger. For the 10 non-completers, all had depression with two (20%) of these having comorbid anxiety. Six (60%) agreed that the intervention would improve their psychological health, three (30%) were neutral, and one (10%) thought their psychological health would not improve. For more information on participant characteristics, including those who dropped out of the study, see Tables 10.1 to 10.10 below.

Table 10.1

Demographic and Biographic Information for Participants A1 – A4 Using SuperBetter

Variable	Participant A1	Participant A2	Participant A3	Participant A4
Age (in years)	26	36	35	19
Sex	Female	Female	Female	Female
Highest education	University	Secondary – Years 7-10	University	Secondary – Years 11-12
Mental health status	Depression, GAD	Depression, panic attacks, PTSD, social anxiety, bipolar disorder	Panic attacks, PTSD	Depression, panic attacks, GAD
Years with mental illness	1-5	1-5	1-5	1-5
Currently receiving counselling?	Yes – Psychologist	Yes – Psychiatrist, psychologist	No	Yes - Psychologist
Currently on medication?	Yes – Antidepressant, antipsychotic, mood stabiliser	Yes - Benzodiazepine	No	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Strongly agree	Strongly agree	Strongly agree	Strongly agree
Ability with technology generally	Excellent	Good	Excellent	Excellent

Table 10.2

Demographic and Biographic Information for Participants A5 – A8 Using SuperBetter

Variable	Participant A5	Participant A6[a]	Participant A7[a]	Participant A8[a]
Age (in years)	30	51	46	40
Sex	Female	Female	Female	Male
Highest education	University	Secondary – Years 7-10	University	University
Mental health status	Depression	Depression	Depression	Depression
Years with mental illness	1-5	> 11	> 11	> 11
Currently receiving counselling?	No	Yes – Psychologist	No	No
Currently on medication?	Yes – Antidepressant	No	No	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Somewhat agree	Strongly agree	Neutral	Somewhat agree
Ability with technology generally	Average	Good	Average	Excellent

[a] Participants that dropped out of the study.

Table 10.3

Demographic and Biographic Information for Participants B1 – B4 Using Smiling Mind

Variable	Participant B1	Participant B2	Participant B3	Participant B4
Age (in years)	20	23	24	55
Sex	Female	Male	Female	Female
Highest education	Secondary – Years 11-12	University	Secondary – Years 11-12	University
Mental health status	Depression, panic attacks, GAD	Depression, social anxiety, autism	Panic attacks, OCD, social anxiety	Depression
Years with mental illness	1-5	1-5	1-5	> 11
Currently receiving counselling?	Yes – Psychologist	Yes – Psychiatrist, psychologist	Yes - Counsellor	No
Currently on medication?	No	Yes – Antidepressant	No	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Somewhat disagree	Neutral	Neutral	Somewhat agree
Ability with technology generally	Good	Excellent	Excellent	Poor

Table 10.4

Demographic and Biographic Information for Participants B5 – B7 Using Smiling Mind

Variable	Participant B5	Participant B6	Participant B7
Age (in years)	18	42	30
Sex	Female	Female	Male
Highest education	Secondary – Years 11-12	Secondary – Years 7-10	University
Mental health status	Panic attacks, social anxiety, GAD	Depression	Depression, GAD
Years with mental illness	1-5	6-10	1-5
Currently receiving counselling?	Yes - Psychologist	No	Psychologist
Currently on medication?	No	No	Yes - ?
I am motivated to do what the mobile app suggests?	Neutral	Somewhat agree	Somewhat disagree
Ability with technology generally	Good	Average	Good

Table 10.5

Demographic and Biographic Information for Participants C1 – C4 Using MoodMission

Variable	Participant C1	Participant C2	Participant C3	Participant C4
Age (in years)	27	47	42	20
Sex	Female	Male	Female	Female
Highest education	Secondary – Years 11-12	Secondary – Years 11-12	University	Secondary – Years 11-12
Mental health status	Depression, panic attacks, OCD	Depression	Panic attacks, GAD	Social anxiety
Years with mental illness	6-10	> 11	> 11	1-5
Currently receiving counselling?	Yes – Psychologist	No	No	Yes - Counsellor
Currently on medication?	No	Yes - Antidepressant	No	No
I am motivated to do what the mobile app suggests?	Neutral	Neutral	Neutral	Somewhat agree
Ability with technology generally	Excellent	Good	Poor	Average

Table 10.6

Demographic and Biographic Information for Participants C5 – C8 Using MoodMission

Variable	Participant C5	Participant C6	Participant C7[a]	Participant C8[a]
Age (in years)	31	44	30	35
Sex	Female	Female	Male	Male
Highest education	University	University	Secondary – Years 11-12	Secondary – Years 11-12
Mental health status	Depression, GAD	Panic attacks, PTSD, GAD	Depression	Depression, PTSD
Years with mental illness	1-5	> 11	6-10	> 11
Currently receiving counselling?	Yes - Psychologist	No	Yes – Psychologist	No
Currently on medication?	No	No	No	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Somewhat disagree	Neutral	Strongly agree	Neutral
Ability with technology generally	Average	Poor	Excellent	Good

[a] Participants that dropped out of the study.

Table 10.7

Demographic and Biographic Information for Participants D1 – D4 Using MindShift

Variable	Participant D1	Participant D2	Participant D3	Participant D4
Age (in years)	49	57	31	22
Sex	Male	Female	Female	Female
Highest education	Secondary – Years 11-12	University	Secondary – Years 11-12	Secondary – Years 11-12
Mental health status	Depression	GAD	Depression, panic attacks, Borderline personality disorder	GAD
Years with mental illness	> 11	> 11	1-5	1-5
Currently receiving counselling?	No	No	Yes - Psychologist	Yes - Psychologist
Currently on medication?	Yes - Antidepressant	Yes - Antidepressant	No	No
I am motivated to do what the mobile app suggests?	Somewhat agree	Neutral	Somewhat disagree	Somewhat agree
Ability with technology generally	Average	Poor	Excellent	Good

Table 10.8

Demographic and Biographic Information for Participants D5 – D8 Using MindShift

Variable	Participant D5	Participant D6[a]	Participant D7[a]	Participant D8[a]
Age (in years)	52	52	39	55
Sex	Male	Female	Female	Female
Highest education	Secondary – Years 7-10	University	University	University
Mental health status	Depression	Depression	Depression	Depression
Years with mental illness	> 11	> 11	6-10	> 11
Currently receiving counselling?	No	Yes - Counsellor	No	Yes - Psychologist
Currently on medication?	Yes - Antidepressant	No	Yes - ?	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Somewhat disagree	Strongly agree	Somewhat agree	Somewhat agree
Ability with technology generally	Good	Good	Good	Excellent

[a] Participants that dropped out of the study.

Table 10.9

Demographic and Biographic Information for Participants E1 – E4 Using Destressify

Variable	Participant E1	Participant E2	Participant E3	Participant E4
Age (in years)	46	39	35	18
Sex	Male	Male	Male	Female
Highest education	University	University	University	Secondary – Years 11-12
Mental health status	Panic attacks, social anxiety, autism	Depression, bipolar disorder	Depression, PTSD	Panic attacks, GAD
Years with mental illness	> 11	> 11	> 11	1-5
Currently receiving counselling?	No	No	No	Yes - Psychologist
Currently on medication?	No	Yes – Antidepressant, antipsychotic	Yes – Antidepressant, benzodiazepine	No
I am motivated to do what the mobile app suggests?	Neutral	Somewhat disagree	Strongly agree	Strongly agree
Ability with technology generally	Excellent	Average	Average	Average

Table 10.10

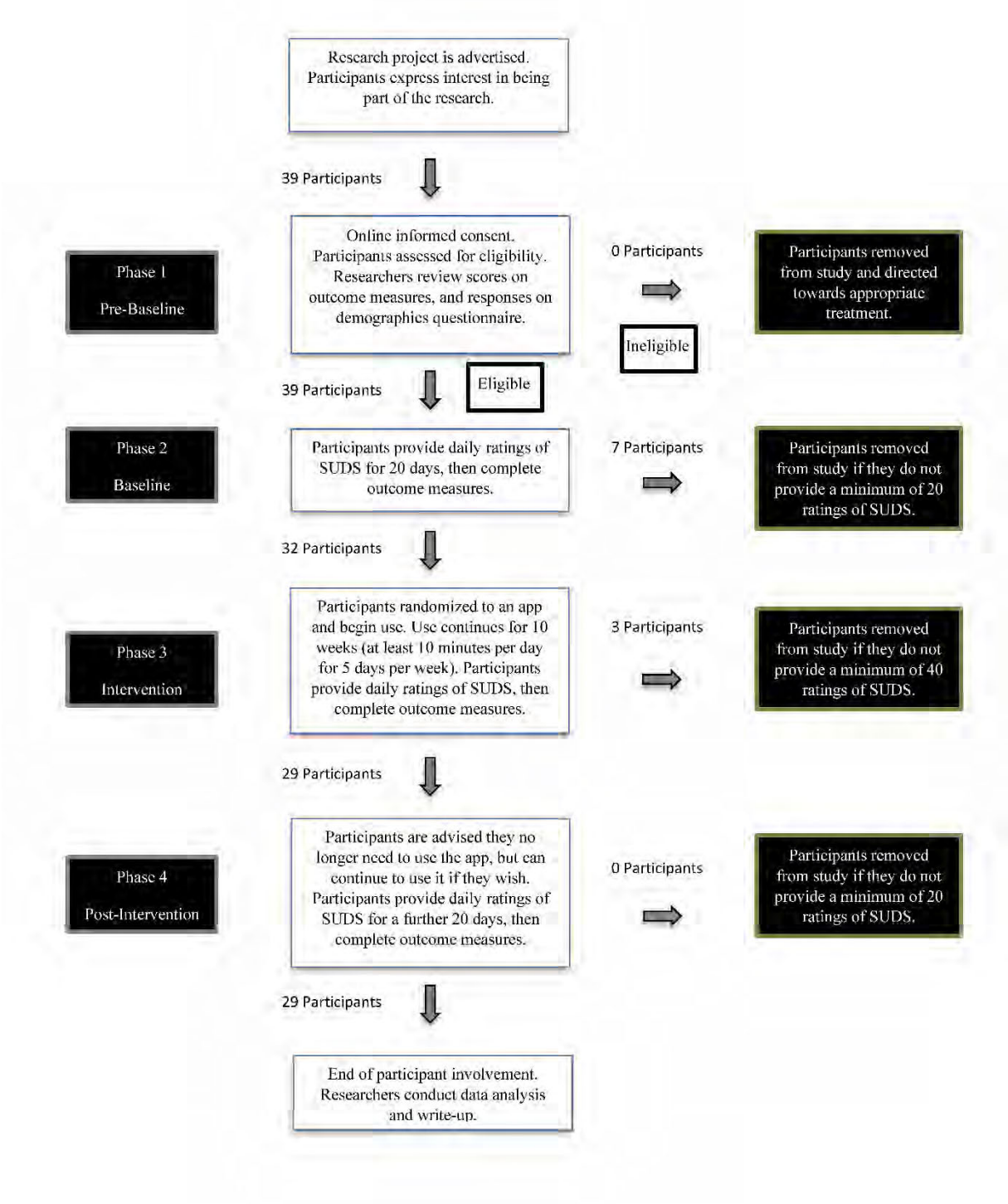
Demographic and Biographic Information for Participants E5 – E8 Using Destressify

Variable	Participant E5	Participant E6	Participant E7[a]	Participant E8[a]
Age (in years)	49	20	29	68
Sex	Female	Male	Female	Female
Highest education	University	Secondary – Years 11-12	University	Secondary – Years 7-10
Mental health status	Depression	Depression, panic attacks, GAD	Depression, PTSD	Depression
Years with mental illness	> 11	1-5	> 11	> 11
Currently receiving counselling?	No	Yes - Psychologist	Yes – Psychologist	No
Currently on medication?	Yes - Antidepressant	No	Yes - Antidepressant	Yes - Antidepressant
I am motivated to do what the mobile app suggests?	Strongly agree	Somewhat agree	Strongly agree	Neutral
Ability with technology generally	Average	Excellent	Excellent	Good

[a] Participants that dropped out of the study.

Figure 10.1

Flowchart of Participant Involvement and Study Phases



Effectiveness of the Apps in Reducing Subjective Distress

Visual Inspection

Visual inspection of the plotted SUDS data and the time series analyses for the 29 participants who completed the study is reported below by app.

SuperBetter. Five participants used the *SuperBetter* app. Visual inspection revealed that four participants were experiencing noticeable feelings of distress at baseline. By post-intervention, three participants had achieved a reduction in subjective distress to a non-noticeable level (i.e., a rating of < 3); but two had deteriorated. Table 10.11 shows the mean SUDS ratings per participant by phase; Figures 10.2 and 10.3 display the continuous data.

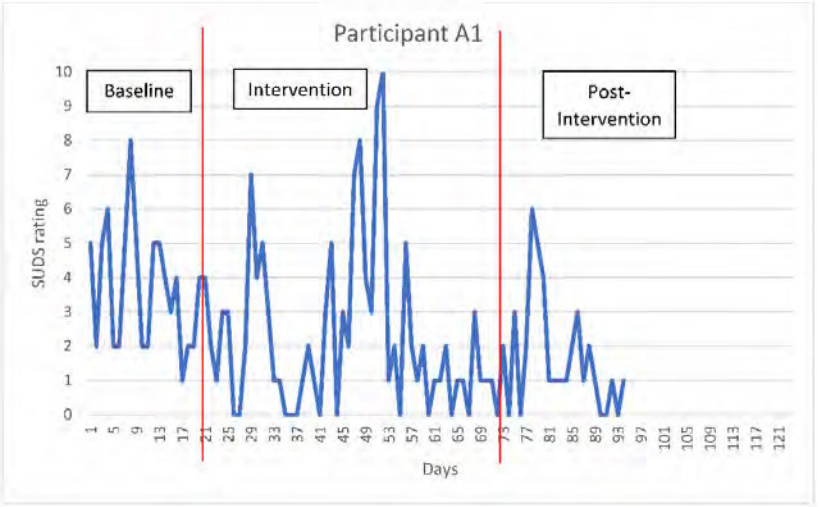
Table 10.11

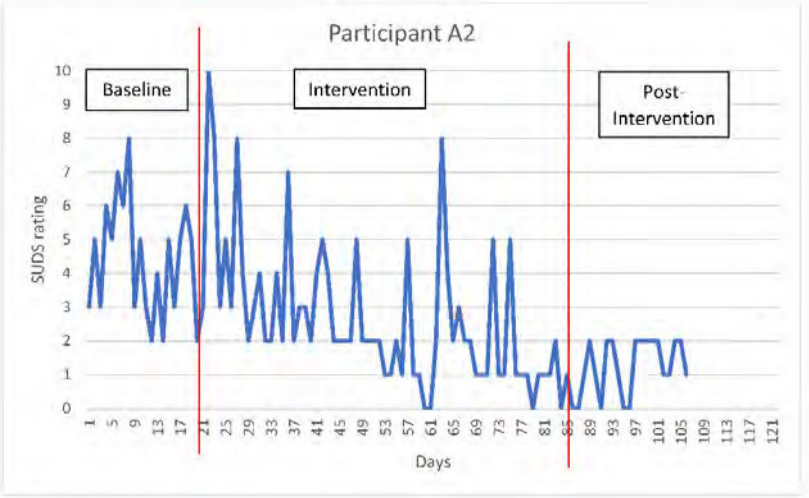
SUDS Data Summary for Participants Using SuperBetter

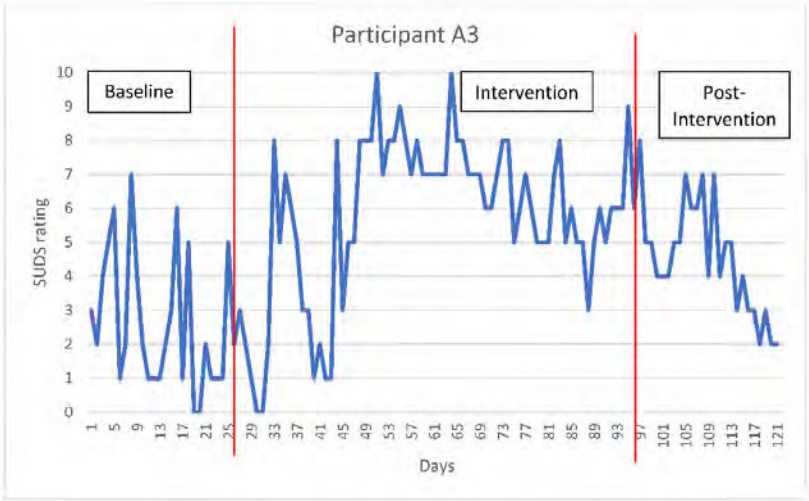
Participant		Baseline SUDS	Intervention SUDS	Post-intervention SUDS
A1	*M*	3.7	2.3	1.7
	SD	1.8	2.5	1.6
	Frequency (*n*)	21	51	22
A2	*M*	4.3	2.7	1.2
	SD	1.7	2.1	0.8
	Frequency (*n*)	22	64	21
A3	*M*	2.6	5.8	4.5
	SD	2.0	2.4	1.7
	Frequency (*n*)	26	70	25
A4	*M*	4.3	6.0	6.1
	SD	1.5	1.5	1.3
	Frequency (*n*)	21	64	28
A5	*M*	5.0	3.5	2.9
	SD	1.2	1.6	1.2
	Frequency (*n*)	23	67	26

Figure 10.2

SUDS Ratings for Participants A1 – A4 Using SuperBetter







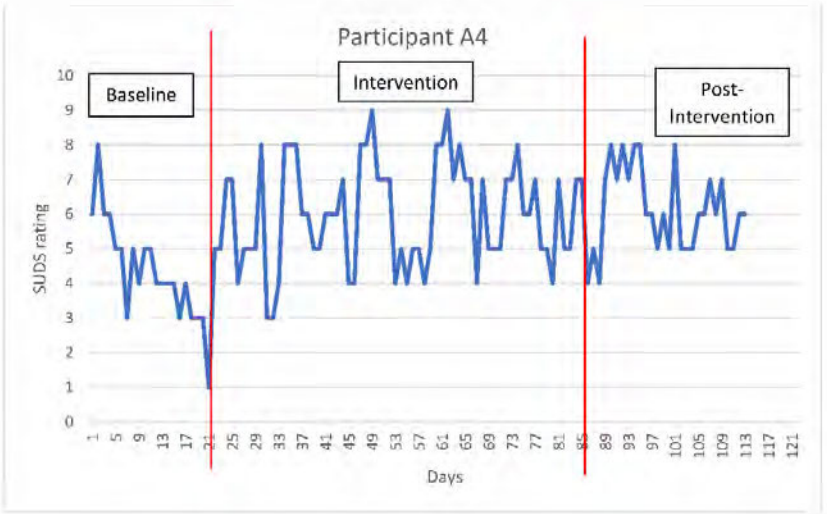
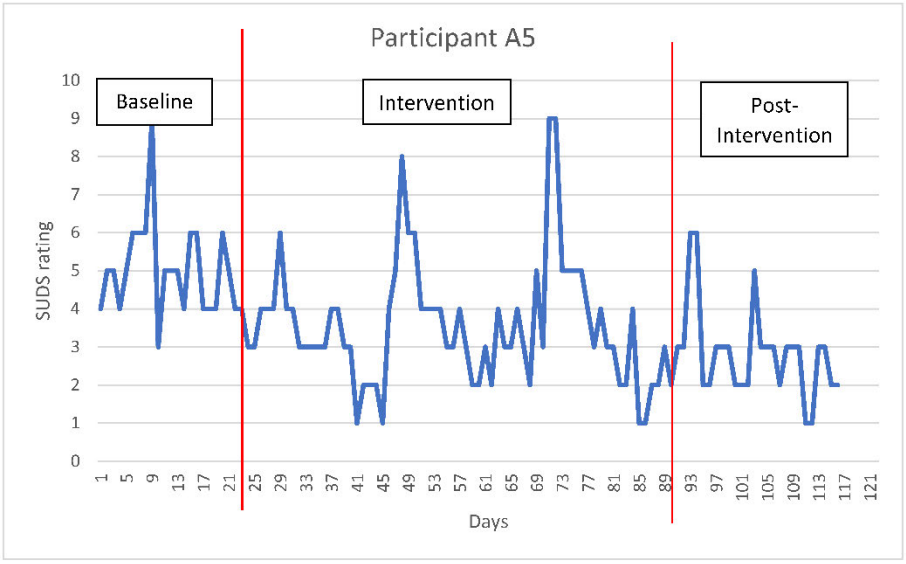


Figure 10.3

SUDS Ratings for Participant A5 Using SuperBetter



Smiling Mind. Seven participants used the *Smiling Mind* app. Visual inspection revealed that all were experiencing noticeable feelings of distress at baseline. By post-intervention, six participants had achieved a reduction in subjective distress; five to a non-noticeable level (i.e., a rating of < 3). Table 10.12 shows the mean SUDS ratings per participant by phase; Figures 10.4 and 10.5 display the continuous data.

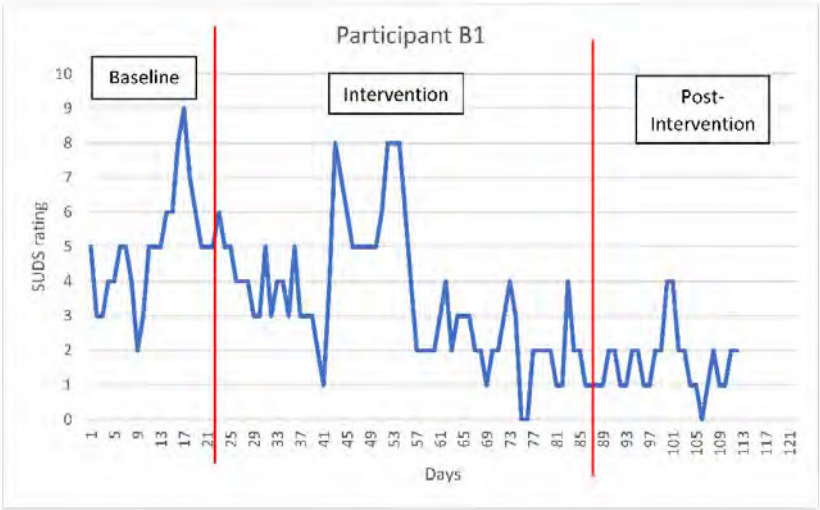
Table 10.12

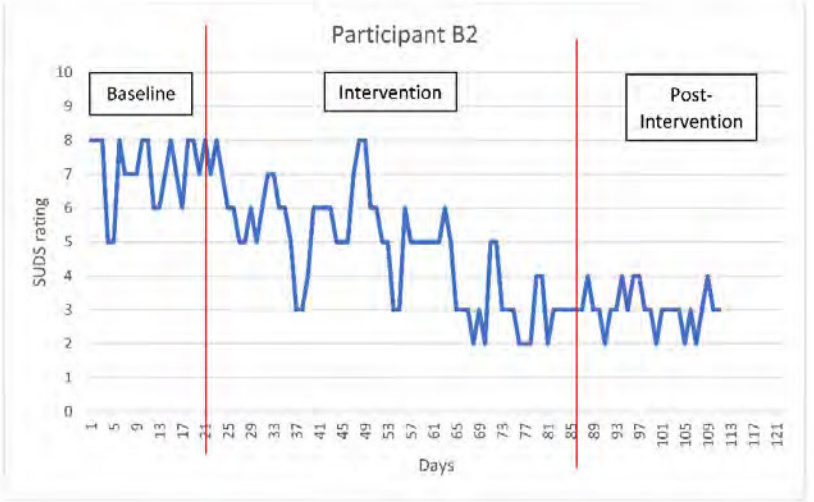
SUDS Data Summary for Participants Using Smiling Mind

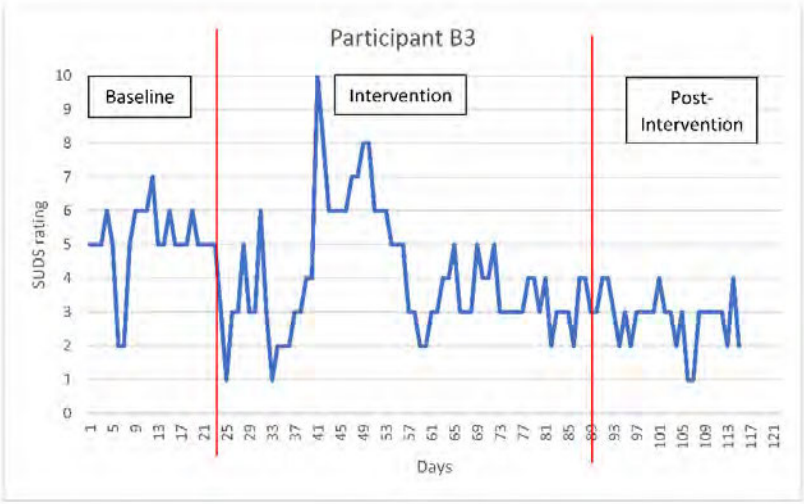
Participant		Baseline SUDS	Intervention SUDS	Post-intervention SUDS
B1	*M*	5.0	3.5	1.6
	SD	1.6	2.0	0.9
	Frequency (*n*)	22	65	25
B2	*M*	7.1	4.7	3.0
	SD	1.0	1.7	0.6
	Frequency (*n*)	21	65	25
B3	*M*	5.1	4.0	2.8
	SD	1.1	1.8	0.8
	Frequency (*n*)	23	66	26
B4	*M*	6.8	5.9	4.9
	SD	1.2	2.0	1.5
	Frequency (*n*)	27	68	27
B5	*M*	7.2	3.0	2.0
	SD	1.2	2.0	1.2
	Frequency (*n*)	23	72	26
B6	*M*	3.7	2.6	2.2
	SD	1.2	1.0	1.0
	Frequency (*n*)	23	66	25
B7	*M*	5.5	4.0	2.1
	SD	0.7	2.1	0.9
	Frequency (*n*)	22	70	26

Figure 10.4

SUDS Ratings for Participants B1 – B4 Using Smiling Mind







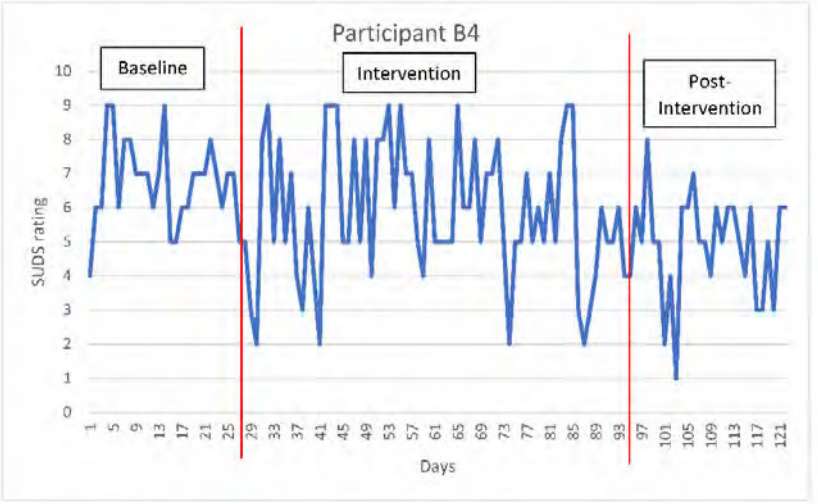
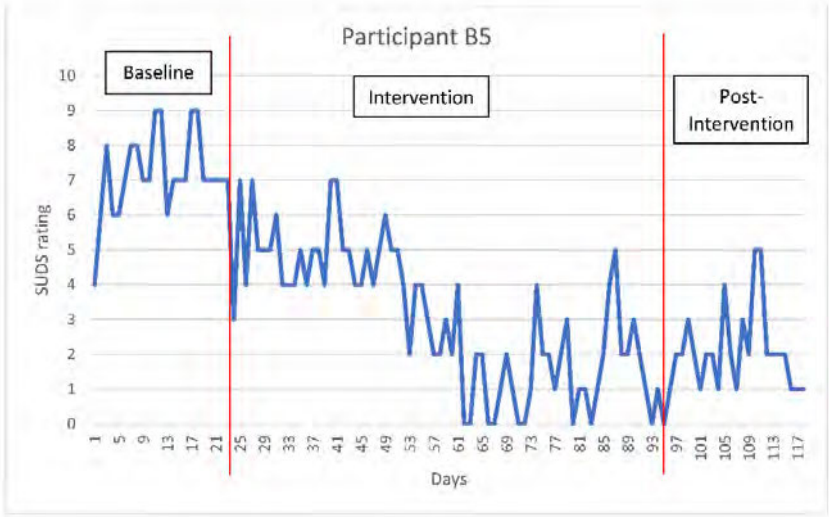
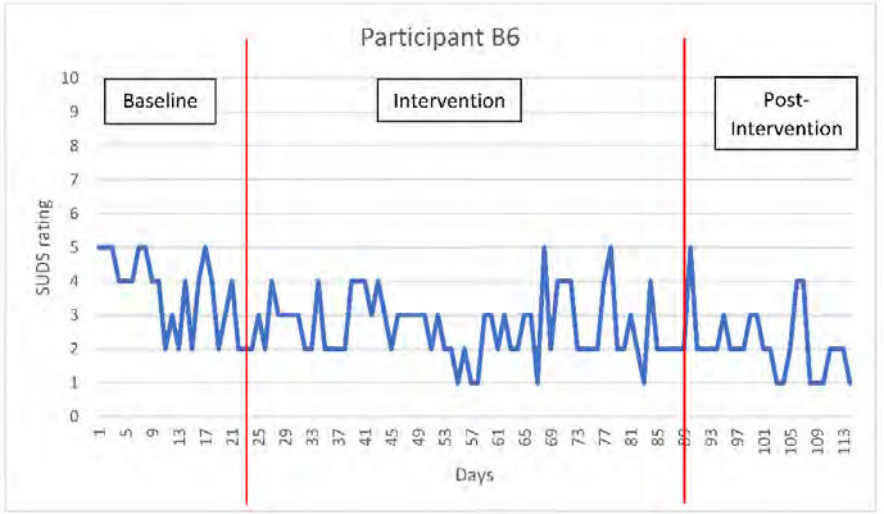
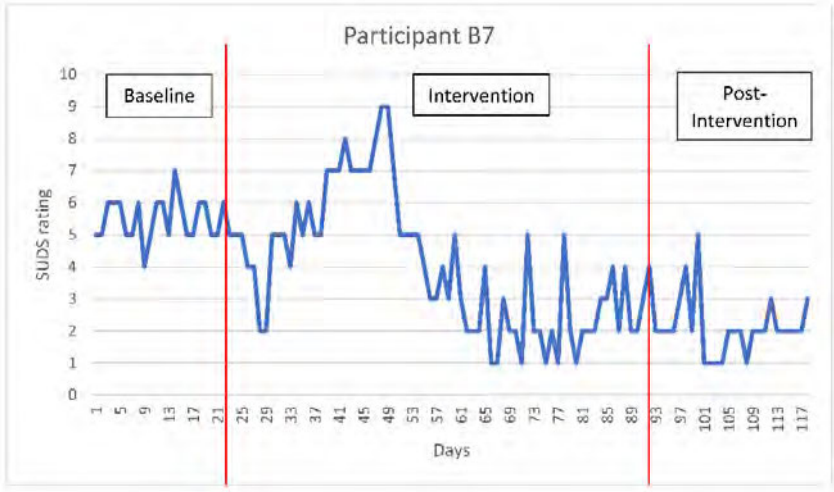


Figure 10.5

SUDS Ratings for Participants B5 – B7 Using Smiling Mind







MoodMission. Six participants used the *MoodMission* app. Visual inspection revealed that all were experiencing noticeable feelings of distress at baseline. By post-intervention, all but one participant (C3) had achieved a reduction in subjective distress; four to a non-noticeable level (i.e., a rating of < 3). Table 10.13 shows the mean SUDS ratings per participant by phase; Figures 10.6 and 10.7 display the continuous data.

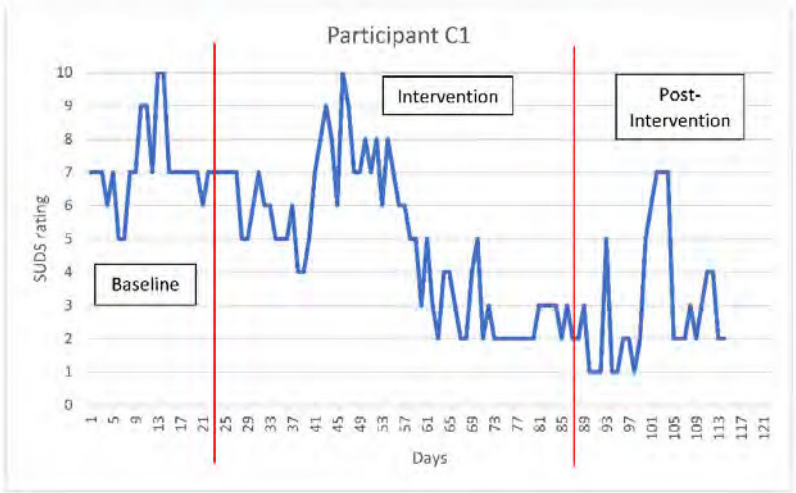
Table 10.13

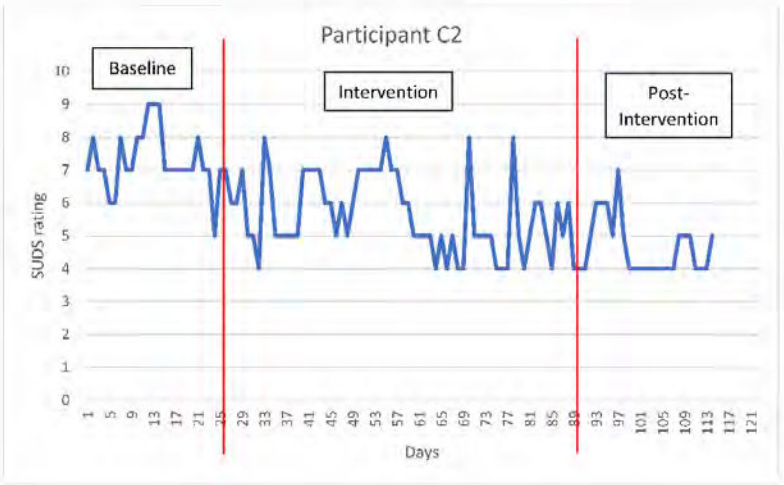
SUDS Data Summary for Participants Using MoodMission

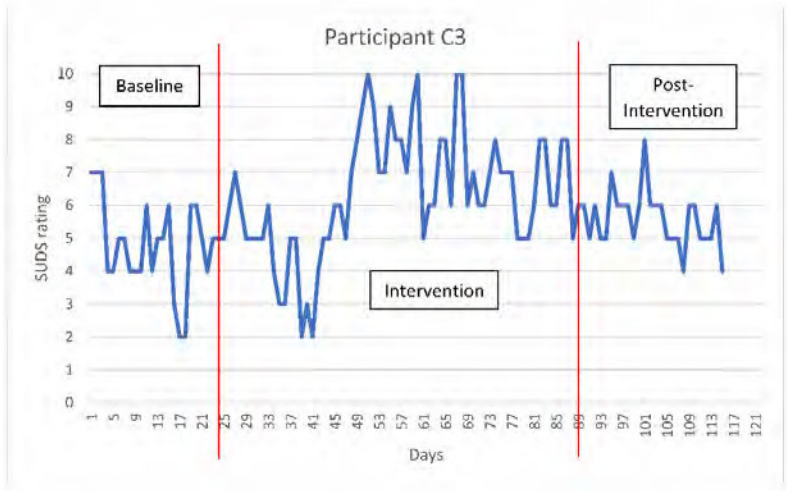
Participant		Baseline SUDS	Intervention SUDS	Post-intervention SUDS
C1	*M*	4.5	4.8	1.8
	SD	1.8	2.1	0.8
	Frequency (*n*)	22	62	27
C2	*M*	7.3	5.6	4.6
	SD	0.9	1.2	0.9
	Frequency (*n*)	26	63	25
C3	*M*	4.8	6.3	5.6
	SD	1.4	1.9	0.9
	Frequency (*n*)	24	65	26
C4	*M*	4.1	3.2	2.6
	SD	1.5	1.6	0.9
	Frequency (*n*)	25	67	27
C5	*M*	5.9	4.9	2.7
	SD	1.5	1.4	0.8
	Frequency (*n*)	26	68	26
C6	*M*	3.6	3.2	2.5
	SD	1.0	1.2	1.4
	Frequency (*n*)	25	67	26

Figure 10.6

SUDS Ratings for Participants C1 – C4 Using MoodMission







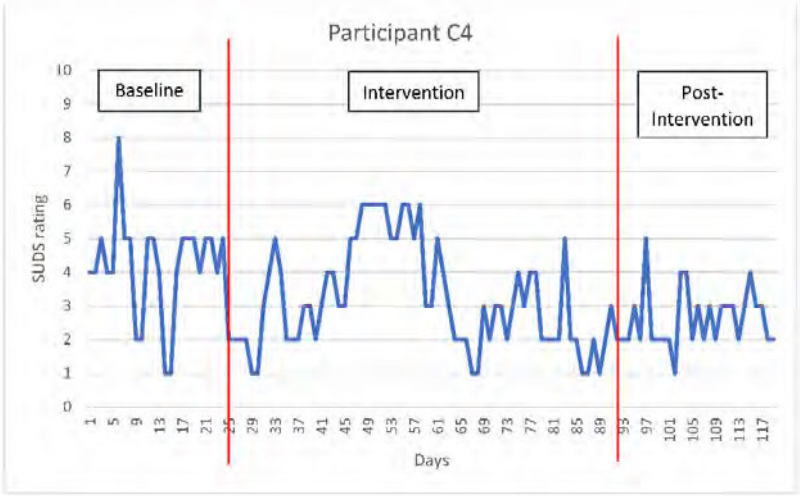
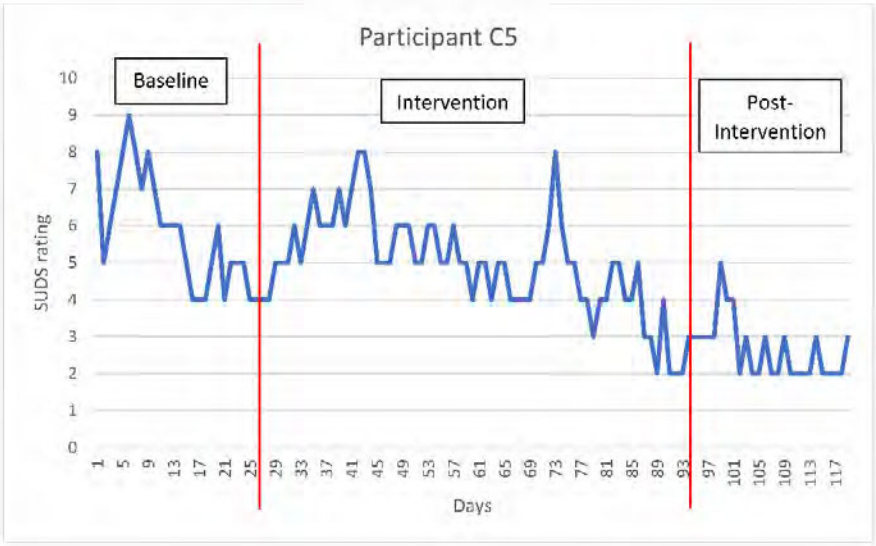
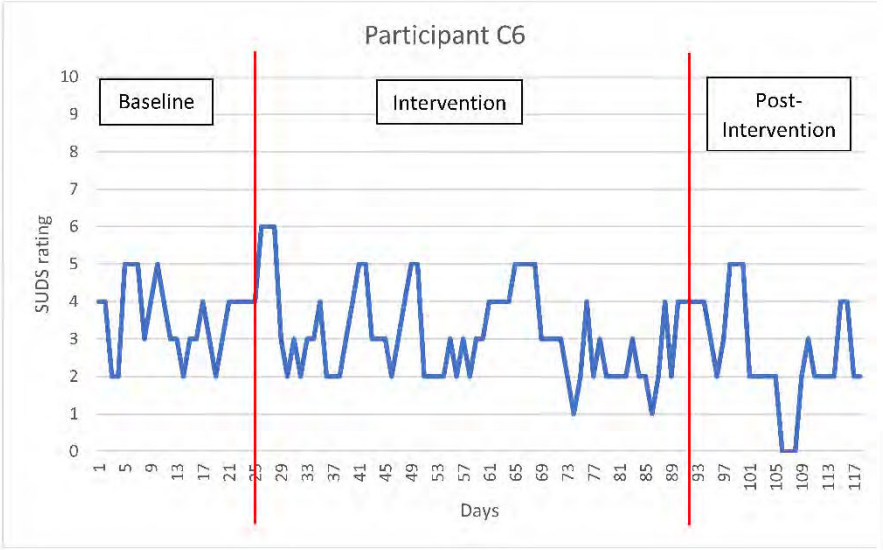


Figure 10.7

SUDS Ratings for Participants C5 – C6 Using MoodMission





MindShift. Five participants used the *MindShift* app. Visual inspection revealed that all were experiencing noticeable feelings of distress at baseline. By post-intervention, three participants had achieved a reduction in subjective distress; two to a non-noticeable level (i.e., a rating of < 3). Table 10.14 shows the mean SUDS ratings per participant by phase; Figures 10.8 and 10.9 display the continuous data.

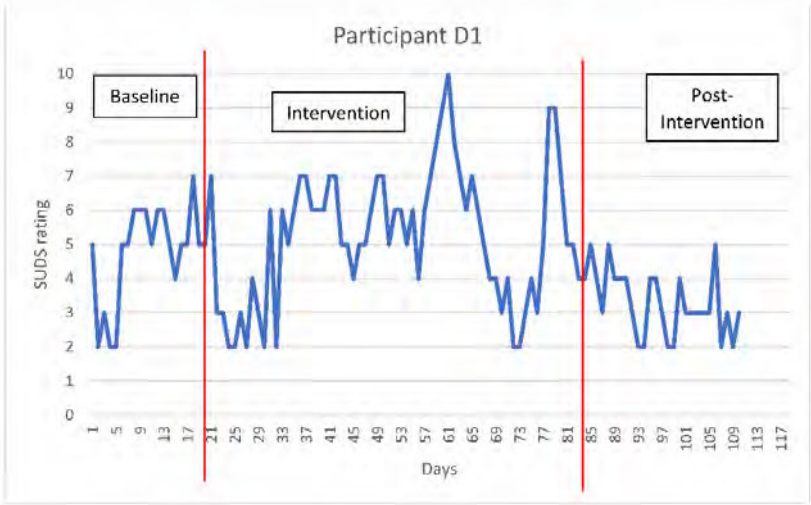
Table 10.14

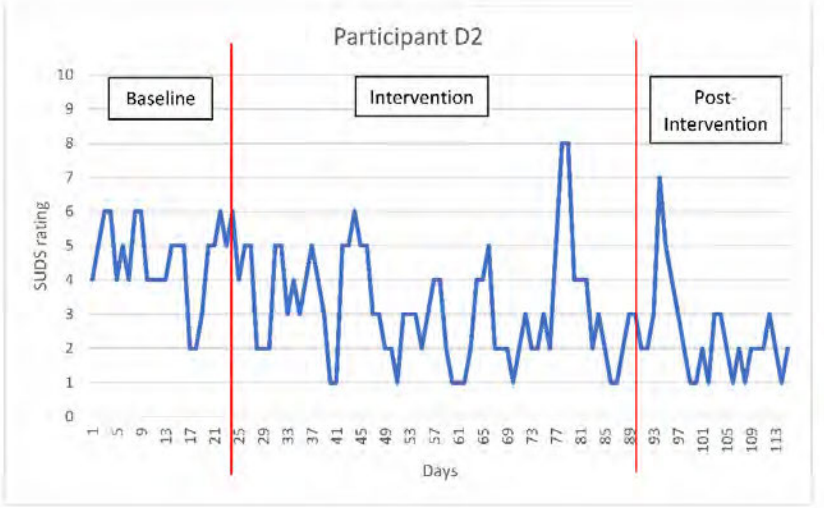
SUDS Data Summary for Participants Using MindShift

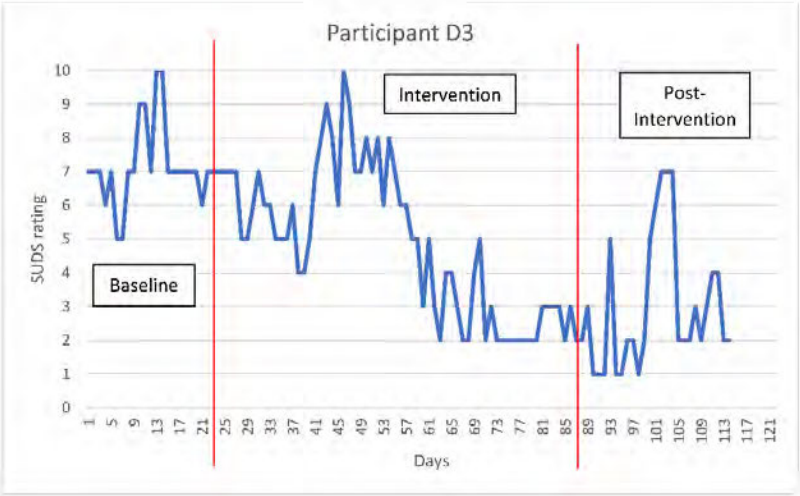
Participant		Baseline SUDS	Intervention SUDS	Post-intervention SUDS
D1	*M*	4.8	5.2	3.3
	SD	1.5	1.8	1.0
	Frequency (*n*)	20	64	26
D2	*M*	4.6	3.2	2.4
	SD	1.2	1.5	1.4
	Frequency (*n*)	24	66	25
D3	*M*	7.2	4.6	3.0
	SD	1.3	1.5	2.0
	Frequency (*n*)	23	64	27
D4	*M*	3.3	3.6	2.6
	SD	1.2	1.0	0.8
	Frequency (*n*)	25	67	27
D5	*M*	5.7	6.6	5.2
	SD	1.4	1.4	1.2
	Frequency (*n*)	20	63	27

Figure 10.8

SUDS Ratings for Participants D1 – D4 Using MindShift







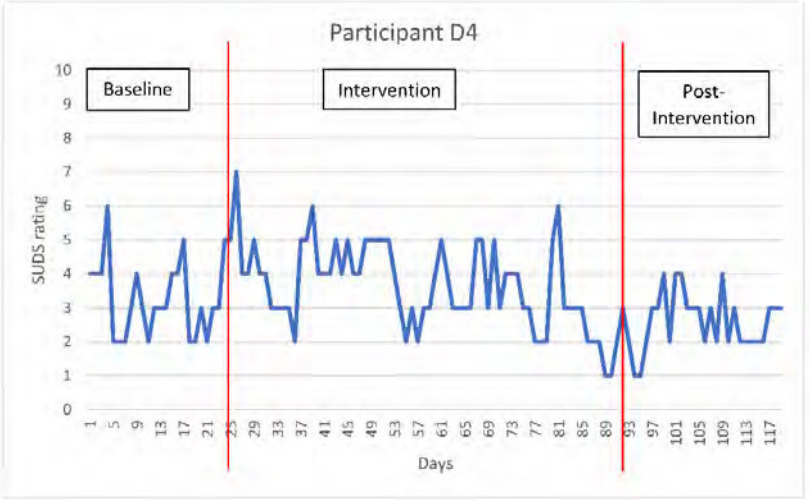
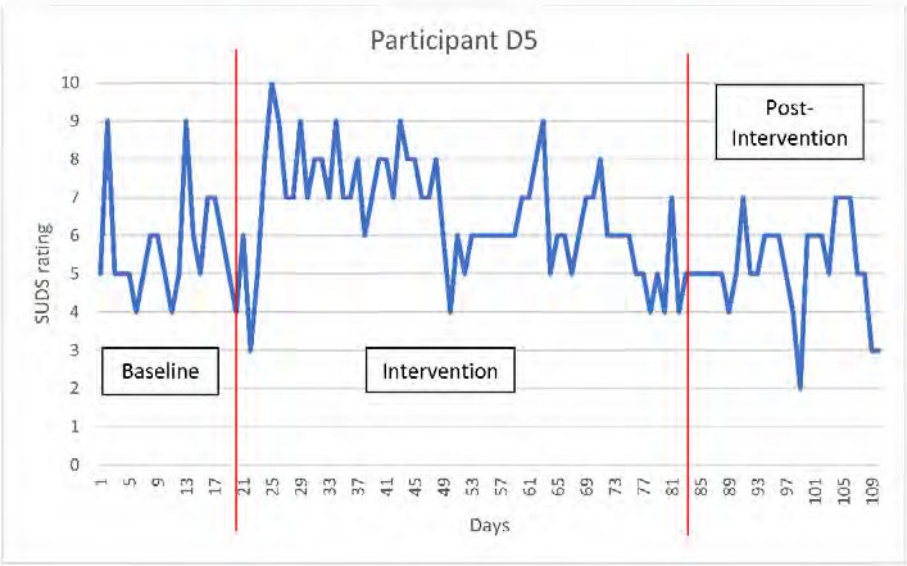


Figure 10.9

SUDS Ratings for Participant D5 Using MindShift



Destressify. Six participants used the *Destressify* app. Visual inspection revealed that all but Participant E1 were experiencing noticeable feelings of distress at baseline. By post-intervention, four participants had achieved a reduction in subjective distress; three to a non-noticeable level (i.e., a rating of < 3). Table 10.15 shows the mean SUDS ratings per participant by phase; Figures 10.10 and 10.11 display the continuous data.

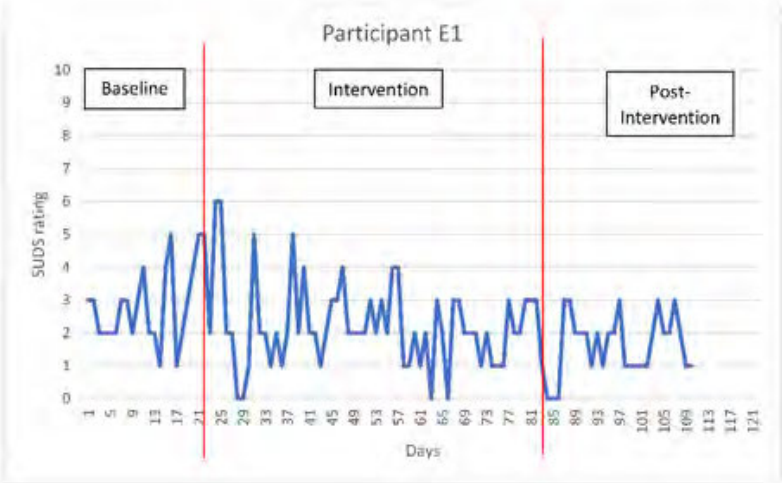
Table 10.15

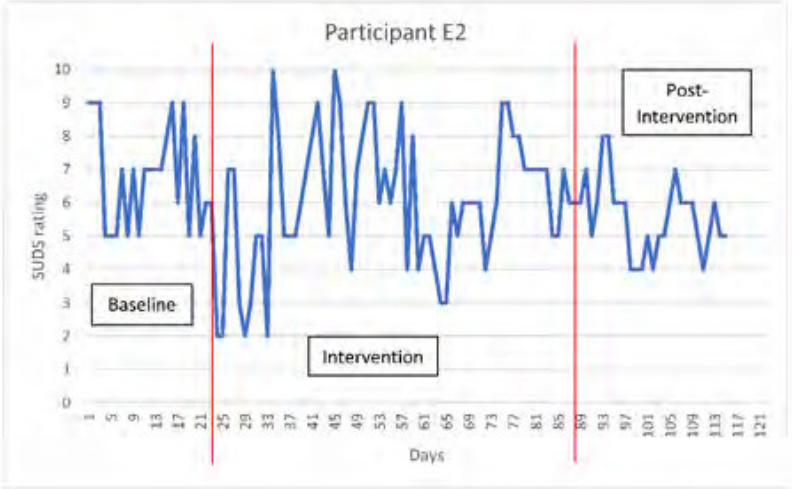
SUDS Data Summary for Participants Using Destressify

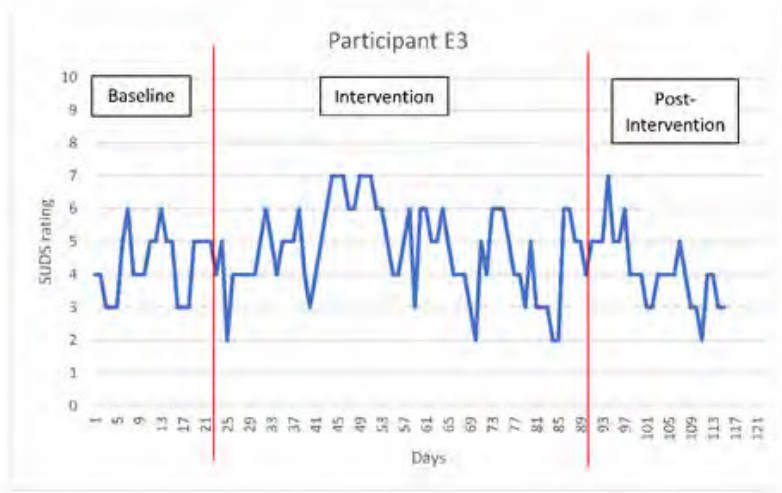
Participant		Baseline SUDS	Intervention SUDS	Post-intervention SUDS
E1	*M*	2.9	2.2	1.6
	SD	1.2	1.3	0.9
	Frequency (*n*)	22	61	27
E2	*M*	6.8	6.1	5.6
	SD	1.5	2.0	1.1
	Frequency (*n*)	23	65	27
E3	*M*	4.3	4.8	4.1
	SD	1.0	1.3	1.1
	Frequency (*n*)	22	68	25
E4	*M*	5.1	4.3	1.8
	SD	0.9	2.3	0.9
	Frequency (*n*)	24	71	24
E5	*M*	7.1	6.2	4.7
	SD	1.1	1.7	0.8
	Frequency (*n*)	22	68	26
E6	*M*	5.1	2.6	1.8
	SD	1.1	1.2	1.0
	Frequency (*n*)	22	65	24

Figure 10.10

SUDS Ratings for Participants E1 – E4 Using Destressify







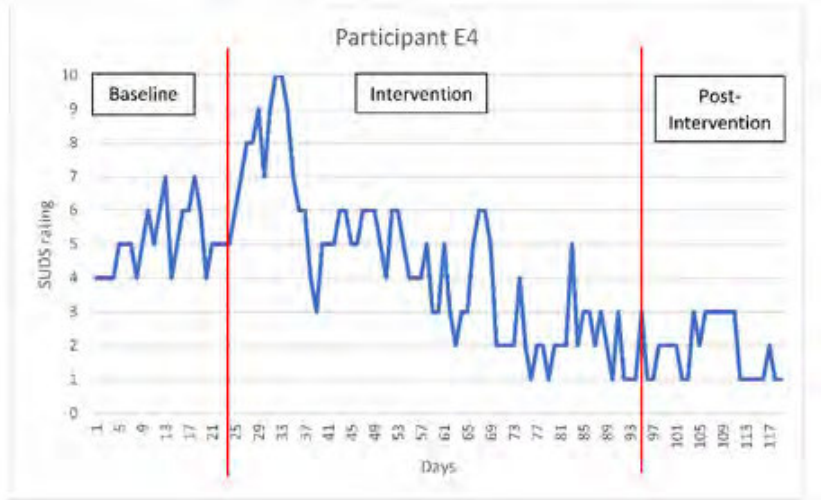
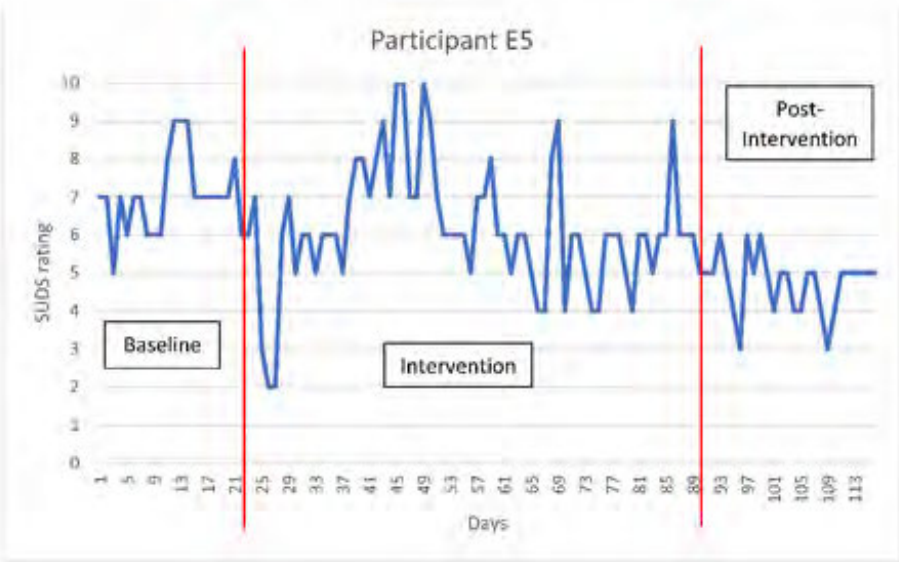
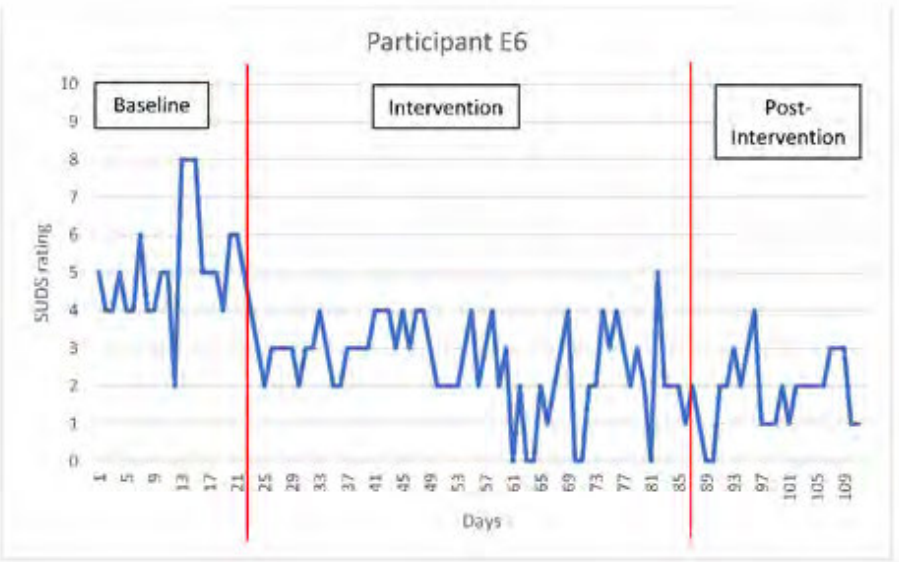


Figure 10.11

SUDS Ratings for Participants E5 – E6 Using Destressify





Time-Series Analysis

Time-series analyses confirmed the findings observed through visual inspection of the plotted SUDS data for each app. Using the statistical package, *R*, version 1.2.5033, an interrupted time series analysis (ITSA) used autoregressive integrated moving average (ARIMA) models to evaluate intervention effects on each participant's data. Autocorrelation effects were addressed using the augmented Dickey-Fuller Test (Mushtaq, 2011) and Ljung-box Q (Burns, 2002). The residuals in the models exhibited independence and normality.

The SUDS time-series analysis data are presented in Tables 10.16 to 10.21 below and can be matched to the relevant participants in Figures to 10.2 to 10.11 above. These are reported below by app.

SuperBetter. Statistically significant decreases in SUDS ratings from baseline to post-intervention were observed for Participants A2, A3, A4, and A5; however, Participant A1 did not show a statistically significant decrease in SUDS from baseline to post-intervention (see Table 10.16), which confirms the conclusion from visual inspection of Participant A1's graph in Figure 10.2. The change in SUDS ratings within the baseline phase was not significant for any participant. This indicates that all four participants had a stable baseline phase prior to the intervention. The change in SUDS ratings within the intervention phase was significant for Participants A2 and A3. The change in SUDS ratings within the post-intervention phase was significant for Participants A3 and A5. This indicates that there continued to be statistically significant improvements for A3 and A5 even after they had ceased using the app (see Table 10.16).

Table 10.16

Times Series Analysis Results for Participants A1 – A5 Using SuperBetter

Participant	Baseline	Intervention	Post-intervention	Overall (baseline to post-intervention)
A1	$t = 1.18$, $p = 0.26$	$t = -0.80$, $p = 0.43$	$t = -1.85$, $p = 0.08$	$t = -1.13$, $p = 0.27$
A2	$t = -1.15$, $p = 0.27$	$t = -4.87$, $p = < .001$	$t = 2.00$, $p = 0.08$	$t = -5.51$, $p = < .001$
A3	$t = -1.57$, $p = 0.13$	$t = 3.23$, $p = 0.00$	$t = -4.04$, $p = 0.00$	$t = 4.89$, $p = < .001$
A4	$t = -6.78$, $p = < .001$	$t = -0.93$, $p = 0.36$	$t = -0.54$, $p = 0.60$	$t = 3.31$, $p = < .001$
A5	$t = -0.75$, $p = 0.46$	$t = -0.64$, $p = 0.53$	$t = -2.23$, $p = 0.04$	$t = -2.33$, $p = 0.02$

Smiling Mind. Statistically significant decreases in SUDS ratings from baseline to post-intervention were observed for Participants B1, B2, B3, B4, B5, and B7; however, Participant B6 did not show a statistically significant decrease in SUDS from baseline to post-intervention (see Table 10.17), which confirms the conclusion from visual inspection of Participant B6's graph in Figure 10.5. The change in SUDS ratings within the baseline phase was significant for Participants B1 and B6, suggesting that they did not have a stable baseline phase. The change in SUDS ratings within the intervention phase was significant for B1, B2, B5, and B7, suggesting that improvements in SUDS ratings during the intervention phase for these participants were due to the intervention and were maintained post-intervention. All seven participants had statistically stable post-intervention phases (see Table 10.17).

Table 10.17

Time Series Analysis Results for Participants B1 – B7 Using Smiling Mind

Participant	Baseline	Intervention	Post-intervention	Overall (baseline to post-intervention)
B1	$t = 3.07$, $p = 0.01$	$t = -4.74$, $p = < .001$	$t = 0.00$, $p = 1.00$	$t = -5.52$, $p = < .001$
B2	$t = 0.45$, $p = 0.66$	$t = -7.87$, $p = < .001$	$t = -1.24$, $p = 0.43$	$t = -10.19$, $p = < .001$
B3	$t = -0.80$, $p = 0.43$	$t = -1.17$, $p = 0.25$	$t = -1.24$, $p = 0.23$	$t = -4.40$, $p = < .001$
B4	$t = -0.39$, $p = 0.70$	$t = -0.90$, $p = 0.37$	$t = -0.44$, $p = 0.66$	$t = -2.52$, $p = 0.01$
B5	$t = 1.61$, $p = 0.12$	$t = -8.59$, $p = < .001$	$t = -0.64$, $p = 0.53$	$t = -9.11$, $p = < .001$
B6	$t = 3.48$, $p = 0.00$	$t = -0.98$, $p = 0.33$	$t = -1.87$, $p = 0.07$	$t = -1.69$, $p = 0.10$
B7	$t = 0.47$, $p = 0.65$	$t = -5.58$, $p = < .001$	$t = -0.33$, $p = 0.75$	$t = -8.78$, $p = < .001$

MoodMission. Statistically significant decreases in SUDS ratings from baseline to post-intervention were observed for Participants C1, C2, C3, C4, and C5; however, Participant C6 did not show a statistically significant decrease in SUDS from baseline to post-intervention (see Table 10.18), which confirms the conclusion from visual inspection of Participant C6's graph in Figure 10.7. The change in SUDS ratings within the baseline phase was significant for Participant C5, suggesting that Participant C5 did not have a stable baseline phase. The change in SUDS ratings within the intervention phase was significant for all six participants, suggesting that improvements in SUDS ratings during the intervention phase for all participants were due to the intervention and were maintained post-intervention. The change in SUDS ratings during the post-intervention phase was significant for Participant C5. This indicates that there continued to be statistically significant improvements for C5 even after she had ceased using the app (see Table 10.18).

Table 10.18

Time Series Analysis Results for Participants C1 – C6 Using MoodMission

Participant	Baseline	Intervention	Post-intervention	Overall (baseline to post-intervention)
C1	$t = 1.58$, $p = 0.13$	$t = -3.15$, $p = 0.00$	$t = 0.31$, $p = 0.76$	$t = -4.52$, $p = < .001$
C2	$t = -0.69$, $p = 0.50$	$t = -2.67$, $p = 0.01$	$t = -1.67$, $p = 0.11$	$t = -5.32$, $p = < .001$
C3	$t = -1.48$, $p = 0.15$	$t = 3.20$, $p = 0.00$	$t = -1.66$, $p = 0.11$	$t = 2.48$, $p = 0.02$
C4	$t = -0.50$, $p = 0.62$	$t = -1.82$, $p = 0.07$	$t = 0.55$, $p = 0.59$	$t = -2.99$, $p = 0.01$
C5	$t = -5.31$, $p = < .001$	$t = -6.32$, $p = < .001$	$t = -2.60$, $p = 0.02$	$t = -4.73$, $p = < .001$
C6	$t = -0.31$, $p = 0.76$	$t = -2.08$, $p = 0.04$	$t = -1.72$, $p = 0.10$	$t = -0.36$, $p = 0.72$

MindShift. Statistically significant decreases in SUDS ratings from baseline to post-intervention were observed for Participants D2 and D3; however, Participants D1, D4, and D5 did not show a statistically significant decrease in SUDS from baseline to post-intervention (see Table 10.19), which confirms the conclusion from visual inspection of Participant D1, D4, and D5s' graphs in Figures 10.8 and 10.9. The change in SUDS ratings within the baseline phase was not significant for any participant, suggesting that all participants had stable baseline phases. The change in SUDS ratings within the intervention phase was significant for Participants D3, D4, and D5, suggesting that improvements in SUDS ratings during the intervention phase for Participants D3, D4, and D5 were due to the intervention, but were only maintained post-intervention by D3. The change in SUDS ratings within the post-intervention phases for Participants D1 and D2 were significant. This indicates that there continued to be statistically significant improvements for D1 and D2 even after they had ceased using the app (see Table 10.19).

Table 10.19

Time Series Analysis Results for Participants D1 – D5 Using MindShift

Participant	Baseline	Intervention	Post-intervention	Overall (baseline to post-intervention)
D1	$t = 2.67$, $p = 0.12$	$t = 1.483$, $p = 0.14$	$t = -2.28$, $p = 0.03$	$t = -1.42$, $p = 0.16$
D2	$t = -0.48$, $p = 0.64$	$t = -1.33$, $p = 0.19$	$t = -2.16$, $p = 0.04$	$t = -2.23$, $p = 0.03$
D3	$t = 0.59$, $p = 0.56$	$t = -8.02$, $p = < .001$	$t = 1.20$, $p = 0.24$	$t = -6.73$, $p = <.001$
D4	$t = -0.02$, $p = 0.98$	$t = -4.37$, $p = < .001$	$t = 0.46$, $p = 0.65$	$t = -1.93$, $p = 0.06$
D5	$t = 0.04$, $p = 0.97$	$t = -3.99$, $p = < .001$	$t = -0.14$, $p = 0.89$	$t = -0.67$, $p = 0.51$

Destressify. Statistically significant decreases in SUDS ratings from baseline to post-intervention were observed for Participants E1, E3, E4, E5, and E6; however, Participant E2 did not show a statistically significant decrease in SUDS from baseline to post-intervention (see Table 10.20), which confirms the conclusion from visual inspection of Participant E2's graph in Figure 10.10. The change in SUDS ratings within the baseline phase was not significant for any participant, suggesting that all had stable baseline phases. The change in SUDS ratings within the intervention phase was significant for Participants E4 and E6, suggesting that improvements in SUDS ratings during the intervention phase for Participants E4 and E6 were due to the intervention and were maintained post-intervention. The change in SUDS ratings in the post-intervention phase for Participant E3 was significant. This indicates that there continued to be a statistically significant improvement for E3 even after he had ceased using the app (see Table 10.20).

Table 10.20

Time Series Analysis Results for Participants E1 – E5 Using Destressify

Participant	Baseline	Intervention	Post-intervention	Overall (baseline to post-intervention)
E1	$t = 1.98$, $p = 0.06$	$t = -1.36$, $p = 0.18$	$t = 1.01$, $p = 0.32$	$t = -3.26$, $p = 0.00$
E2	$t = -0.67$, $p = 0.51$	$t = 1.64$, $p = 0.11$	$t = -1.74$, $p = 0.09$	$t = -0.98$, $p = 0.33$
E3	$t = 1.25$, $p = 0.23$	$t = -1.21$, $p = 0.23$	$t = -4.498$, $p = < .001$	$t = 2.67$, $p = 0.01$
E4	$t = 1.99$, $p = 0.06$	$t = -11.97$, $p = < .001$	$t = -0.67$, $p = 0.51$	$t = -7.16$, $p = < .001$
E5	$t = 1.25$, $p = 0.22$	$t = -0.53$, $p = 0.60$	$t = -0.65$, $p = 0.52$	$t = -5.10$, $p = < .001$
E6	$t = 1.47$, $p = 0.16$	$t = -2.71$, $p = 0.01$	$t = 1.18$, $p = 0.25$	$t = -6.23$, $p = < .001$

Effectiveness of the Apps for Reducing Anxiety

The severity scores of each participant's symptoms of anxiety, measured by the DASS-21 Anxiety subscale, shown in Tables 10.21 to 10.25, and illustrated in Figures 10.12 to 10.16, are reported by app below.

SuperBetter

Three participants showed improvement in their symptoms of anxiety but two did not. See Figure 10.12 below for a summary of the *SuperBetter* anxiety outcomes.

Table 10.21

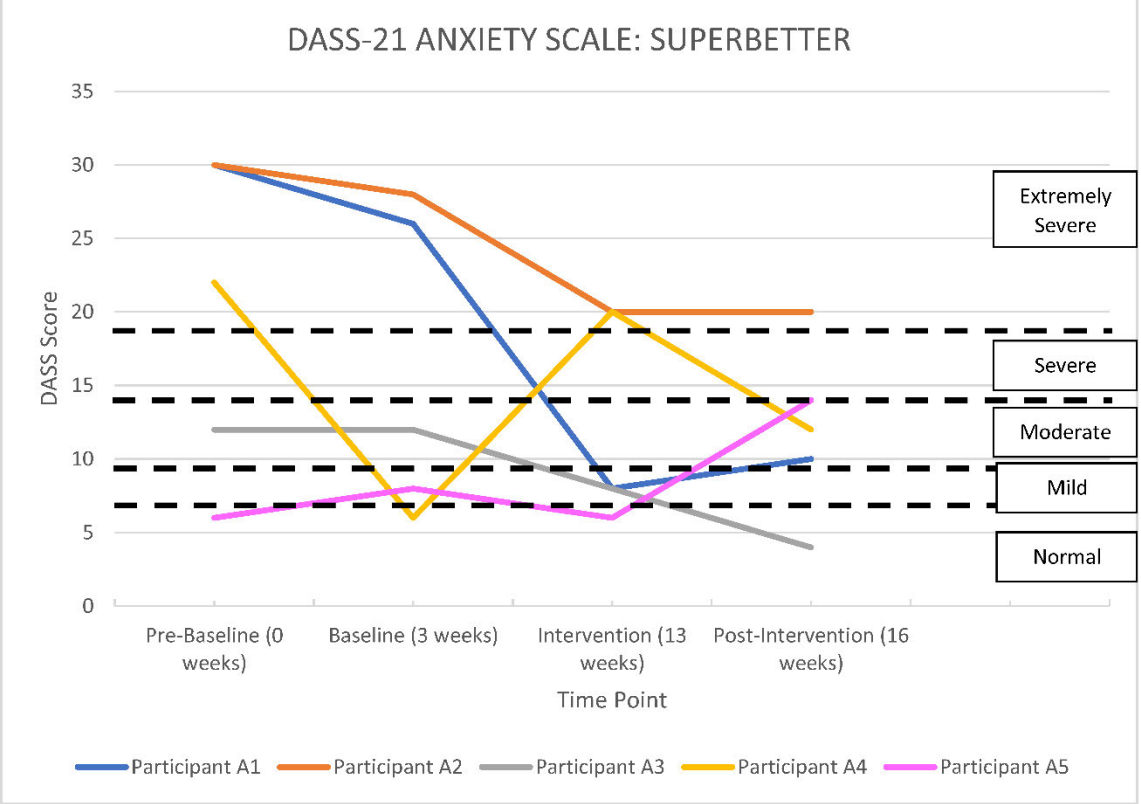
DASS-21 Scores for Participants A1-A5 Using SuperBetter

Participant	Scale	DASS Score at Beginning of Baseline	DASS Score at Beginning of Intervention	DASS Score at Beginning of Post-intervention	DASS Score at end of Post-intervention
A1	Depression	30	36	2	12
	Anxiety	30	26	8	10
A2	Depression	32	20	18	20
	Anxiety	30	28	20	20
A3	Depression	18	6	10	8
	Anxiety	12	12	8	4
A4	Depression	10	2	30	22
	Anxiety	22	6	20	12
A5	Depression	26	24	16	16
	Anxiety	6	8	6	14

Note. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).

Figure 10.12

DASS-21 Anxiety Scale Scores for Participants A1 – A5 Using SuperBetter



Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Smiling Mind

All Smiling Mind participants showed improvement in their symptoms of anxiety. See Figure 10.13 below for a summary of the *Smiling Mind* anxiety outcomes.

Table 10.22

DASS-21 Scores for Participants B1 – B7 Using Smiling Mind

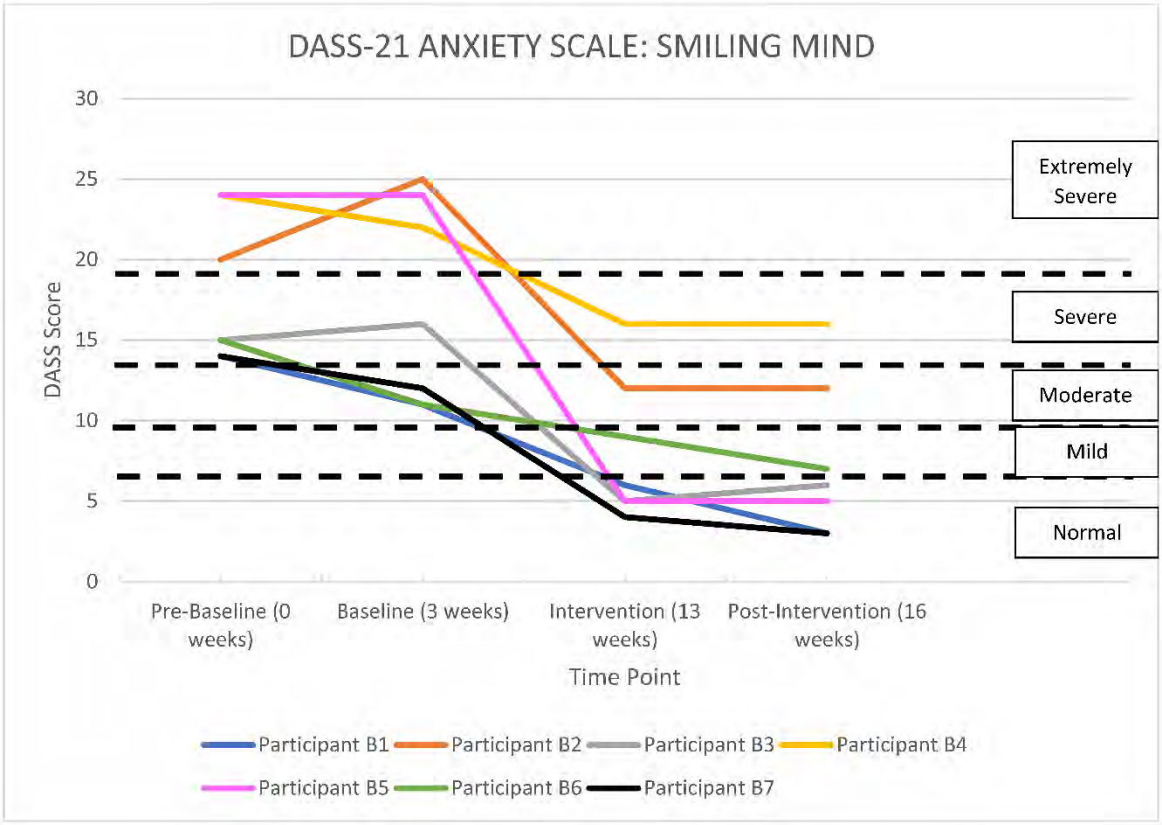
Participant	Scale	DASS Score at Beginning of Baseline	DASS Score at Beginning of Intervention	DASS Score at Beginning of Post-intervention	DASS Score at end of Post-intervention
B1	Depression	15	16	9	8
	Anxiety	14	11	6	3
B2	Depression	25	27	12	10
	Anxiety	20	25	12	12
B3	Depression	16	16	6	7
	Anxiety	15	16	5	6
B4	Depression	28	24	16	19
	Anxiety	24	22	16	16
B5	Depression	25	26	8	5
	Anxiety	24	24	5	5
B6	Depression	12	10	8	7

	Anxiety	15	11	9	7
B7	Depression	14	18	6	6
	Anxiety	14	12	4	3

Note. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).

Figure 10.13

DASS-21 Anxiety Scale Scores for Participants B1 – B7 Using Smiling Mind



Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

MoodMission

Five participants showed improvement in their symptoms of anxiety, but one did not. See Figure 10.14 below for a summary of the *MoodMission* anxiety outcomes.

Table 10.23

DASS-21 Scores for Participants C1-C6 Using MoodMission

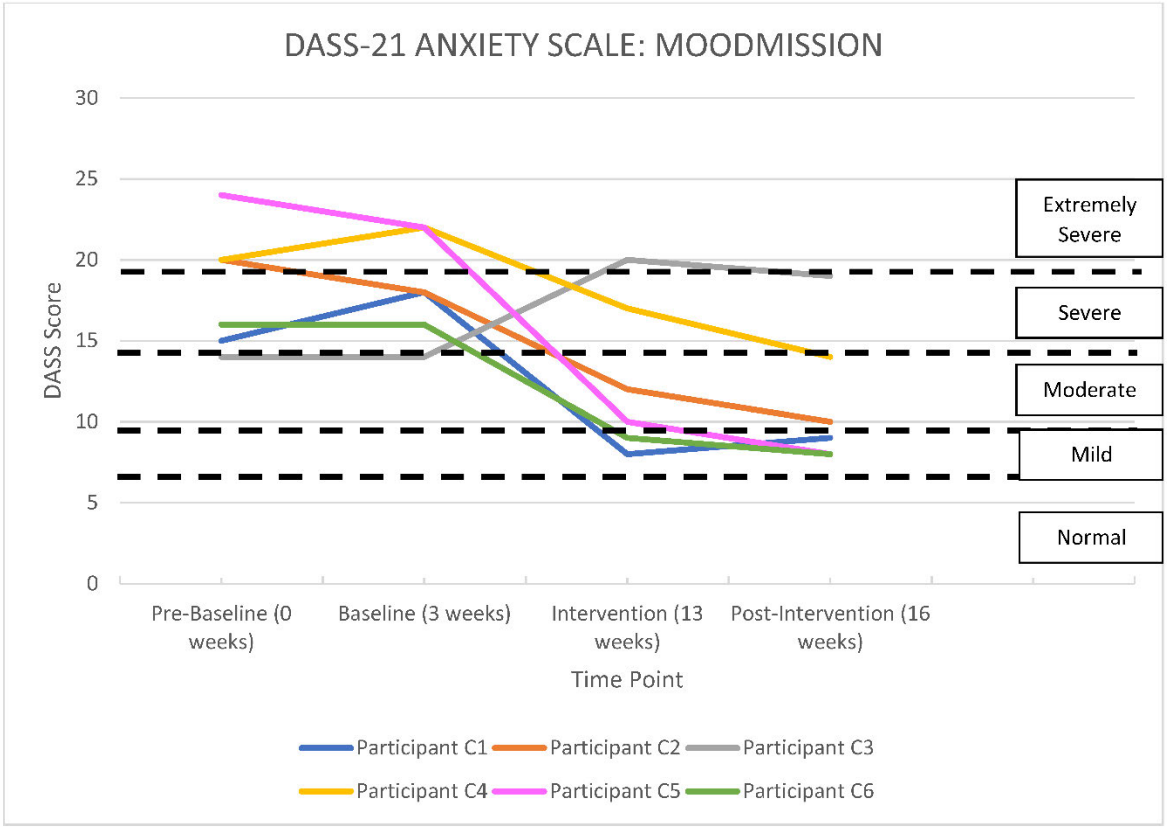
Participant	Scale	DASS Score at Beginning of Baseline	DASS Score at Beginning of Intervention	DASS Score at Beginning of Post-intervention	DASS Score at end of Post-intervention
C1	Depression	26	20	14	10
	Anxiety	15	18	8	9
C2	Depression	26	27	20	21
	Anxiety	20	18	12	10
C3	Depression	13	11	18	14
	Anxiety	14	14	20	19
C4	Depression	12	12	7	9
	Anxiety	20	22	17	14
C5	Depression	14	18	8	6
	Anxiety	24	22	10	8
C6	Depression	8	10	7	7

Anxiety	16	16	9	8

Note. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).

Figure 10.14

DASS-21 Anxiety Scale Scores for Participants C1 – C6 Using MoodMission



Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

MindShift

All participants showed improvement in their symptoms of anxiety. See Figure 10.15 below for a summary of the *MindShift* anxiety outcomes.

Table 10.24

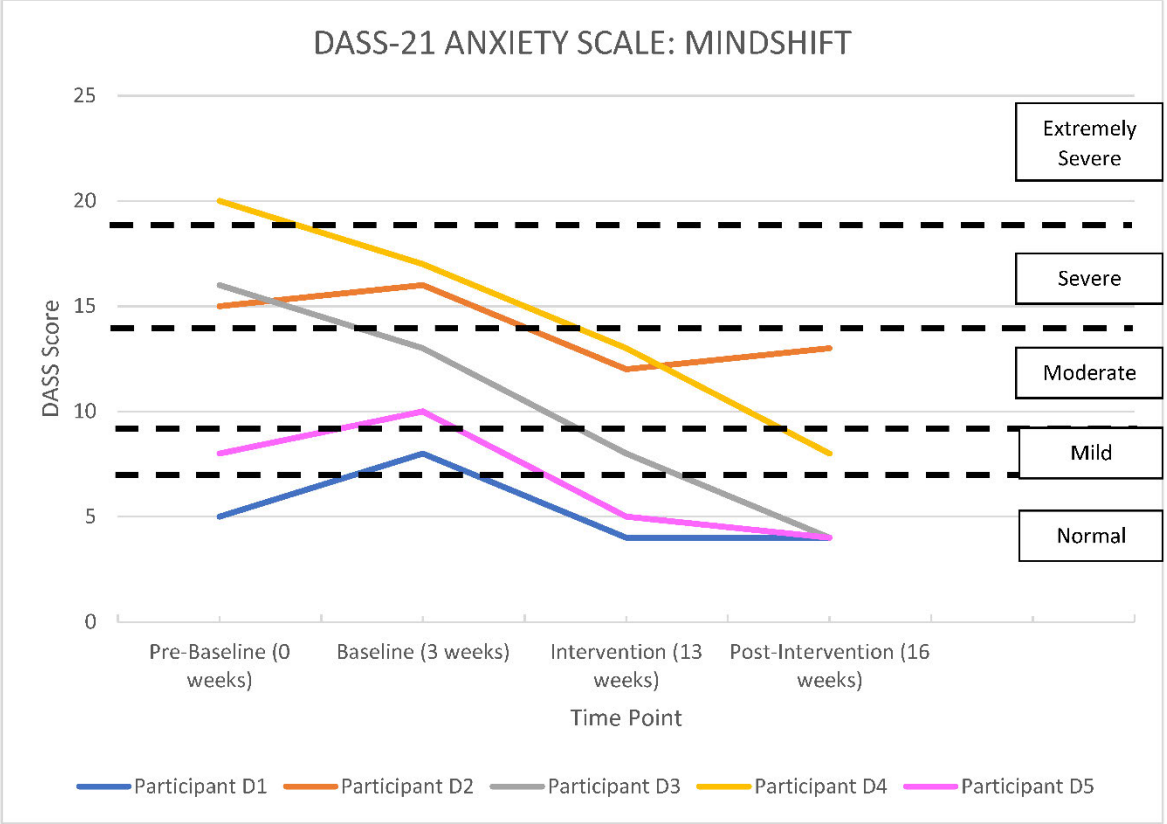
DASS-21 Scores for Participants D1 – D5 Using MindShift

Participant	Scale	DASS Score at Beginning of Baseline	DASS Score at Beginning of Intervention	DASS Score at Beginning of Post-intervention	DASS Score at end of Post-intervention
D1	Depression	12	19	21	17
	Anxiety	5	8	4	4
D2	Depression	12	10	8	7
	Anxiety	15	16	12	13
D3	Depression	14	13	9	6
	Anxiety	16	13	8	4
D4	Depression	8	8	6	7
	Anxiety	20	17	13	8
D5	Depression	25	24	22	22
	Anxiety	8	10	5	4

Note. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).

Figure 10.15

DASS-21 Anxiety Scale Scores for Participants D1 – D5 Using MindShift



Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Destressify

Five participants showed improvement in their symptoms of anxiety, but one did not. See Figure 10.16 below for a summary of the *Destressify* anxiety outcomes.

Table 10.25

DASS-21 Scores for Participants E1 – E6 Using Destressify

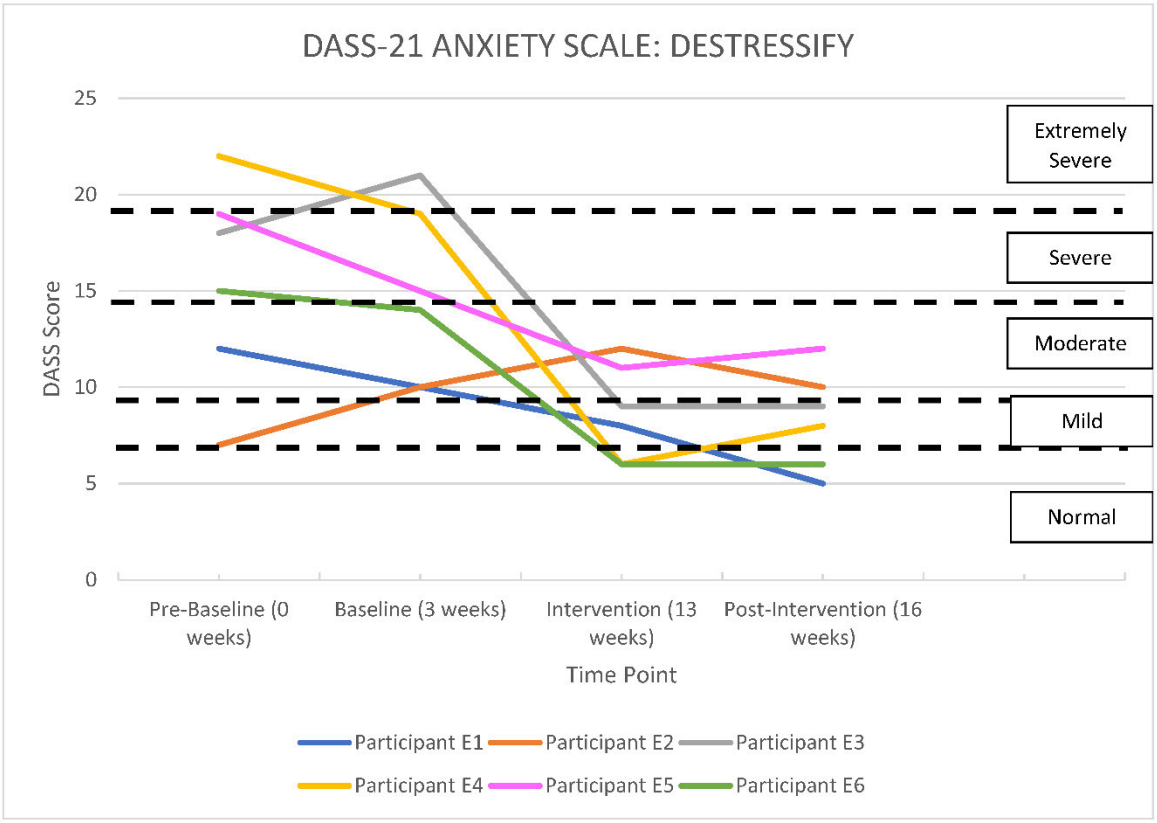
Participant	Scale	DASS Score at Beginning of Baseline	DASS Score at Beginning of Intervention	DASS Score at Beginning of Post-intervention	DASS Score at end of Post-intervention
E1	Depression	8	7	6	6
	Anxiety	12	10	8	5
E2	Depression	28	24	19	21
	Anxiety	7	10	12	10
E3	Depression	19	22	18	16
	Anxiety	18	21	9	9
E4	Depression	7	6	6	7
	Anxiety	22	19	6	8
E5	Depression	22	24	24	25
	Anxiety	19	15	11	12
E6	Depression	15	17	10	10

Anxiety	15	14	6	6

Note. The ratings for the depression subscale are 0-9 (*Normal*), 10-13 (*Mild*), 14-20 (*Moderate*), 21-27 (*Severe*) and 28+ (*Extremely Severe*); and, for the anxiety subscale are 0-7 (*Normal*), 8-9 (*Mild*), 10-14 (*Moderate*), 15-19 (*Severe*) and 20+ (*Extremely Severe*).

Figure 10.16

DASS-21 Anxiety Scale Scores for Participants E1 – E6 Using Destressify



Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Effectiveness of the Apps for Reducing Depression

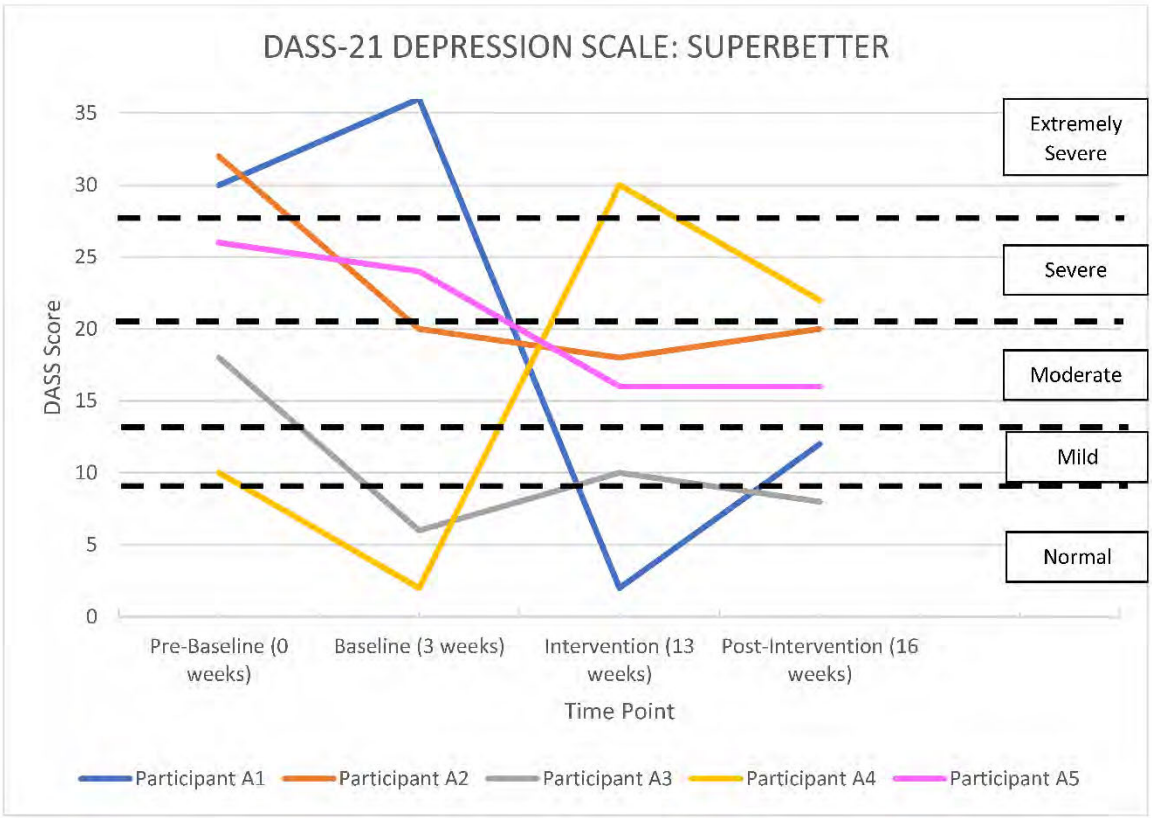
The severity scores of each participant's symptoms of depression, measured by the DASS-21 Depression subscale, shown in Tables 10.21 to 10.25 above, and illustrated in Figures 10.17 to 10.21 below, are reported by app below.

SuperBetter

Four of the five participants showed improvement in symptoms of depression. See Figure 10.17 for a summary of *SuperBetter* depression outcomes.

Figure 10.17

DASS-21 Depression Scale Scores for Participants A1 – A5 Using SuperBetter



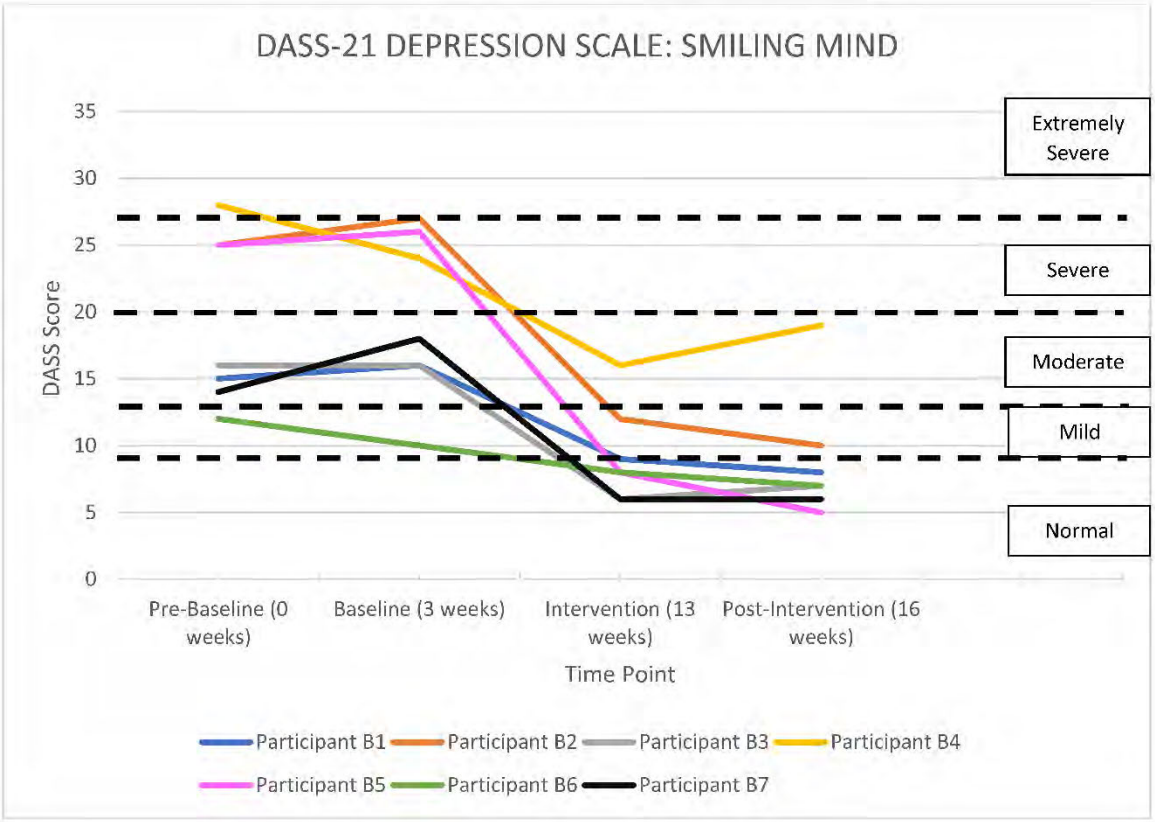
Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Smiling Mind

All seven participants showed improvement in symptoms of depression. See Figure 10.18 for a summary of *Smiling Mind* depression outcomes.

Figure 10.18

DASS-21 Depression Scale Scores for Participants B1 – B7 Using Smiling Mind



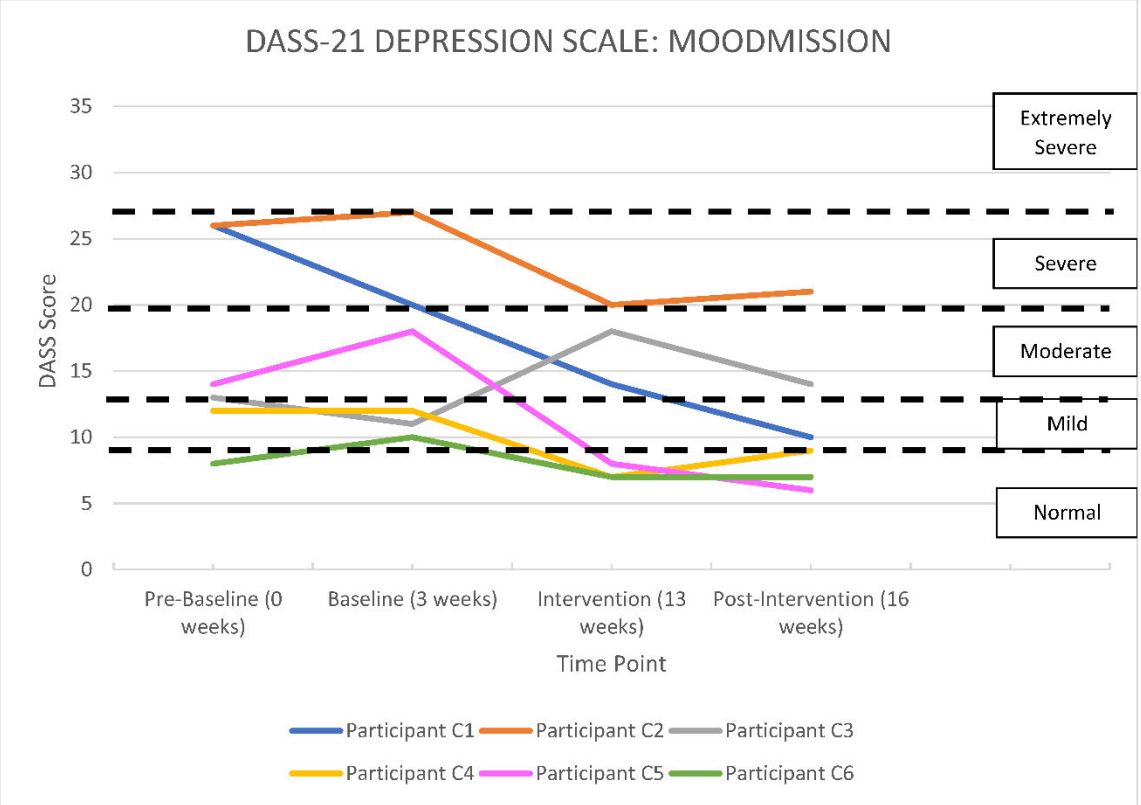
Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

MoodMission

Two participants showed improvement in symptoms of depression, three remained unchanged, and one had worsened. See Figure 10.19 for a summary of *MoodMission* depression outcomes.

Figure 10.19

DASS-21 Depression Scale Scores for Participants C1- C6 Using MoodMission



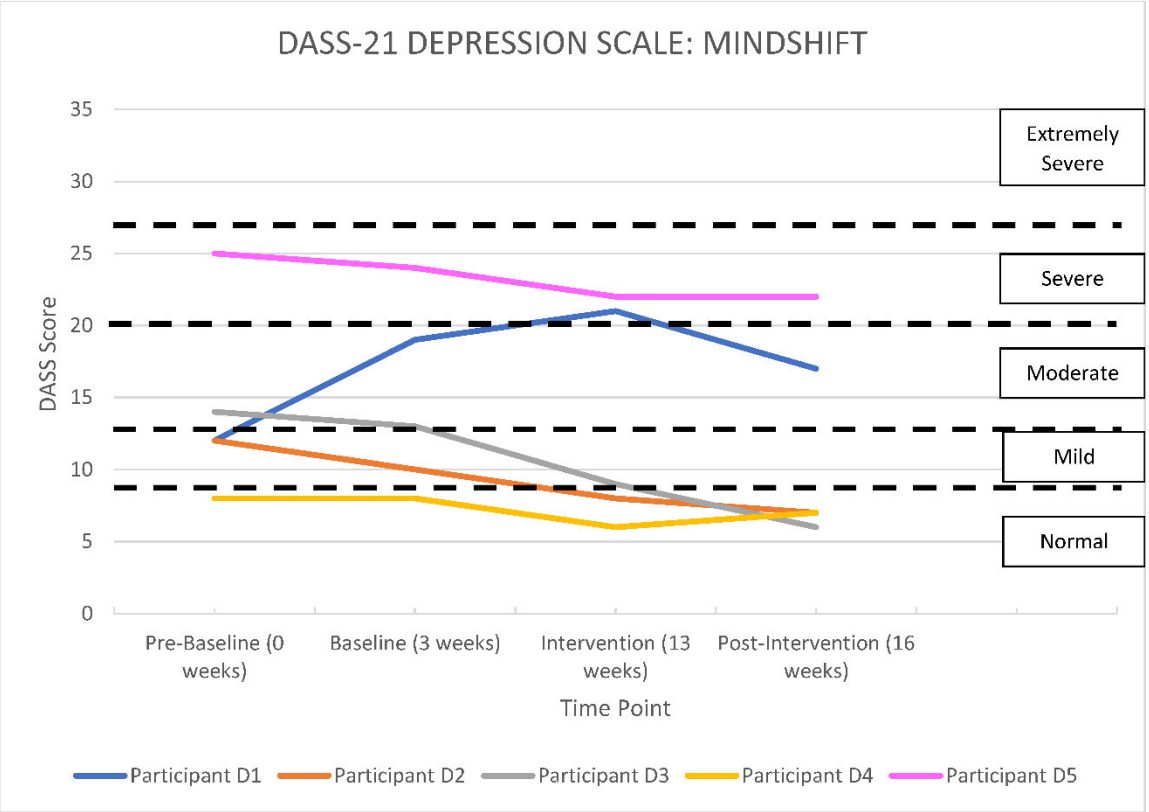
Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

MindShift

Two participants showed improvement in symptoms of depression, two remained unchanged (one of whom was not depressed throughout), and one had worsened. See Figure 10.20 for a summary of *MindShift* depression outcomes.

Figure 10.20

DASS-21 Depression Scale Scores for Participants D1 – D5 Using MindShift



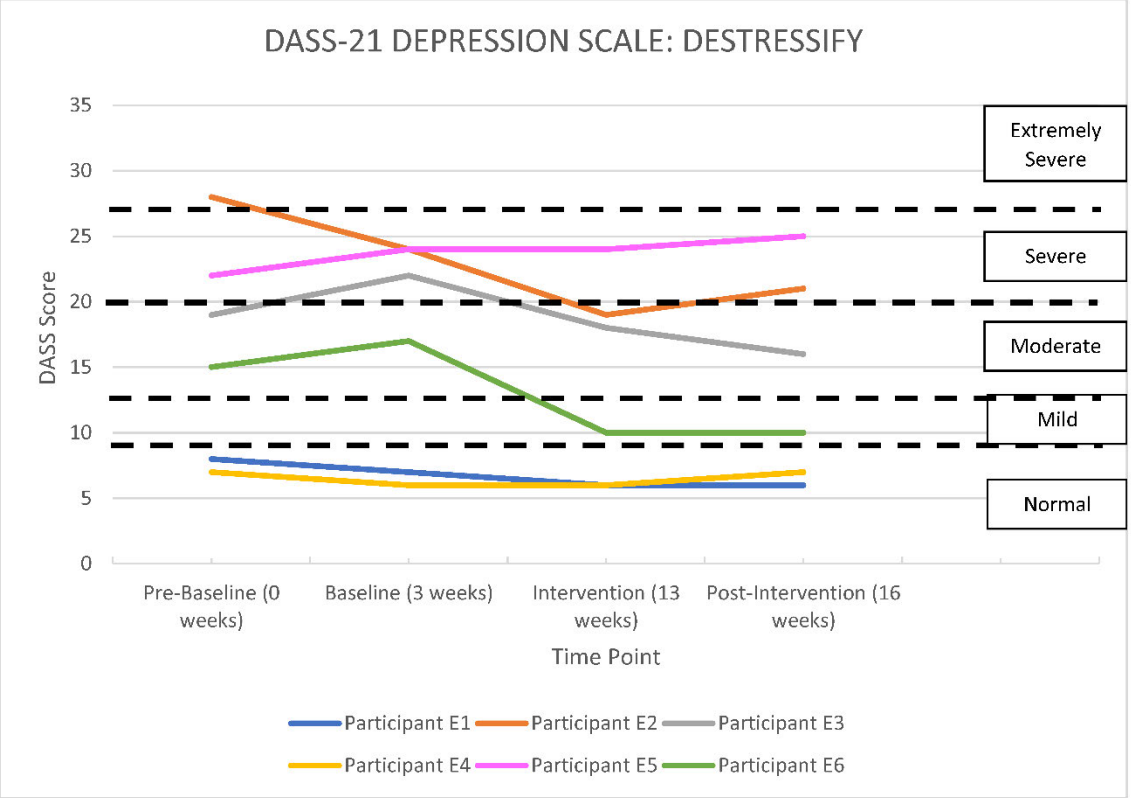
Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Destressify

Two participants showed improvement in symptoms of depression and four remained unchanged (including two who were not depressed throughout). See Figure 10.21 for a summary of *Destressify* depression outcomes.

Figure 10.21

DASS-21 Depression Scale Scores for Participants E1 – E6 Using Destressify



Note. Descriptive severity labels from the DASS manual (Lovibond & Lovibond, 1995).

Effectiveness of the Apps in Improving Life Functioning

Each participant's overall functioning, measured by the OQ-45 Total Score, is shown in Tables 10.26 to 10.30, and illustrated in Figures 10.22 to 10.26, and reported by app below.

SuperBetter

Table 10.26 and Figure 10.22 show a summary of life functioning outcomes for participants using *SuperBetter*. As can be seen in Figure 10.22, at pre-baseline (Phase 1) all participants recorded a clinically significant impairment in life functioning. At post-intervention (Phase 4), all participants were improved except for Participant A4. The improvements shown by Participants A1, A2, A3 and A5 were clinically significant, with Participants A1 and A3 dropping to the non-clinical range.

Table 10.26

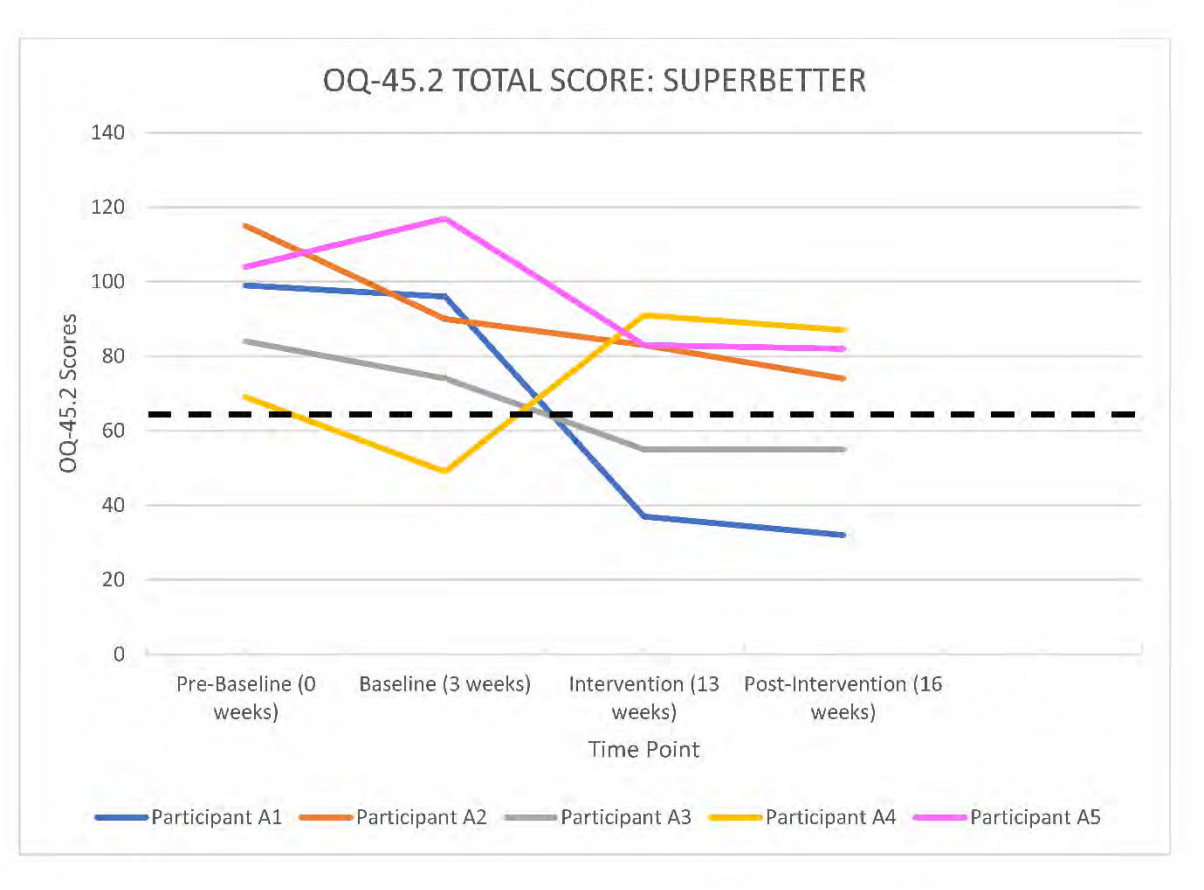
OQ-45.2 Total Scores for Participants A1 – A5 Using SuperBetter

Participant	Pre-baseline to post-baseline	Post-baseline to post-intervention	Post-intervention to post-3-week follow up	Overall (pre-baseline to post-3-week follow up)	Classification (pre-baseline to post-3-week follow up)	Improvement (pre-baseline to post-3-week follow up)
A1	99-96 ^	96-37 *+	37-32 ^	99-32 *+	Recovered	67.68%
A2	115-90 *	90-83 ^	83-74 ^	115-74 *	Improved	35.65%
A3	84-74 ^	74-55 *+	55-55 ^	84-55 *+	Recovered	34.52%
A4	69-49 *+	(49-91) *+	91-87 ^	(69-87) *	Deteriorated	26.09%
A5	(104-117) ^	117-83 *	83-82 ^	104-82 *	Improved	21.15%

Note. Based on RCI = 14 and CSI = 63 for Total Score. ^ denotes not statistically reliable, not clinically significant; * denotes statistically reliable ($p < .05$); + denotes clinically significant; *+ denotes statistically reliable and clinically significant. Parentheses indicate change in a worsening direction. *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

Figure 10.22

OQ-45.2 Total Scores for Participants A1 – A5 Using SuperBetter



Note. Clinical cut-off of 63 from the OQ-45.2 technical manual (Lambert et al., 2004).

Smiling Mind

Table 10.27 and Figure 10.23 show a summary of life functioning outcomes for participants using *Smiling Mind*. As can be seen in Figure 10.23, at pre-baseline (Phase 1) all participants except for B6 recorded a clinically significant impairment in life functioning. At post-intervention (Phase 4), all participants were improved, with B4 the only participant still in the clinically significant range. The improvements shown by all participants were clinically significant, with Participants B1, B2, B3, B5, and B7 classified as recovered.

Table 10.27

OQ-45.2 Total Scores for Participants B1 – B7 Using Smiling Mind

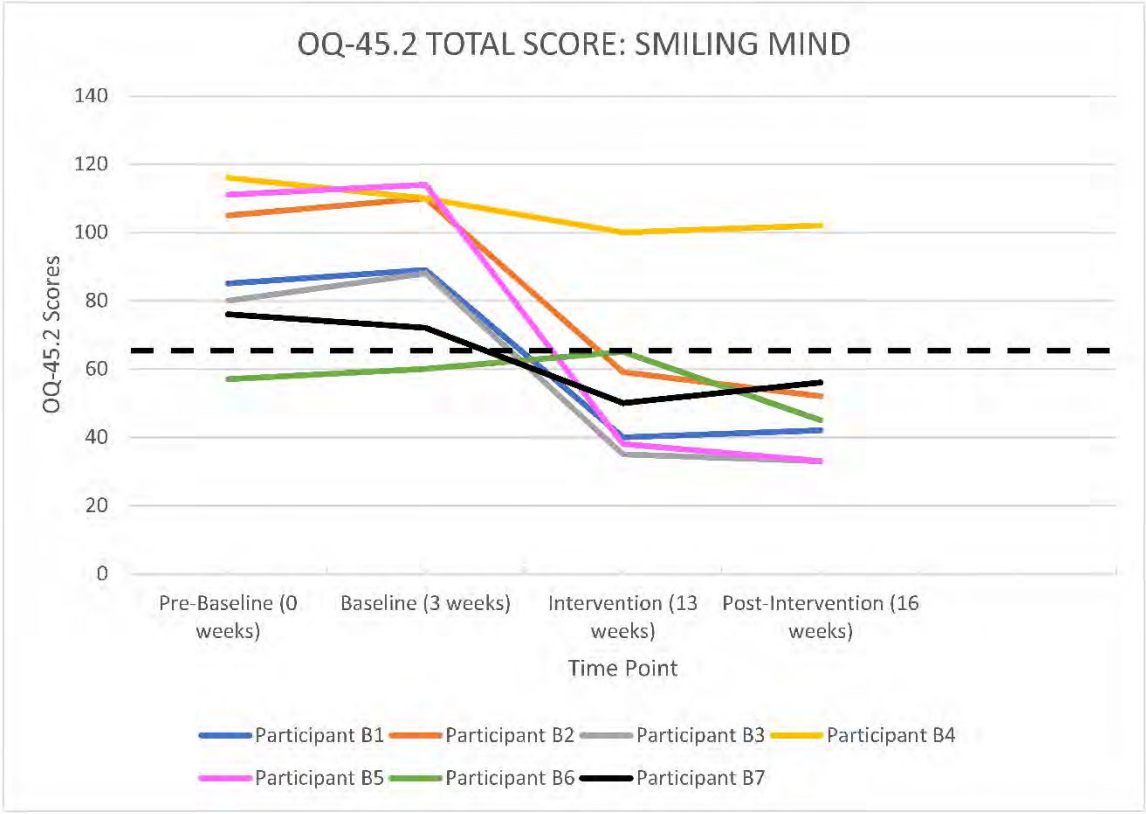
Participant	Pre-baseline to post-baseline	Post-baseline to post-intervention	Post-intervention to post-3-week follow up	Overall (pre-baseline to post-3-week follow up)	Classification (pre-baseline to post-3-week follow up)	Improvement (pre-baseline to post-3-week follow up)
B1	(85-89) ^	89-40 *+	(40-42) ^	85-42 *+	Recovered	50.59%
B2	(105-110) ^	110-59 *+	59-52 ^	105-52 *+	Recovered	50.47%
B3	(80-88) ^	88-35 *+	35-33 ^	80-33 *+	Recovered	58.75%
B4	116-110 ^	110-100 ^	(100-102) ^	116-102 *	Improved	12.07%
B5	(111-114) ^	114-38 *+	38-33 ^	111-33 *+	Recovered	70.27%
B6	(57-60) ^	(60-65) +	65-45 *+	57-45 *	Unchanged [a]	21.05%
B7	76-72 ^	72-50 *+	(50-56) ^	76-56 *+	Recovered	26.32%

Note. Based on RCI = 14 and CSI = 63 for Total Score. ^ denotes not statistically reliable, not clinically significant; * denotes statistically reliable ($p < .05$); + denotes clinically significant; *+ denotes statistically reliable and clinically significant. Parentheses indicate change in a worsening direction. *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

[a] Non-clinical to begin with.

Figure 10.23

OQ-45.2 Total Scores for Participants B1 – B7 Using Smiling Mind



Note. Clinical cut-off of 63 from the OQ-45.2 technical manual (Lambert et al., 2004).

MoodMission

Table 10.28 and Figure 10.24 show a summary of life functioning outcomes for participants using *MoodMission*. As can be seen in Figure 10.24, at pre-baseline (Phase 1) all participants except for C3 recorded a clinically significant impairment in life functioning. At post-intervention (Phase 4), all participants were improved except for Participant C3, who remained under the clinically significant threshold. The improvements shown by Participants C1, C4, C5, and C6 were clinically significant, with Participant C2 the only remaining participant in the clinically significant range at post-intervention.

Table 10.28

OQ-45.2 Total Scores for Participants C1 – C6 Using MoodMission

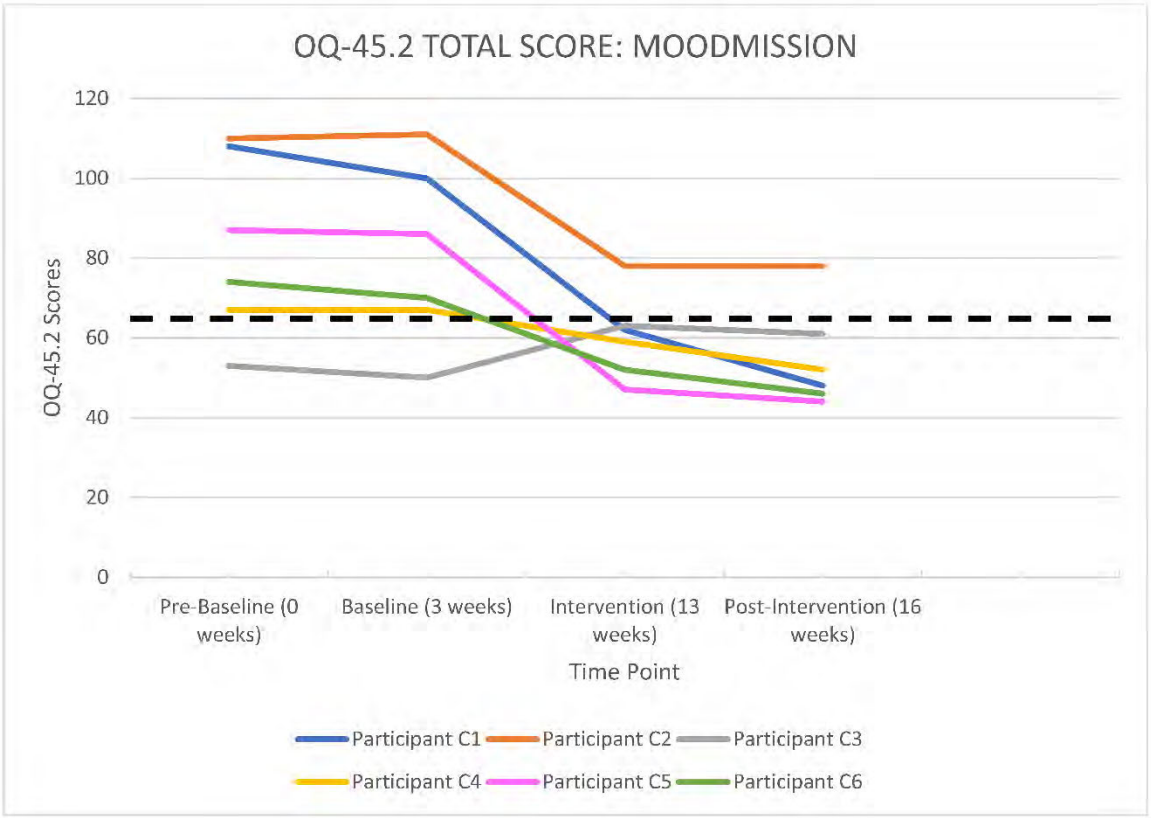
Participant	Pre-baseline to post-baseline	Post-baseline to post-intervention	Post-intervention to post-3-week follow up	Overall (pre-baseline to post-3-week follow up)	Classification (pre-baseline to post-3-week follow up)	Improvement (pre-baseline to post-3-week follow up)
C1	108-100 ^	100-62 *+	62-48 *	108-48 *+	Recovered	55.56%
C2	(110-111) ^	111-78 *	78-78 ^	110-78 *	Improved	29.09%
C3	53-50 ^	(50-63) +	63-61 ^	(53-61) ^	Unchanged [a]	15.09%
C4	67-67 ^	67-59 +	59-52 ^	67-52 *+	Recovered	22.39%
C5	87-86 ^	86-47 *+	47-44 ^	87-44 *+	Recovered	49.43%
C6	74-70 ^	70-52 *+	52-46 ^	74-46 *+	Recovered	37.84%

Note. Based on RCI = 14 and CSI = 63 for Total Score. ^ denotes not statistically reliable, not clinically significant; * denotes statistically reliable ($p < .05$); + denotes clinically significant; *+ denotes statistically reliable and clinically significant. Parentheses indicate change in a worsening direction. *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

[a] Non-clinical to begin with.

Figure 10.24

OQ-45.2 Total Scores for Participants C1 – C6 Using MoodMission



Note. Clinical cut-off of 63 from the OQ-45.2 technical manual (Lambert et al., 2004).

MindShift

Table 10.29 and Figure 10.25 show a summary of life functioning outcomes for participants using *MindShift*. As can be seen in Figure 10.25, at pre-baseline (Phase 1) all participants recorded a clinically significant impairment in life functioning. At post-intervention (Phase 4), all participants were improved. The improvements shown by Participants D2, D3, D4, and D5 were clinically significant, with Participant D5 being the only participant still in the clinically significant range at post-intervention.

Table 10.29

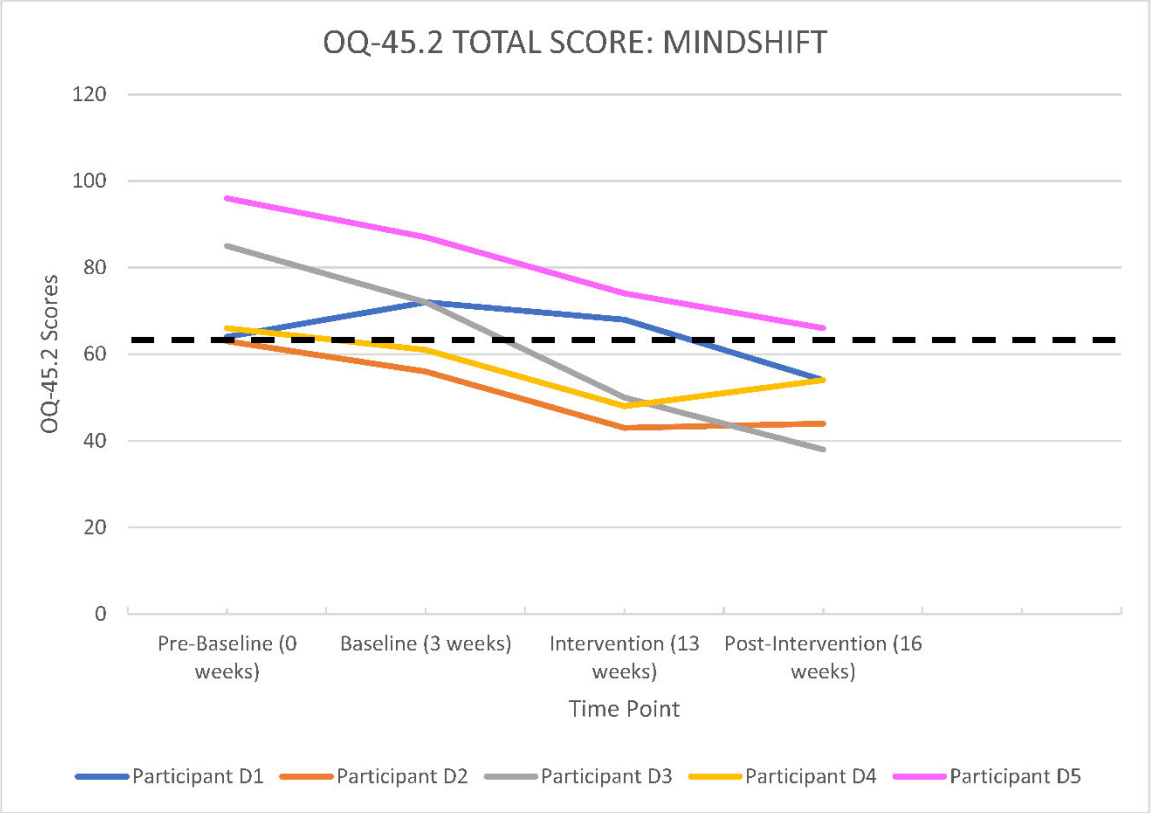
OQ-45.2 Total Scores for Participants D1 – D5 Using MindShift

Participant	Pre-baseline to post-baseline	Post-baseline to post-intervention	Post-intervention to post-3-week follow up	Overall (pre-baseline to post-3-week follow up)	Classification (pre-baseline to post-3-week follow up)	Improvement (pre-baseline to post-3-week follow up)
D1	(64-72) ^	72-68 ^	68-54 *+	64-54 +	Unchanged	15.63%
D2	63-56 +	56-43 ^	(43-44) ^	63-44 *+	Recovered	30.16%
D3	85-72 ^	72-50 *+	50-38 ^	85-38 *+	Recovered	55.29%
D4	66-61 +	61-48 ^	(48-54) ^	66-54 +	Unchanged	18.18%
D5	96-87 ^	87-74 ^	74-66 ^	96-66 *	Improved	31.25%

Note. Based on RCI = 14 and CSI = 63 for Total Score. ^ denotes not statistically reliable, not clinically significant; * denotes statistically reliable ($p < .05$); + denotes clinically significant; *+ denotes statistically reliable and clinically significant. Parentheses indicate change in a worsening direction. *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

Figure 10.25

OQ-45.2 Total Scores for Participants D1 – D5 Using MindShift



Note. Clinical cut-off of 63 from the OQ-45.2 technical manual (Lambert et al., 2004).

Destressify

Table 10.30 and Figure 10.26 show a summary of life functioning outcomes for participants using *Destressify*. As can be seen in Figure 10.26, at pre-baseline (Phase 1) all participants except for E1 recorded a clinically significant impairment in life functioning. At post-intervention (Phase 4), all participants were improved. The improvements shown by Participants E2, E4, and E6 were clinically significant, with Participants E1 and E6 being in the non-clinical range at post-intervention.

Table 10.30

OQ-45.2 Total Scores for Participants E1 – E6 Using Destressify

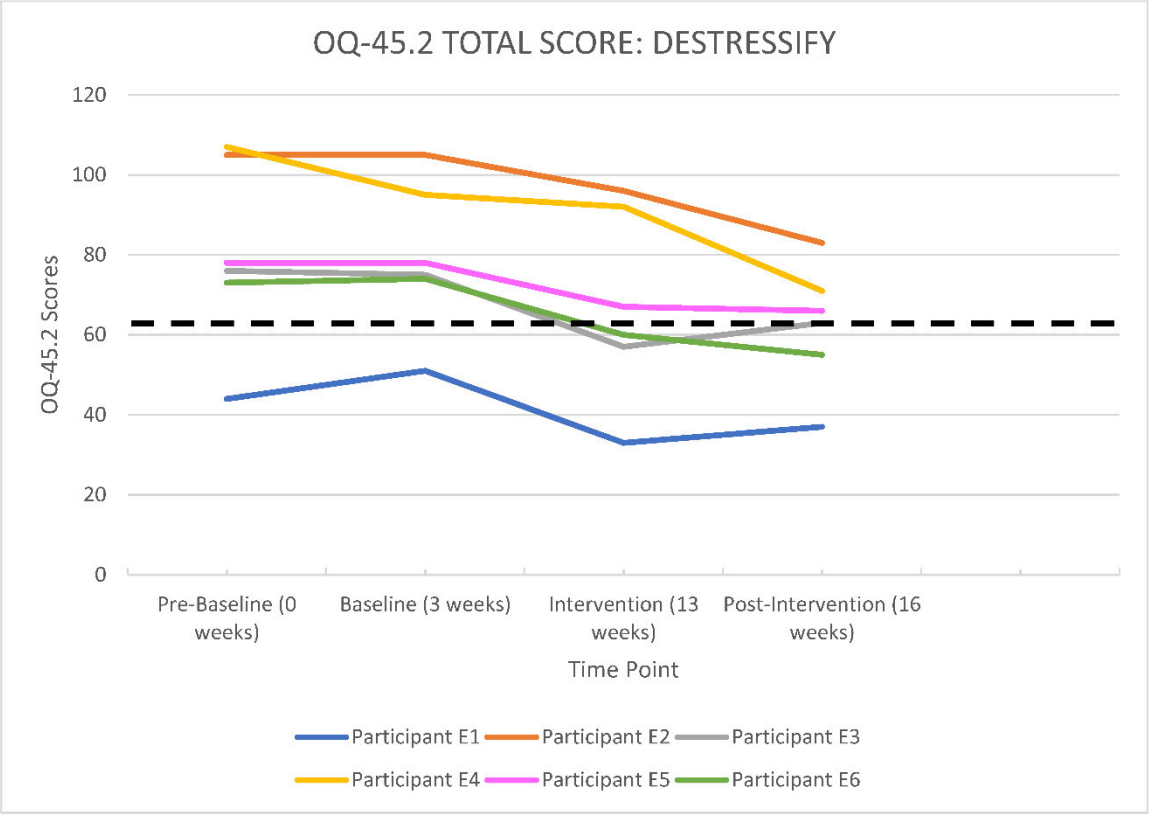
Participant	Pre-baseline to post-baseline	Post-baseline to post-intervention	Post-intervention to post-3-week follow up	Overall (pre-baseline to post-3-week follow up)	Classification (pre-baseline to post-3-week follow up)	Improvement (pre-baseline to post-3-week follow up)
E1	(44-51) ^	51-33 *	(33-37) ^	44-37 ^	Unchanged [a]	15.91%
E2	105-105 ^	105-96 ^	96-83 ^	105-83 *	Improved	20.95%
E3	76-75 ^	75-57 *+	(57-63) +	76-63 ^	Unchanged	17.11%
E4	107-95 ^	95-92 ^	92-71 *	107-71 *	Improved	33.65%
E5	78-78 ^	78-67 ^	67-66 ^	78-66 ^	Unchanged	15.39%
E6	(73-74) ^	74-60 *+	60-55 ^	73-55 *+	Recovered	24.66%

Note. Based on RCI = 14 and CSI = 63 for Total Score. ^ denotes not statistically reliable, not clinically significant; * denotes statistically reliable ($p < .05$); + denotes clinically significant; *+ denotes statistically reliable and clinically significant. Parentheses indicate change in a worsening direction. *Recovered* = clinically significant and statistically reliable; *Improved* = not clinically significant, but statistically reliable; *Unchanged* = not clinically significant or statistically reliable; *Deteriorated* = clinically significant and/or statistically reliable in a worsening direction.

[a] Non-clinical to begin with.

Figure 10.26

OQ-45.2 Total Scores for Participants E1 – E6 Using Destressify



Note. Clinical cut-off of 63 from the OQ-45.2 technical manual (Lambert et al., 2004).

Summary of Improvements

Participants were classified into different groups based on their overall improvements in daily distress, symptomatology, and life functioning, and according to the criteria in the *Note* in Table 10.31 below. The classifications were: Highly Effective (those participants who responded with the greatest improvements), Moderately Effective (those participants who responded with moderate improvements in some areas), Less Effective (those participants who responded with smaller overall improvements), and Did Not Finish. See Table A.1 in Appendix A for how participants were classified on all measures.

Table 10.31

Summary of Improvements for all Participants

Improvement summary classification	Participant	Total
Highly effective	A2, B1, B2, B3, B5, B7, C5, D3	8
Moderately effective	A1, A3, A5, B4, C1, C2, C4, C6, D2, E4, E6	11
Less effective	A4, B6, C3, D1, D4, D5, E1, E2, E3, E5	10
Did not finish	A6, A7, A8, C7, C8, D6, D7, D8, E7, E8	10

Note. Classification summary of each participant into an effectiveness group is based on author consensus of the following criteria relating to Table A.1 in Appendix A: Highly effective = 1 x "High" rating in three different categories, or "High" and "Mod" ratings across all categories; Moderately effective = 1 x "Mod" rating in three different categories, or at least 1 x "High" rating; and Less effective = 1 x "Less" rating in three different categories.

Participant Factors that Impacted the Results

There were differences between participants who either finished the study or dropped out, and those who gained greater or lesser benefit from the intervention. Four main characteristics stood out amongst the group of participants who failed to finish the study: (a) having a diagnosis of depression alone; (b) having a longer duration of mental illness; (c) being older; and (d) having increased motivation and beliefs prior to commencing the study that their app would provide an improvement to their mental health. All participants who finished the study showed some level of improvement in one or more areas examined by the self-report measures. However, the following participant factors were associated with greater treatment effectiveness.

Treatment effectiveness tended to be greater in younger participants compared to older participants ($t(29) = 3.24$; $p = .002$). The average age of the highly effective group was 26.63 (SD = 6.28), the moderately effective group was 34.46 (SD = 14.19), and the less effective group was 39.50 (SD = 11.25).

All participants in the highly effective group, and all but two participants (A3 and C6) in the moderately effective group, were receiving concurrent psychotherapy and/or psychotropic medication.

All participants in the highly effective group had their diagnosis for less than five years. Whereas eight of the 10 participants (80%) in the less effective group had their diagnosis for greater than six years. Furthermore, all 10 participants who failed to finish had their diagnosis for greater than six years.

All participants in the highly effective group had an anxiety disorder, with six out of the eight (75%) having co-morbid depression. Whereas five of the 10 participants (50%) in the less effective group had standalone depression, and eight of the 10 non-finishers (80%) had standalone depression.

All but one participant (A2) in the highly effective group (87.5%) had neutral or negative views about their motivation to do what the app suggests prior to starting the intervention (and prior

to knowing which app they had been randomised to). In the less effective group, four out of 10 participants (40%) had neutral or negative views, and in the non-finishers group, 3 out of 10 participants (30%) had neutral or negative views.

App Ratings

Overall, participants who finished the study rated their app using the uMARS as follows: *SuperBetter* 2.60 out of 5 stars (*SD* = 0.55) and 79.60 out of 130 (*SD* = 12.16); *Smiling Mind* 3.86 (*SD* = 1.46) and 100.00 (*SD* = 20.12); *MoodMission* 3.50 (*SD* = 0.55) and 98.17 (*SD* = 13.00); *MindShift* 2.60 (*SD* = 1.14) and 83.60 (*SD* = 18.20); and *Destressify* 3.00 (*SD* = 0.63) and 82.17 (*SD* = 15.83). See Table A.2 in Appendix A for further details about individual participant responses.

Discussion

This study has evaluated the effectiveness of five mental health apps: *SuperBetter*, *Smiling Mind*, *MoodMission*, *MindShift*, and *Destressify*. Participants who completed the study were diverse in terms of age, level of subjective distress, symptoms of anxiety and depression, and impairment in life functioning. Even so, all 29 participants who completed the study recorded at least a moderate improvement in some aspect of their pre-intervention presentation, indicating that using these mobile apps may aid in improving the mental health of people with symptoms of anxiety and/or depression, including in the context of a global pandemic such as COVID-19. However, in interpreting the effectiveness of these mental health apps, several potential limitations need to be considered (discussed later in this section), including that actual app usage by participants was unable to be recorded.

Answering the Research Questions

The present study sought to answer three research questions:

1. Can a range of mental health apps, employing diverse theoretical orientations, reduce subjective distress and clinically significant symptoms of anxiety and/or depression, and improve functioning in a sample of heterogeneous participants?

Given the results outlined above, it appears that the apps may be effective at improving multiple dimensions of mental health and wellbeing. These results enhance the previous evidence obtained from a range of methodologies, including RCT designs (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; and Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017).

2. Are there specific factors about the participants that impact on the results?

The apps have potential to offer a range of positive results over a broad cross-section of individuals. Although some observations can be made about participant characteristics that may have influenced treatment outcomes, these need to be read with the limitation in mind that data on actual app usage was not collected. Treatment effectiveness tended to be greater in younger participants compared to older participants. This may suggest that younger people get more benefit from mental health apps than older people. This may be reflective of the fact that younger people have grown up using mobile Internet-enabled devices and may therefore have greater affinity with and comfort using this technology, and who may simply rely more on their smartphone to perform a variety of tasks.

All participants in the highly effective group, and all but two participants (A3 and C6) in the moderately effective group, were receiving concurrent psychotherapy and/or psychotropic medication. This may potentially signal that those who are receiving concurrent psychotherapy and/or psychotropic medication may see greater benefit from using apps as an adjunct to this treatment compared to people who are using apps as a standalone treatment for their symptoms of anxiety and/or depression. Even though the best results were obtained by people receiving concurrent therapy and/or psychotropic medication, their stable baselines indicate that their recovery had plateaued, and improvement came after adding the app intervention. One possible explanation for this mechanism of action is that participants were instructed to use their app regularly (for at least 10 minutes per day, five days per week), which effectively means that they were doing regular “homework” to help

improve their mental health. Doing such regular homework is known to aid in treatment effectiveness of therapies such as CBT (Addis and Jacobson, 2000).

All 10 participants that dropped out of this study had long-term chronicity of their anxiety and/or depression diagnoses; two of these participants had had their diagnosis for 6 – 10 years, and the other eight had their diagnoses for over 11 years. This may suggest that the use of apps to treat symptoms of anxiety and/or depression are potentially more effective for individuals with a shorter chronicity of their diagnosis.

There is further evidence of this possibility when looking at the summary classifications of participants who completed the study. Of the ten participants classified as having less effective results overall, eight (80%) had a long-term diagnosis over six years. In contrast, all participants classified in the highly effective group had shorter chronicity of symptoms (less than five years). This, again, is evidence that perhaps apps for anxiety and/or depression are more effective for individuals who have experienced their symptoms for shorter periods.

Furthermore, there is a possibility of an interaction effect. That is, the combination of younger individuals with symptoms of a shorter duration, *and* who are engaged in concurrent psychotherapy and/or psychotropic medication, may produce the most effective conditions for successful outcomes from using one of the five apps in this study.

Another participant characteristic that was present in the highly effective group of participants was a neutral or negative view about their motivation to do what the app suggests prior to starting the intervention (and prior to knowing which app they had been randomised to). Only one participant in the highly effective group had positive motivation, whereas 60% and 70% respectively had positive motivation in the less effective and non-finishers group. This runs counter to previous research that has found motivation to be compliant with a therapy predicts successful outcomes in treatment (Addis

& Jacobson, 2000). Future research could delve deeper into this finding to determine if mental health apps offer an effective treatment option for those individuals with low motivation towards treatment.

However, the multiple single-case design approach revealed exceptions to other demographic factors that have previously been used to predict outcomes for in-person treatment e.g., attitudes towards mental health professionals, and the presence of another chronic medical condition, did not necessarily predict successful outcomes generally. Similarly, factors that may intuitively feel as though they should predict outcomes here, such as app star ratings, technology and smartphone abilities, or having a particular belief prior to treatment that technology has the potential to help anxiety and/or depression, did not necessarily predict outcomes in this study. What the findings do reveal is that the smartphone apps used in this study have the potential to help a wide cross-section of people manage their symptoms of anxiety and/or depression, even in the context of a major worldwide crisis such as the COVID-19 pandemic.

One interesting participant characteristic, also found in the pilot study, is that level of mental health literacy (Jorm, 2012) does not necessarily predict outcomes, except to say that the mental health apps used here were effective for some individuals who had negative attitudes towards mental health professionals. For instance, participants B1, B3, and B7, who all used the *Smiling Mind* app, reported that psychiatrists or psychologists were “harmful” for an individual’s mental health, yet all achieved excellent outcomes from using the app in areas of reducing daily levels of distress and symptoms of anxiety and/or depression, and improving general life functioning. Another finding that replicates the results of the pilot study is that ability with technology and smartphones did not seem to be a factor in influencing the results, as there was a cross-section amongst participants of “average” to “excellent” abilities, according to their own self-ratings.

Comparing each of the apps, it is apparent that there are more similarities in the results than differences. This is interesting because the design / aesthetic qualities and theoretical orientations of

the apps are different and did not appear to influence the results. That is, *SuperBetter* is based on CBT, positive psychology and neuroplasticity; *Smiling Mind* is based on mindfulness; *MoodMission* on CBT; *MindShift* on CBT; and *Destressify* on mindfulness. Overall, it seems that the biggest impact on successful treatment effectiveness may have been individual participant characteristics such as being of younger age, having shorter chronicity of mental illness, receiving concurrent psychotherapy and/or taking psychotropic medication, having anxiety with or without depression (rather than standalone depression), and not having high motivation to engage with a mental health app. However, this needs confirming in studies that are able to record data on actual app usage by participants.

3. What are the participants' experiences of using the apps?

Although it is difficult to directly compare the results of app ratings by participants of this study to ratings from studies with a different experimental design, it would appear that the finishing participants rated the apps at least moderately positively. Participants using the *Smiling Mind* app rated it the highest using the combined scores from the different sections on participants' ratings on the uMARS questionnaire (see Table A.2 in Appendix A for details), although all the apps can claim some measure of positive outcomes based on the uMARS data.

The following question emerged from the results from the preceding three research questions:

What factors emerge as benefits, facilitators or barriers to using single-case research, conducted by practicing clinicians, to develop the evidence base for mental health apps?

The findings of this study indicate that a single-case research methodology is able to provide nuanced information about the effectiveness of an app that is not available in an RCT. For instance, not only was it established that the apps are effective in facilitating improvements in a number of domains (subjective distress, symptoms of psychopathology, and life functioning) across a wide age range (18 – 57 years of age), this also occurred in a variety of contexts. These include: with individuals receiving concurrent psychotherapy and/or taking antidepressant medication; having co-

morbid anxiety and depression; and having relatively low levels of motivation. Therefore, not only can a single-case research design adequately and comprehensively assess the effectiveness of a mental health app in a real-world setting, but given the automated nature of the methodology, it could easily be scaled up to accommodate more participants with the same level of detailed information available. The assertive follow-up approach was effective at reducing participant dropout rates, but this should be noted by clinician researchers as an extra demand on their time if they wish to employ this methodology in practice-based research.

Comparison with Results from the Pilot Study Using *SuperBetter*

SuperBetter demonstrated effectiveness in this study and the pilot study. The pilot study had a greater percentage of participants drop-out with assertive follow-up likely to have reduced attrition in the current study. The participants who dropped out of both studies and who made the least effective gains were more likely to have had their mental health diagnoses for longer. However, perhaps the most noteworthy comment to make is that *SuperBetter* was effective during a non-COVID-19 period (in the pilot study) and during the COVID-19 period (in the present study).

Comparison with Results from Other Previous Research

The previous research on *SuperBetter* (Roepke et al., 2015) and *MindShift* (Paul and Fleming, 2019) found evidence for improvements in both symptoms of anxiety and depression. The previous research on *Smiling Mind* (Flett et al., 2019) and *Destressify* (Lee and Jung, 2018) found improvements in anxiety, but not depression, which is in line with the findings of the present study. The previous research on *MoodMission* (Bakker et al., 2018b) found improvements in symptoms of depression but not anxiety, which runs counter to the findings of the present study. None of the previous research on these apps provides enough information about the participant characteristics to compare the notable results of the present research, including: if there were any age differences between the participants who achieved the greatest benefits; if there were any differences in

chronicity of mental illness; if there were any differences between the app as a standalone treatment compared to app usage with concurrent pharmacotherapy and /or psychological therapy; and if there were any differences in motivation levels prior to treatment. These are all areas for future research to consider.

COVID-19 Summary

This study provides an interesting snapshot of a digitised and automated method of psychological intervention during the worldwide COVID-19 pandemic. In many countries, mental health services have seen an increase in demand due to the psychological burden of COVID-19, and this has highlighted a lack of available in-person services (Marshall et al., 2020f). Governments and individuals have therefore been turning to digital or telephone options to fill the gap. While some of these methods have evidence of their efficacy, the use of smartphone mental health apps to treat trauma has limited evidence. The present study has shown that five specific apps with an assortment of evidence-based theoretical frameworks, activities, and approaches have been effective in managing symptoms of anxiety and/or depression during the height of the COVID-19 emergency for a wide variety of participants. It provides evidence that these smartphone apps can be used for people with mild to moderate symptoms so that other limited in-person resources may be allocated to individuals who may have more acute and/or severe presentations, or who require sustained in-person treatment during an emergency such as COVID-19.

Strengths of the Present Study

This study was able to closely examine the outcomes of 29 diverse participants in response to using five different apps to improve their mental health. The multiple baseline across-individuals design allowed detailed information on participants to be considered when examining how effective an app had been at improving their mental health and wellbeing. In this case, the apps were effective

on a number of variables across a range of participants, providing evidence in support of existing research findings.

A further strength of the study was the relatively low attrition rate (25%). It is believed that this was primarily due to an assertive follow-up approach that was used in response to the findings of the pilot study that contained a higher attrition rate (60%). It also compares favourably to the attrition rates of other studies on the efficacy and effectiveness of mental health apps (e.g., Roepke et al., 2015).

Clear instructions on how to use the apps were given to participants. That is, they were told to use their app for at least ten minutes per day, five days per week, for ten weeks. Most mental health apps do not offer such precise instructions, which means they are not comparable to other evidence-based treatments, such as having weekly 50-minute psychotherapy sessions, or taking a prescribed dose of psychotropic medication. Taking a "dosage" approach makes the overall use of mental health apps clearer. Researchers and mental health app developers should be examining this issue closely as it may have an impact on a person's overall fidelity to a mental health app.

This study has also been successful in providing research evidence on the effectiveness of five apps, all of which have other published research, but lack replicated independent evidence. This study has provided much needed independent data that has been widely called for in the literature.

Limitations

The findings of this study should be read with the following limitations in mind.

Firstly, the small number of participants means that generalisations need to be made in the context of the demographics of these participants. The evidence presented in this paper does not mean that the apps will be effective for all consumers. However, it does suggest that the apps can be effective for people aged between 18 and 57 with a range of presentations, but particularly if they

have a shorter duration of mental illness, and even if they have some level of doubt about an app's ability to improve mental health.

Secondly, a number of participants did not exhibit statistically stable baseline periods prior to intervention commencement, which is a usual prerequisite for a multiple baseline across-individuals design (Barlow et al., 2009). Specifically, participants A4, B1, B6, C5, and D1 all had statistically significant changes in their SUDS ratings during their baseline phase. This instability in subjective distress prior to the implementation of the intervention means that caution is required in interpreting the results achieved in response to the intervention, especially if the participant's baseline readings were trending downwards, as in the case of A4, B6, and C5. However, three of the five participants who did not have statistically stable baseline periods (A4, B6, and D1) were classified in the Less Effective group at the end of the post-intervention phase. In all five apps, there were at least four other participants who did show a statistically stable baseline period, and therefore provided an offset for comparison.

Thirdly, the researchers were unable to use data directly from the apps themselves, and therefore were unable to know exactly how much time participants spent using their apps which means that comments on treatment fidelity cannot be made. Although the technology exists to do this, it would require the resources of a study with greater funding, and would have required the involvement and cooperation of app developers. This means that judgments about the extent that the apps were responsible for any change in symptomatology or life functioning are limited. Furthermore, if access to app data was available, those participants that dropped out of the study would also have been able to contribute data that may have impacted on these findings. Having access to app data is a consideration for future research in this area.

Fourthly, the use of assertive follow-up, while solving one problem, created another potential difficulty. That is, assertive follow-up may have acted as a form of adjunctive therapist support.

While it resulted in an increase in the amount of usable data at the study's completion, it also places a limitation on the extent that these apps can be said to offer truly automated support.

Fifthly, by not incorporating SUDS data from participants who dropped out of the study, there is a risk that the sample is biased, and therefore that the time-series analysis results are skewed. Interpreting the time-series analysis results must be made with caution in light of this.

Sixthly, participants who dropped out of the study were not followed up in relation to reasons why they did so. If such participants were given the opportunity to describe why they dropped out, it may have shed further light on particular characteristics of the apps themselves and how these may detrimentally interact with participants who have specific characteristics.

Seventhly, while the applicability of the results may stretch across a wide variety of participant characteristics, the level of applicability may be limited when it comes to specific minority groups, such as First Nations peoples and culturally and linguistically diverse (CALD) communities. Specific research into such minority groups is needed to confirm the level of this acceptability towards those groups.

Eighthly, the survey instrument used to obtain demographic information from participants did not seek details relating to ethnicity, language, country of birth, and Indigenous status, thereby not allowing any investigation into participant characteristics related to First Nations peoples or CALD communities. Furthermore, the survey instrument contained a label with the heading "other mental health professional" which may have alienated professionals who identify as mental health nurses, peer workers, social workers, occupational therapists, and Indigenous health workers. Such omissions may have skewed the research methodology, the interpretation of data, and the reporting of results.

Ninthly, participants were only sought from Australia. By not seeking and including participants from other countries, the utility of the findings may be narrower as a consequence. However, the study was deliberately seated within the Australian context to have applicability to

funded mental health services such as Medicare, and to align with national strategies, priorities and frameworks.

Finally, a validated brief symptom measure such as the Patient Health Questionnaire (PHQ-2 or PHQ-9) could have been used more regularly to monitor changes in symptomatology over time in the same fashion as the SUDS ratings. This would have provided additional data to the larger measures used less regularly and provided a useful comparison with the daily SUDS ratings.

Future Directions

The present study has shown that a single-case research methodology is useful for examining the effectiveness of mental health apps, and could demonstrate efficacy if sufficient data was gained from participants with homogenous presentations. As a result, this study provides a number of possible avenues of future research through the comprehensive information obtained about participant characteristics. These include age of participants, level of motivation to engage with the app, chronicity of illness, and standalone or co-morbid diagnoses of anxiety and depression. There may also be other moderating factors impacting on the effectiveness of these mental health apps. Participant characteristics of future research also needs to embody specific populations within the community, such as First Nations peoples and CALD communities who may traditionally experience unique barriers to accessing in-person mental health treatment. Future research also needs to establish how the principles of human computer interaction and design interact with participant characteristics to contribute to treatment outcomes. A further key area for future research should focus on the amount of concurrent intervention is required to support positive outcomes with apps. The present study has produced results by implementing a vigorous methodology to obtain data that can inform the future research of the five apps used here, and the future research of other mental health apps whose developers would like to produce evidence of efficacy in a timely manner.

Conclusion

Whilst all apps demonstrated some level of effectiveness, it is the personal characteristics of the user that may have the greatest impact on treatment success. All of the participants whose outcomes revealed their app to be “highly effective” were either receiving concurrent psychotherapy and/or taking psychotropic medication. This finding suggests that caution is required when directing someone towards one of these apps as a standalone treatment. It is not suggested that all available mental health apps have the potential to produce similar results. The apps used in this study all contained evidence-based frameworks and were produced with the input of mental health professionals and/or mental health organisations (which should be the minimum standard required of a mental health app), and all had prior evidence of their efficacy published in peer-reviewed journals, something that the vast majority of available mental health apps do not have. The methodology of the present study demonstrates that a single-case research design is able to provide useful evidence to complement the evidence of efficacy obtained in larger RCTs. The evidence base for the apps examined in this study is further enhanced with the results obtained here, which have shown them to be effective to varying degrees across several domains, including subjective distress, symptoms of psychopathology, and life functioning. It is possible to scale up this methodology so that it can be used for many other apps, with the ultimate aim to give consumers greater choice of evidence-based mental health apps when searching the app stores, and ultimately more options in treating anxiety and depression.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Chapter 11: Discussion

There has been an increasing trend towards utilising technology in healthcare. The growth in global smartphone ownership has seen an equivalent growth in the development and use of apps to manage health-related issues. Approximately 60% of adults have downloaded a health app (Carras et al., 2014), and more than this have indicated they would consider using an app to manage a health-related condition (Makovsky, 2015). The area of mental health apps is no different. This has been highlighted in the COVID-19 pandemic where record numbers of mental health apps have been downloaded (Heilwell, 2020; Statista, 2020) in response to difficulties accessing face-to-face mental health services (Liu et al., 2020).

When an individual sees a health professional or is prescribed medication, there is an assumption that the health professional is qualified to practice and that the medication has undergone a scientific evaluation process. It remains unknown if individuals assume the same level of accreditation about the mental health apps they download. It is also unknown whether health professionals make similar assumptions. Whether they do or do not assume this, there remains very little research evidence about the efficacy or effectiveness of mental health apps (Orman & O'Dea, 2018; Marshall et al., 2020b). There may be a variety of reasons for this, but two salient points of contention for app developers in putting their apps through a research process are the time and financial constraints of conducting RCTs.

Aims of This Thesis

This thesis had two primary aims and attempted to address several research questions. The first aim was to ascertain the current level of research into the efficacy and effectiveness of mental health apps, with a specific focus on apps that addressed symptoms of anxiety and/or depression. In doing so, the research also sought to identify the barriers to research. Throughout several chapters,

this thesis explored seven specific research questions related to the first aim as follows: What proportion of apps for anxiety and/or depression claim to have research evidence for their effectiveness; what proportion of that research is independent and/or replication research (that is, what proportion of research has not involved a member of the development team of the app being studied, or anyone else who is otherwise connected with the app being studied and who potentially stands to gain financially or professionally from the success of that app); what proportion of apps for anxiety and/or depression have involved a mental health expert in their development; what proportion of apps for anxiety and/or depression have used an evidence-based framework in their development; what are the proportions of specific frameworks that have been used; what proportion of apps for anxiety and/or depression have been developed in affiliation with an academic institution, medical facility, or other government-funded body; and what proportion of apps for anxiety and/or depression are free to download?

The second aim of this thesis was to examine the feasibility of a single-case research design, specifically the *multiple baseline across-individuals* design (Barlow et al., 2009) which incorporates a modern, automated approach, to enhance the existing evidence base and offer a solution to increasing the speed and comprehensiveness of future research. Single-case designs have previously been suggested as a method for increasing the evidence-base of mental health apps (Clough & Casey, 2015a; Mehrotra & Tripathi, 2018), yet none could be located in the literature – see Chapter 4 of this thesis. Throughout several chapters, this thesis explored three specific research questions related to the second aim as follows: Can a range of mental health apps, employing diverse theoretical orientations, reduce clinically significant symptoms of anxiety and/or depression in a sample of heterogeneous participants; are there specific factors about the participants or the apps that impact on the results; and what factors emerge as benefits, facilitators or barriers to using single-case research, conducted by practicing clinicians, to develop the evidence base for mental health apps?

Summary of Outcomes

This thesis was able to illuminate further information on both of the main aims and all of the research questions.

Study 1 (Chapter 4) was a systematic review of the literature on studies conducted by independent researchers on mental health apps. Only four articles examining five apps could be located that met criteria for inclusion in the review. All articles found statistically significant results for improvements in anxiety symptoms, but only one study contained significant results for reductions in depression. Given the low number of studies, it was decided that a meta-analysis would not produce meaningful data and was therefore not undertaken. There was a high level of heterogeneity between these studies in many areas, including measures of symptomatology used, quality of research, and numbers and demographics of participants.

This was the first systematic review that focused exclusively on independent research. The consideration of independence is crucial to this area of research because so much of the previous research has been conducted by researchers who have had a relationship to the app being tested. Without studies by independent researchers, the field is leaving itself open to accusations of bias. This situation is not without precedent in mental health research. Past research on the efficacy of psychotropic medication was carried out by the pharmaceutical companies that developed the drugs and has been found to have larger effect sizes than more recent research by independent researchers on the same medications (Cipriani et al., 2018). In the area of mental health apps, previous reviews including all published research have found statistically significant improvements in symptoms of anxiety and depression (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017).

Study 2 (Chapter 5) had the aim of finding the proportion of mental health apps that had research evidence which could be located in the app stores. This was seen as complementary to Study 1 which focused on locating peer-reviewed publications as part of a literature review. Study 2 found that only 3.41% of apps that offered a therapeutic treatment for anxiety and/or depression had

research evidence, with only about 1.02% representing independent research. Of the 293 apps shortlisted as offering a therapeutic treatment for anxiety and/or depression, 55.3% mentioned an evidence-based framework in their app store descriptions. When broken down, this equated to 30.0% claiming to use CBT, 15.7% claimed to use mindfulness, 9.2% claimed to use positive psychology, 3.4% claimed to use DBT, 1.7% claimed to use ACT, and 6.8% claimed to use other techniques. Only 6.2% of the apps that claimed to use an evidence-based framework had published evidence for their efficacy. Other results indicated that 30.38% of shortlisted apps had expert development input; 20.48% had an affiliation with a government body, academic institution, or medical facility; and 74.06% were free to download.

Specifically, in relation to the first aim, the systematic reviews that were conducted of the literature and the app stores revealed a distinct lack of research on the efficacy and effectiveness of mental health apps, which has been noted in other published reviews (e.g. Alyami et al, 2017; Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017; Wang et al., 2018). Also, very little of this research is independent, and none involves a single-case design or replication study.

Study 3 (Chapter 7) was a pilot study conducted to ensure that the proposed single-case methodology was practical and feasible in answering the remaining research questions for the second aim of this thesis. The pilot study examined the effect of the app, *SuperBetter*, on four participants' symptoms of anxiety and depression, subjective daily distress, and life functioning. Statistically significant reductions in subjective distress were found for three out of four participants and these were maintained post-intervention. All participants showed beneficial changes on measures of symptoms and life functioning. Some demographic variables also showed promise in explaining successful outcomes - or presented as variables that, at a minimum, could be used as ideas for future research. These included: receiving concurrent psychotherapy; taking antidepressant medication; having co-morbid anxiety and depression; and, levels of motivation prior to beginning the

intervention. This last variable provided interesting results because no participant answered "Strongly agree" to the statement "I am motivated to do what the mobile app suggests". This implies a level of scepticism on the part of all participants, yet each received benefit from using the app. Participants rated the app favourably overall, indicating that they would recommend it to others.

The findings suggested that *SuperBetter* is effective at improving ratings of distress, anxiety, depression, and life functioning across a diverse range of presentations. The single-case research design seemed particularly adept for collecting and interpreting data, and suggested that this approach has high utility for research in this field. There was a high attrition rate (60%) in the pilot study, therefore it was decided that an assertive follow-up approach would be applied in the main intervention study (Study 4) as a way of mitigating this response.

Study 4 (Chapter 10) used the same single-case methodology with the addition of an assertive follow-up approach. As a result of this approach, the attrition rate in Study 4 was 25.6%, substantially lower than in the pilot study. Seven participants were followed up in this fashion during the study, and all seven resumed their participation to complete the study. It is believed that all seven would have been lost to the study were they not followed up. This suggests that an assertive follow-up approach was effective at keeping participants in the study, and it provides an example for future research projects with similar designs.

Five apps were studied in total: *SuperBetter*, *Smiling Mind*, *MoodMission*, *MindShift*, and *Destressify*. Gains of clinical and statistical significance were made across all measures and all apps, with individual demographic factors providing some guidance on which of the 29 finishing participants had greater levels of success. Overall, more favourable outcomes were achieved by younger participants, those concurrently undertaking psychotherapy and/or psychotropic medication, those with anxiety or anxiety and comorbid depression (as opposed to depression on its own), and those with a shorter history of mental illness. In keeping with findings from the pilot study, levels of motivation prior to commencement of the intervention did not predict successful outcomes, with

several examples of individuals with low motivation and expectation finishing with some of the highest gains. Using the uMARS questionnaire, the overall mean star rating out of five for each of the apps was: *Smiling Mind* 3.86; *MoodMission* 3.50; *Destressify* 3.00; *SuperBetter* 2.60; and *MindShift* 2.60. However, there were examples of participants who gave lower subjective ratings of their app, but had high level outcomes. This is important because the main method used by consumers to select a mental health app is its star rating (Huang & Bashir, 2017), assuming that the higher the rating the more effective the app has been for individuals. This research suggests that often this may not be the case, and thus calls into question the appropriateness of this method of selection, and highlighting the need for a more effective method of categorising mental health apps, and health apps generally, in the app stores.

In relation to the second aim of this thesis, Study 3 and Study 4 were able to demonstrate that a single-case research design was able to successfully test the effectiveness of mental health apps, firstly by conducting a pilot study on a single app, and then by scaling up the methodology to test five mental health apps simultaneously. These apps demonstrated effectiveness in a number of areas (daily subjective distress, symptoms of anxiety and depression, and life functioning) for a range of presentations and demographically diverse participants, all within the context of a global pandemic, COVID-19.

Impact of COVID-19

The first signs of COVID-19 occurred towards the end of 2019 in Wuhan Provence, China. Over the course of the next few months, the virus spread around the world, and had significant negative health consequences for many. It became clear that mental health was one such negative impact across countries such as: Australia (Koh, 2020), China (Feng, 2020), India (Mukherjee, 2020), New Zealand (Mindfood, 2020), U.K. (Chowdhury, 2020), U.S. (Heilweil, 2020), and others.

The timing of the pandemic in relation to Study 4 is both serendipitous and intriguing. The baseline period for all participants commenced on 30 January 2020, and all participants were using

their assigned app by the 28 February. In Australia, where the study was completed, the Federal Government made several key announcements (including lockdown orders and making available additional government services and rebates for people who became unemployed) between the 12 and 23 March (Klapdor, 2020). This 12-day period saw an increase in stress across communities, including panic buying at grocery stores (Wright, 2020).

In relation to Study 4, all participants had been using their app for at least two weeks before these Australian Government announcements were made. The methodology was able to detect reliable spikes in SUDS between the crucial period of the 12 to 23 March, and in the weeks and months afterwards. Therefore, the study has been able to provide data on how well these five mental health apps were able to assist people to manage symptoms of anxiety and/or depression during what has arguably been the most significant global health event in a century. More broadly, the results of this study provide quality evidence of the effectiveness of these apps to help manage anxiety and/or depression during a period of massive global upheaval.

Implications for the Evidence Base and Clinical Use

There is currently a great deal of interest in research into the efficacy and effectiveness of mental health apps. This is evidenced by the high number of views that a publication from this thesis has had. Study 2A was published online in November 2019 in *Frontiers in Psychiatry* and has had over 22,000 views at time of writing (https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00831/full), placing it in the top 3% of all articles ever published by Frontiers. These metrics indicate the level of awareness by mental health researchers, clinicians, government departments, and organisations that smartphone mental health apps are being developed and downloaded in increasingly larger volumes, and there is a thirst for knowledge about them.

This thesis adds to the existing research on apps for reducing symptoms of anxiety and/or depression in a number of ways. It provides independent research evidence of effectiveness for five

mental health apps, *SuperBetter*, *Smiling Mind*, *MoodMission*, *MindShift*, and *Destressify*. It also demonstrates that these apps can be effective in times of global upheaval, such as the COVID-19 pandemic. It identifies specific demographic factors about participants that may be key to understanding the effectiveness of these apps; a finding that has not been published in any peer-reviewed forum that could be located. These demographic factors include whether a participant is receiving psychotherapy and/or taking psychotropic medication, chronicity of diagnosis, age, gender, attitudes towards technology, and motivation to engage with an app. It suggests a fresh approach towards a methodology that may expand the evidence base for mental health apps in a timely fashion, at a reduced cost, and in a manner that may include the involvement of practicing clinicians.

The use of a *multiple baseline across-individuals design* proved to be a thorough method of examining the effectiveness of a mental health app allowing for the incorporation of a larger spectrum of participant characteristics than is usually achievable in an RCT. This research has also demonstrated that using an assertive follow-up procedure to re-engage participants, who may otherwise have dropped out of the study, meant that more data was obtained with a lower attrition rate than is usual in previous research on mental health apps and other automated / digitised mental health interventions (Andrews et al., 2010; Richardson et al., 2010). However, this is a factor that practicing clinicians who are interested in taking part in this kind of practice-based research need to be aware of, as it will increase the demands on their already limited time.

Overall, this thesis provides a blueprint on how research into mental health apps can be accessed, harvested, collated and shared by researchers and clinicians around the world. The idea to develop an ever-increasing database of mental health app knowledge and research would make such information more easily accessible than any current method. Finally, this research used a unique protocol to guide an app store search that could potentially be developed into a more formal document and published as a guide for other researchers to use in the absence of other such guides.

One of the most meaningful findings from this research for practicing mental health clinicians is that individuals who were undertaking concurrent psychotherapy and/or psychotropic medication were more likely to experience positive outcomes than individuals who were using the app as a standalone treatment. This suggests that there is an opportunity for clinicians to include the recommendation and use of mental health apps by their patients and clients as an item in their treatment "toolkit". It should be noted, though, that the vast majority of mental health apps have been developed for individuals to use in isolation and therefore provide little opportunity for clinicians to integrate these into face-to-face therapy (Lan et al., 2018). Although there is evidence that consumers have more favourable attitudes towards e-mental health resources than clinicians (Gun et al., 2011; Waller & Gilbody, 2009), there appears to be scope and benefits for clinicians to become more educated about digital tools, such as mental health apps, given their potential as a positive adjunct to therapy.

However, while some websites have been developed to offer "expert reviews" of mental health apps, such as *PsyberGuide* (https://onemindpsyberguide.org/), these are usually aimed at the consumer. Clinicians who are interested in discovering what scientific evidence exists for the efficacy of a mental health app, do not have many options, but this thesis has identified a possible remedy.

The methodology employed in Study 3 and Study 4 could be used by practicing clinicians to generate a database of evidence. By having their patients and clients use an appropriate mental health app in between sessions, and reporting the outcomes in the same fashion as outlined in this thesis, there could be a clinician-led expansion of the knowledge base of evidence. With results on specific apps uploaded to a centralised database, evidence could be accessed by other clinicians (and consumers, researchers, mental health organisations, and government departments) seeking information. The resulting database would be an ever-increasing hub of knowledge about which apps work for who. See Chapter 3 for more information.

In summary, although there are many avenues in this area of research on mental health apps still to be explored, this thesis improves the level of understanding about how and why mental health apps work most effectively by identifying specific demographic variables that may be key predictors. It also identifies a role for practicing clinicians working within a single-case research design model to accelerate the evidence base. While the work in this research has progressed the existing knowledge base, it has also supplied hypotheses for future research. Generalisations about the findings contained in this thesis should be read with the following limitations in mind.

Summary of Clinical Implications

For clinicians seeking to integrate the use of apps into their therapy with clients and patients, this research suggests that adding an app as an adjunctive treatment may be quite effective. For clinicians wishing to engage in practice-based research, this research has demonstrated a single-case design approach that may be incorporated into a clinical practice setting. By doing so, clinicians could contribute to an expanding knowledge base of the efficacy and effectiveness of mental health apps. For clinicians seeking information about which mental health apps are effective, access to this knowledge base could be invaluable. Finally, this research has provided evidence that certain mental health apps may be effective at alleviating mild to moderate symptoms of anxiety and depression when access to other mental health services may be limited, especially during times of global unrest such as the COVID-19 pandemic.

Limitations of This Research

Limitations of the Literature and App Store Reviews

There are several limitations to consider in the literature and app store reviews. Firstly, in the literature review, only articles related to apps that claimed to offer a therapeutic treatment for anxiety and/or depression were examined. Articles that dealt with apps which did not specifically measure outcomes related to this were not included. It is, however, important to acknowledge this because

there are apps that have deliberately attempted to market themselves without referring to such potentially stigmatising terms as “anxiety” and “depression”, yet may still benefit people experiencing these conditions (Bakker et al., 2016). Further, the literature review only searched for peer-reviewed articles in English, and this may have ruled out other relevant but unpublished studies, thus creating potential publication bias, and may have missed relevant articles in a language other than English. Also, there are other databases that could have been selected in place of the ones that were used, but time and resources meant that no further databases were searched beyond the five mentioned. The literature review did not consider apps that acted as an adjunct to other types of treatment, and this is important because apps potentially offer consumers a more enhanced face-to-face therapy experience.

Secondly, in the app store reviews, it is important to recognise that there are challenges to conducting a search of the Apple App Store and Google Play, with differences in the way each delivers search results. This is not so much a limitation of the research itself, but of the search capabilities of the app stores. The algorithms used in each case are unknown and remain the product of corporate intellectual property. Search results can differ on different days, and it is impossible to explain how one app can appear earlier in a search compared to another similar app. The app store searches were also limited by their minimal search options. For example, there is no filter by date, or developer, or other options as are available when searching literature using a standard database. However, analysing the app stores remains an important exercise because this is how most people currently find mental health apps to download. A further limitation of the app store searches is that some of the terms used in our searches may not necessarily reflect words that consumers may use. For example, consumers may be more likely to use terms like “stress”, “depressed mood”, or “worry”, instead of terms like “dialectical behavior therapy” or “interpersonal therapy”. The app store reviews are also limited to results based on information contained in the app store description. That is, none of the apps were downloaded to ensure their store description actually matched the app’s

content. Furthermore, if no research is highlighted in the app store description, this does not necessarily mean that research does not exist. This is similarly true for the issues of level of expert input and association with a relevant organisation. To confirm exact numbers of apps with these elements, every app listed in each app store search would also have to be put through a literature search.

Limitations of the Pilot and Main Intervention Studies

There are also some limitations of the pilot and main intervention studies that should be considered when interpreting the findings. Firstly, generalisations need to be measured given the small number of participants in the pilot and main intervention studies. A great deal of information was gained on each participant, so it is advisable to consider the demographic profiles of each when drawing conclusions. The evidence presented in the main intervention study does not mean that all five apps will be effective for all consumers. However, it does suggest that the apps can be effective for a wide cross-section of people aged between 18 and 57 with a diverse range of demographic profiles.

Secondly, a number of participants in the main intervention study did not exhibit statistically stable baseline periods prior to intervention commencement, which is a usual prerequisite for a multiple baseline across-individuals design (Barlow et al., 2009). Specifically, participants A4, B1, B6, C5, and D1 all had statistically significant changes in their SUDS ratings during their baseline phase. This instability in subjective distress prior to the implementation of the intervention means that the reader must use caution when interpreting the results achieved as a result of the intervention, especially if the participant's baseline readings were trending downwards, as in the case of A4, B6, and C5. However, three of these participants (A4, B6, and D1) were classified in the Less Effective group at the end of the post-intervention phase, and in all cases, there were at least four other participants who did show a statistically stable baseline period in each of the apps selected, and therefore provided an offset for comparison.

Strengths of This Research

Strengths of the Literature and App Store Reviews

The literature search used a thorough, systematic methodology to filter studies on apps that offered a therapeutic treatment of symptoms of anxiety and/or depression. It was the first time that independent and replication studies were deliberately sought in this research area, and uncovered a worrying trend in a lack of this type of research.

The app store reviews provided information that is directly linked to the general public, as opposed to reflecting the results obtained from experimental groups. In doing so, the app store reviews highlighted significant shortcomings in the way mental health apps are represented (as described in their app store descriptions) and in what order they appear in search results (and whether they appear at all). The app store reviews also contained a proposed search protocol adapted from the PRISMA (Moher et al., 2009) and AMSTAR (Shea et al., 2017) models. No such existing, published protocol for app store reviews could be located.

Strengths of the Pilot and Main Intervention Studies

By using a multiple baseline across-individuals design, the pilot and main intervention studies were able to offer a detailed examination of the mental health outcomes of 33 participants with diverse demographic profiles. This detailed approach was able to occur in the same or shorter period of time as would be the case with an RCT design, yet far more information could be gleaned from this data. This large amount of nuanced data has the added benefit of providing many possible hypotheses for future research. This study has also been successful in providing research evidence on the effectiveness of five apps, all of which have other published research, but lack replicated independent evidence. This study has provided this much needed data that has been widely called for in the literature and adds to the growing amount of evidence for digital mental health interventions generally including those that are Internet-based. Finally, the low attrition rate of the main

intervention study (25.60%) is in contrast to other similar studies, including the pilot study (60%), and the Roepke et al. (2015) study on the *SuperBetter* app, which had an attrition rate of 81%. This is attributed to the assertive follow-up procedure used in the main intervention study, which involved contacting participants after a period of non-response for the SUDS ratings and encouraging them to re-engage with the study. This approach may not be feasible in larger RCTs.

Future Research

The pilot and main intervention studies in this research have shown that a single-case methodology is capable of, and useful for, examining the effectiveness of mental health apps, and potentially the efficacy if sufficient data is gained from participants with homogenous presentations. As a result, this research provides a number of possible avenues of future research through the comprehensive information obtained about participants. These characteristics include: age of participants (younger participants generally made greater gains), whether they are also receiving psychotherapy and/or on antidepressant medication (apps tended to be more effective when they were adjunctive), level of motivation to engage with the app (apps could be effective even when motivation and optimism about treatment success were low), length of time with mental illness (greater gains were generally seen in people with diagnoses of less than five years compared to those with diagnoses of greater than 11 years), and standalone or co-morbid diagnoses of anxiety and depression (greater gains were generally seen in people with anxiety, or comorbid anxiety and depression, compared to those with standalone depression). The pilot and main intervention studies have produced results by implementing a vigorous single-case methodology to obtain data that can inform the future research of the five apps used here, and the future research of other mental health apps whose developers would like to produce evidence of efficacy in a timely manner. In summary, the specific variables identified by this research that still need to be adequately explored in future research include: Gender; age; chronicity of diagnosis; motivation to engage with apps; attitudes towards technology; attitudes towards mental health providers and services; the impact of being an

adjunct to in-person psychotherapy and/or taking psychotropic medication versus using an app as a standalone treatment; type of evidence-based framework; and individual interventions within the same evidence-based framework.

While the single-case methodology used in this research seemed applicable to studying the efficacy and effectiveness of mental health apps, it may be that other approaches have similar applicability beyond the traditional RCT design. For example, a stepped-wedge trial may be applicable to studying this type of digital mental health treatment. It is hoped that future research consider alternatives to traditional RCT approaches because these may not be the most appropriate method for assessing mental health apps (Clough and Casey, 2015a).

The assertive follow-up procedure used in the main intervention study showed a substantial improvement in participant attrition rate. Using the single-case research design allowed this to occur in a way that did not drain the limited resources of this study, and meant that an additional seven participants remained in the study when they would have otherwise dropped out if they were not followed up. It is recommended that similar future research of a comparable nature and design to this study consider this approach to improve attrition rates, and to capture more data overall.

Summary of Recommendations

This research has produced findings that inform the following recommendations. It is recommended that clinicians become more aware of the issues related to mental health apps and the general lack of evidence that exists for their efficacy. For those apps that do have evidence, it is recommended that they be given consideration by clinicians to integrate these into their clinical practice, as adjunctive use of such apps may improve treatment outcomes. It is recommended that governments consider greater oversight of mental health apps and provide development and certification frameworks for their effective implementation into clinical practice. It is recommended that mental health apps be developed with the involvement of mental health experts and use evidence-

based theoretical frameworks to underpin their functionality. Finally, it is recommended that researchers consider using alternative methodologies such as single-case designs in addition to RCTs in an effort to expand the evidence base for mental health apps.

This research is applicable to and aligns with many Australian Federal and State Government priorities, including the National Digital Health Strategy by the Australian Digital Health Agency (https://www.digitalhealth.gov.au/sites/default/files/2020-11/Australia%27s%20National%20Digital%20Health%20Strategy%20-%20Safe%2C%20seamless%20and%20secure.pdf), the National Suicide Policy Platform by Suicide Prevention Australia (https://www.suicidepreventionaust.org/our-work/policy-and-advocacy/), National Mental Health Strategy by the Federal Department of Health (https://www.health.gov.au/health-topics/mental-health-and-suicide-prevention/mental-health-in-australia), National Safety and Quality in Digital Mental Health Standards by the Australian Commission on Safety and Quality in Health Care (https://www.safetyandquality.gov.au/standards/national-safety-and-quality-digital-mental-health-standards), and the Closing The Gap priorities (https://www.closingthegap.gov.au/).

Conclusion

The overarching reason for research on the effectiveness of mental health apps is to protect consumers. While it is desirable that mental health apps work to improve symptoms of mental illness, it is vital that mental health apps do not harm consumers. A consumer may turn to a mental health app during a time of significant emotional distress, and even during a time when they might be contemplating suicide. If that person in their vulnerable state reads misinformed or poorly articulated information, or engages in a questionable intervention as directed by an app without research evidence or credible development credentials, it may, sadly, potentially lead to that person taking their own life. At the very least, publicly available mental health apps claiming to offer therapeutic

treatments for mental illnesses should be developed by mental health professionals or organisations with mental health expertise and using an evidence-based framework. Furthermore, mental health professionals do not ordinarily engage in therapies (psychological or pharmaceutical) that have not proved to be efficacious and effective with thorough research, so why should mental health apps offering therapeutic treatments to mental illnesses be any different? It is incumbent on government organisations and professional associations to provide oversight on mental health apps in a similar manner to the way they provide oversight on other medical treatments. It is hoped this thesis has provided a guiding light on one method by which this can be achieved with mental health apps.

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF ORIGINALITY

We, the PhD candidate and the candidate's Principal Supervisor, certify that the following text, figures and diagrams are the candidate's original work.

Type of work	Page number/s
All aspects, except for the assistance described in the Statement of Authors' Contribution below.	N/A

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020
Date

[redacted]

Principal Supervisor

15/12/2020
Date

Higher Degree Research Thesis by Publication
University of New England

STATEMENT OF AUTHORS' CONTRIBUTION

We, the PhD candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated in the *Statement of Originality*.

	Author's Name (please print clearly)	**% of contribution**
Candidate	Jamie Marshall	65%
Other Authors	Debra Dunstan	25%
	Warren Bartik	10%

Name of Candidate: Jamie Marshall

Name/title of Principal Supervisor: Prof Debra Dunstan

[redacted]

Candidate

15/12/2020

Date

[redacted]

Principal Supervisor

15/12/2020

Date

Consolidated Reference List

A-Tjak, J. G. L., Davis, M. L., Morina, N., Powers, M. B., Smits, J. A. J., & Emmelkamp, P. M. G. (2015). A meta-analysis of the efficacy of acceptance and commitment therapy for clinically relevant mental and physical health problems. *Psychotherapy and Psychosomatics*, *84*(1), 30-36. https://doi.org/10.1159/000365764

Accenture. (2018). *Meet today's healthcare team: Patients + doctors + machines - Accenture 2018 consumer survey on digital health*. Retrieved November 22, 2019, from https://www.accenture.com/us-en/insight-new-2018-consumer-survey-digital-health?utm_source=newsletter&utm_medium=email&utm_campaign=newsletter_axiosvitals&stream=top-stories

Addis, M. E., & Jacobson, N. S. (2000). A closer look at the treatment rationale and homework compliance in cognitive-behavioral therapy for depression. *Cognitive Therapy and Research*, *24*, 313-326. https://doi.org/10.1023/A:1005563304265

Alyami, M., Giri, B., Alyami, H., & Sundrum, F. (2017). Social anxiety apps: A systematic review and assessment of app descriptors across mobile store platforms. *Evidence-Based Mental Health*, *20*(3), 65-70. https://doi.org/10.1136/eb-2017-102664

American Psychiatric Association. (2009). *Practice Guideline for the Treatment of Patients with Panic Disorder* (2nd ed.). https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/panicdisorder.pdf

American Psychiatric Association. (2010). *Practice Guideline for the Treatment of Patients with Major Depressive Disorder* (3rd ed.). https://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Ameringen, M. V., Turna, J., Khalesi, Z., Pullia, K., & Patterson, B. (2017). There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. *Depression and Anxiety*, *34*(6), 526-539. https://doi.org/10.1002/da.22657

Andrews, G., CuijPers, P., Craske, M. G., McEvoy, P., & Titov, N. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. *PloS One*, *5*(10), Article e13196. https://doi.org/10.1371/journal.pone.0013196

Antony, M. M., Bieling, P. J., Cox, B. J., Enns, M. W., & Swinson, R. P. (1998). Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. *Psychological Assessment*, *10*(2), 176-181. https://doi.org/10.1037/1040-3590.10.2.176

Anxiety Canada. (2019). *MindShift CBT*. Retrieved October 10, 2019, from https://anxietycanada.com/resources/mindshift-cbt/

App Annie. (2019). *The state of mobile in 2019*. Retrieved August 20, 2019, from https://www.appannie.com/en/insights/market-data/the-state-of-mobile-2019/

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications board task force report. *American Psychologist*, *73*(1), 3-25. https://doi.org/10.1037/amp0000191

Armontrout, J. A., Torous, J., Cohen, M., McNeil, D. E., & Binder, R. (2018). Current regulation of mobile mental health applications. *The Journal of the American Academy of Psychiatry and the Law*, *46*, 204-211. https://doi.org/10.29158/JAAPL.003748-18

Armontrout, J. A., Torous, J., Fisher, M., & Drogin, E. (2016). Mobile mental health: Navigating new rules and regulations for digital tools. *Current Psychiatry Reports*, *18*, Article 91. https://doi.org/10.1007/s11920-016-0726-x

Armstrong, C. M., Ciulla, R. P., Edwards-Stewart, A., Hoyt, T., & Bush, N. (2018). Best practices of mobile health in clinical care: The development and evaluation of a competency-based provider training program. *Professional Psychology: Research and Practice*, *49*(5-6), 355-363. https://doi.org/10.1037/pro0000194

Aryana, B., Brewster, L., & Nocera, J. A. (2018). Design for mobile mental health: An exploratory review. *Health and Technology*, 9, 401-424. https://doi.org/10.1007/s12553-018-0271-1

Australian Commission on Safety and Quality in Health Care. (2020). *National Safety and Quality Digital Mental Health Standards*. Retrieved August 12, 2020, from https://www.safetyandquality.gov.au/standards/national-safety-and-quality-digital-mental-health-standards

Australian Communications and Media Authority. (2016). *Aussie Teens and Kids Online*. Retrieved October 2, 2019, from http://www.acma.gov.au/theACMA/engage-blogs/engage-blogs/Research-snapshots/Aussie-teens-and-kids-online

Australian Government Department of Health. (2019a). *Head to Health*. Retrieved October 7, 2019, from https://headtohealth.gov.au

Australian Government Department of Health. (2019b). *Medicare Benefits Schedule - Item 80110*. Retrieved October 15, 2019, from http://www9.health.gov.au/mbs/fullDisplay.cfm?type=item&qt=ItemID&q=80110

Australian Government Department of Health. (2020). *COVID-19: $48.1 Million for National Mental Health and Wellbeing Pandemic Response Plan*. Retrieved August 12, 2020, from https://www.health.gov.au/ministers/the-hon-greg-hunt-mp/media/covid-19-481-million-for-national-mental-health-and-wellbeing-pandemic-response-plan

Australian National University. (2019). *beacon 2.0*. Retrieved October 7, 2019, from https://beacon.anu.edu.au/

Azarang, A., Pakyurek, M., Giroux, C., Nordahl, T. E., & Yellowlees, P. (2019). Information technologies: An augmentation to Post-Traumatic Stress Disorder treatment among trauma survivors. *Telemedicine and E-Health*, *25*(4), 263-271. https://doi.org/10.1089/tmj.2018.0068

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

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2018a). Development and pilot evaluation of smartphone-delivered cognitive behavior therapy strategies for mood- and anxiety-related problems: MoodMission. *Cognitive and Behavioral Practice*, *25*, 496-514. https://doi.org/10.1016/j.cbpra.2018.07.002

Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2018b). 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

Bakker, D., & Rickard, N. (2018). Engagement in mobile phone app for self-monitoring of emotional wellbeing predicts changes in mental health: MoodPrism. *Journal of Affective Disorders*, *227*, 432-442. https://doi.org/10.1016/j.jad.2017.11.016

Bakker, D., & Rickard, N. (2019). Engagement with a cognitive behavioural therapy mobile phone app predicts changes in mental health and wellbeing: MoodMission. *Australian Psychologist*, *54*(4), 245-260. https://doi.org/10.1111/ap.12383

Barak, A., & Grohol, J. M. (2011). Current and future trends in Internet-supported mental health interventions. *Journal of Technology in Human Sciences, 29(3),* 155-196. https://doi.org/10.1080/15228835.2011.616939

Barboutov, K., Furuskar, A., & Inam, R. (2017). *Ericsson Mobility Report*. Ericsson. https://www.ericsson.com/assets/local/mobility-report/documents/2017/ericsson-mobility-report-june-2017.pdf

Barlow, D. H., Nock, M. K., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior change* (3rd ed.). Pearson.

Barth, J., Munder, T., Gerger, H., Nuesch, E., Trelle, S., Znoj, H., Juni, P., & Cuijpers, P. (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: A network meta-analysis. *PLoS Medicine*, *10*(5), Article e1001454. https://doi.org/10.1371/journal.pmed.1001454

Basu, T. (2020, March 20). *The coronavirus pandemic is a game changer for mental health care*. MIT Technology Review. Retrieved April 11, 2020, from https://www.technologyreview.com/2020/03/20/905184/coronavirus-online-therapy-mental-health-app-teletherapy/

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression*. Guilford Press.

Beckstead, D. J., Hatch, A. L., Lambert, M. J., Eggett, D. L., Boates, M. K., & Vermeersch, D. A. (2003). Clinical significance of the outcome questionnaire (OQ-45.2). *The Behavior Analyst Today*, *4*(1), 86-97. https://doi.org/10.1037/h0100015

Beiwinkel, T., Kindermann, S., Maier, A., Kerl, C., Moock, J., Barbian, G., & Rossler, W. (2016). Using smartphones to monitor bipolar disorder symptoms: A pilot study. *JMIR Mental Health*, *3*(1), Article e2. https://doi.org/10.2196/mental.4560

Bennett-Levy, J., McManus, F., Westling, B. E., & Fennell, M. (2009). Acquiring and refining CBT skills and competencies: Which training methods are perceived to be most effective? *Behavioural and Cognitive Psychotherapy*, *37*(5), 571-583. https://doi.org/10.1017/S1352465809990270

Bennett-Levy, J., Richards, D. A., Farrand, P., Christensen, H., Griffiths, K. M., Kavanagh, D. J., Klein, B., Lau, M. A., Proudfoot, J., Ritterband, L., White, J., & Williams, C. (Eds.). (2010). *Oxford Guide To Low-Intensity CBT Interventions*. Oxford University Press Inc. https://doi.org/10.1093/med:psych/9780199590117.001.0001

Bentley, K. H., Kleiman, E. M., Elliott, G., Huffman, J. C., & Nock, M. K. (2019). Real-time monitoring technology in single-case experimental design research: Opportunities and challenges. *Behavior Research and Therapy*, *117*, 87-96. https://doi.org/10.1016/j.brat.2018.11.017

Beutler, L. E., Albanese, A. L., Fisher, D., Karno, M., Sandowicz, M., Williams, O. B., Gallagher-Thompson, D., & Thompson, L. (1999, June). *Selecting and matching to patient variables* [Paper presentation]. Annual Meeting of the Society for Psychotherapy Research, Braga, Portugal.

Beutler, L. E., & Clarkin, J. F. (1990). *Systematic treatment selection: Toward targeted therapeutic interventions*. Brunner Mazel.

Black Dog Institute. (2016). *Saving lives with nationally integrated e-mental health services*. Retrieved November 9, 2019, from http://www.blackdoginstitute.org.au/docs/BDI001_297x210_eHealth_Digital.pdf

Boisseau, C. L., Schwartzman, C. M., Lawton, J., & Mancebo, M. C. (2017). App-guided exposure and response prevention for obsessive compulsive disorder: an open pilot trial. *Cognitive Behaviour Therapy*, *46*(6), 447-458. https://doi.org/10.1080/16506073.2017.1321683

Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health*, *13*, Article 119. https://doi.org/10.1186/1471-2458-13-119

Borckardt, J. J., Nash, M. R., Murphy, M. D., Moore, M., Shaw, D., & O'Neil, P. (2008). Clinical practice as natural laboratory for psychotherapy research - A guide to case-based time series analysis. *American Psychologist*, *63*(2), 77-95. https://doi.org/10.1037/0003-066X.63.2.77

Boswell, D. L., White, J. K., Sims, W. D., Harrist, S., & Romans, J. S. C. (2013). Reliability and validity of the Outcome Questionnaire-45.2. *Psychological Reports: Mental and Physical Health*, *112*(3), 689-693. https://doi.org/10.2466/02.08.PR0.112.3.689-693

Bowers, H., Manion, I., Papadopolous, D., & Gauvreau, E. (2013). Stigma in school-based mental health: Perceptions of young people and service providers. *Child and Adolescent Mental Health*, *18*(3), 165-170. https://doi.org/10.1111/j.1475-3588.2012.00673.x

Broussard, J. D., & Teng, E. J. (2019). Models for enhancing the development of experiential learning approaches within mobile health technologies. *Professional Psychology: Research and Practice*, *50*(3), 195-203. https://doi.org/10.1037/pro0000234

Bry, L. J., Chou, T., Miguel, E., & Comer, J. S. (2018). Consumer smartphone apps marketed for child and adolescent anxiety: A systematic review and content analysis. *Behavior Therapy*, *49*(2), 249-261. https://doi.org/10.1016/j.beth.2017.07.008

Buckley, J. A., Speece, D. L., & McLaughlin, J. E. (2014). The role of single-case designs in supporting rigorous intervention development and evaluation at the Institute of Education Sciences. In T. R. Kratochwill & J. R. Levin (Eds.), *Methodological and statistical advances* (pp. 283-296). American Psychological Association. https://psycnet.apa.org/doi/10.1037/14376-000

Burns, P. (2002). *The robustness of the Ljung-Box Test and its rank equivalent*. SSRN. https://doi.org/10.2139/ssrn.443560

Bushnell, M. (2020). *What is the difference between an app and a mobile website?* Business News Daily. Retrieved September 19, 2021, from https://www.businessnewsdaily.com/6783-mobile-website-vs-mobile-app.html

Butler, A. C., Chapman, J. E., Forman, E. M., & Beck, A. T. (2006). The empirical status of cognitive-behavioral therapy: A review of meta-analyses. *Clinical Psychology Review*, *26*, 17-31. https://doi.org/10.1016/j.cpr.2005.07.003

Byambasuren, O., Sanders, S., Beller, E., & Glasziou, P. (2018). Prescribable mHealth apps identified from an overview of systematic reviews. *NPJ Digital Medicine*, *1*, Article 12. https://doi.org/10.1038/s41746-018-0021-9

Carey, T. A., Haviland, J., Tai, S. J., Vanags, T., & Mansell, W. (2016). MindSurf: a pilot study to assess the usability and acceptability of a smartphone app designed to promote contentment, wellbeing, and goal achievement. *BMC Psychiatry*, *16*, Article 442. https://doi.org/10.1186/s12888-016-1168-z

Carpenter-Song, E., Noel, V. A., Acquilano, S. C., & Drake, R. E. (2018). Real world technology use among people with mental illnesses: Qualitative study. *JMIR Mental Health*, *5*(4), Article e10652. https://doi.org/10.2196/10652

Carras, M. C., Mojtabai, R., Furr-Holden, C. D., Eaton, W., & Cullen, B. (2014). Use of mobile phones, computers and Internet among clients of an inner-city community psychiatric clinic. *Journal of Psychiatric Practice*, *20*(2), 94-103. https://doi.org/10.1097/01.pra.0000445244.08307.84

Carthaus, A. (2020). *Coronavirus and mental health: "We are not made for social isolation"*. DW News. Retrieved June 16, 2020, from https://www.dw.com/en/coronavirus-and-mental-health-we-are-not-made-for-social-isolation/a-52971173

Charney, M. E., Hellberg, S. N., Bui, E., & Simon, N. (2018). Evidence-based treatment of Posttraumatic Stress Disorder: An updated review of validated psychotherapeutic and pharmacological approaches. *Harvard Review of Psychiatry*, *26*(3), 99-115. https://doi.org/10.1097/HRP.0000000000000186

Chowdhury, H. (2020). *Mental health apps downloaded more than 1m times since start of virus outbreak*. The Telegraph. Retrieved June 15, 2020, from https://www.telegraph.co.uk/technology/2020/05/16/mental-health-apps-downloaded-1m-times-since-start-virus-outbreak/

Christensen, H., Proudfoot, J., Woodward, A., Hosie, A., Klein, B., Morgan, C., Kavanagh, D., Andrews, G., Harman, G., Burns, J., Blanchard, M., Zachan, J., Griffiths, K., Stroud, K., Teeson, M., Kay-Lambkin, F., Thomas, N., Scharr, S., Spence, S., & Humphreys, T. (2014). *e-Mental Health Services In Australia 2014: Current And Future*. e-Mental Health Alliance. https://emhalliance.fedehealth.org.au/wp-content/uploads/sites/42/2014/10/e-Mental-Health-in-Australia-2014.pdf

Christoforou, M., Fonseca, J. A. S., & Tsakanikos, E. (2017). Two new cognitive behavioral therapy-based mobile apps for agoraphobia: Randomized controlled trial. *Journal of Medical Internet Research*, *19*(11), Article e398. https://doi.org/10.2196/jmir.7747

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., Leucht, S., Ruhe, H. G., Turner, E. H., Higgins, J. P. T., Egger, M., Takeshima, N., Hayasaka, Y., Imai, H., Shinohara, K., Tajika, A., Ioannidis, J. P. A., & Geddes, J. R. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *Lancet*, *391*(10128), 1357-1366. https://doi.org/10.1016/S0140-6736(17)32802-7

Cipriani, A., Furukawa, T. A., Salanti, G., Geddes, J. R., Higgins, J. P. T., Churchill, R., Watanabe, N., Nakagawa, A., Omori, I. M., McGuire, H., Tansella, M., & Barbui, C. (2009). Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *The Lancet*, *373*(9665), 746-758. https://doi.org/10.1016/S0140-6736(09)60046-5

Clough, B. A., & Casey, L. M. (2015a). Smart designs for smart technologies: Research challenges and emerging solutions for scientist-practitioners within e-mental health. *Professional Psychology: Research and Practice*, *46*(6), 429-436. https://doi.org/10.1037/pro0000053

Clough, B. A., & Casey, L. M. (2015b). The smart therapist: A look to the future of smartphones and mHealth technologies in psychotherapy. *Professional Psychology: Research and Practice*, *46*(3), 147-153. https://doi.org/10.1037/pro0000011

Concilio. (2019). *By 2018, New Zealand will have 90% smartphone and 78% tablet ownership*. Retrieved April 20, 2019, from http://www.concilio.co.nz/blog-layout/by-2018-new-zealand-will-have-90-smartphone-and-78-tablet-ownership

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T. F., & Reynolds, C. F. (2013). The efficacy of psychotherapy and pharmacotherapy in treating depressive and anxiety disorders: A meta-analysis of direct comparisons. *World Psychiatry*, *12*(2), 137-148. https://doi.org/10.1002/wps.20038

Davies, E. B., Morriss, R., & Glazebrook, C. (2014). Computer-delivered and web-based interventions to improve depression, anxiety, and psychological wellbeing of university students: A systematic review and meta-analysis. *Journal of Medical Internet Research*, *16*(5), Article e130. https://doi.org/10.2196/jmir.3142

de Mello, M. F., Mari, J. d. J., Bacaltchuk, J., Verdeli, H., & Neugebauer, R. (2005). A systematic review of research findings on the efficacy of interpersonal therapy for depressive disorders. *European Archives of Psychiatry and Clinical Neuroscience*, *255*(2), 75-82. https://doi.org/10.1007/s00406-004-0542-x

Deloitte. (2015). *Mobile Consumer Survey 2015 - The Australian Cut*. http://landing.deloitte.com.au/rs/761-IBL-328/images/deloitte-au-tmt-mobile-consumer-survey-2015-291015.pdf

Deloitte. (2018). *Behaviour Unlimited Mobile Consumer Survey 2018 The Australian Cut*. https://www2.deloitte.com/au/mobile-consumer-survey

Department of Health. (2019). *Better Access Telehealth Services for People in Rural and Remote Areas*. Retrieved April 10, 2020, from https://www1.health.gov.au/internet/main/publishing.nsf/Content/mental-ba-telehealth

Department of Health. (2020). *COVID-19 National Health Plan - Primary Care Package - MBS Telehealth Services and Increased Practice Incentive Payments*. Retrieved April 11, 2020, from https://www.health.gov.au/resources/publications/covid-19-national-health-plan-primary-care-package-mbs-telehealth-services-and-increased-practice-incentive-payments

Des Jarlais, D. C., Lyles, C., Crepaz, N., & The TREND Group. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND Statement. *American Journal of Public Health*, *94*, 361-366. https://doi.org/10.2105/AJPH.94.3.361

Dillon, A., Kelly, M., Robertson, I., & Robertson, D. A. (2016). Smartphone applications utilizing biofeedback can aid stress reduction. *Frontiers in Psychology*, *7*, Article 832. https://doi.org/10.3389/fpsyg.2016.00832

Dimidjian, S., Arch, J. J., Schneider, R. L., Desormeau, P., Felder, J. N., & Segal, Z. V. (2016). Considering meta-analysis, meaning, and metaphor: A systematic review and critical examination of "Third Wave" cognitive and behavioral therapies. *Behavior Theory*, *47*(6), 886-905. https://doi.org/10.1016/j.beth.2016.07.002

Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., & Christensen, H. (2013). Smartphones for smarter delivery of mental health programs: A systematic review. *Journal of Medical Internet Research*, *15*(11), Article e247. https://doi.org/10.2196/jmir.2791

Donley, N., & Gillam, C. (2018). *Roundup Case Demonstrates Importance of Independence in Scientific Evidence*. San Francisco Chronicle. Retrieved January 7, 2020, from

https://www.sfchronicle.com/opinion/openforum/article/SF-Roundup-case-demonstrates-importance-of-13192330.php

Dunn, A. (2020). *Experts are calling for expanded Medicare rebates during the COVID-19 crisis*. Special Broadcasting Service (SBS) Bulgarian. Retrieved April 11, 2020, from https://www.sbs.com.au/language/english/audio/experts-are-calling-for-expanded-medicare-rebates-during-the-covid-19-crisis

Ebert, D. D., Zarski, A.-C., Christensen, H., Stikkelbroek, Y., CuijPers, P., Berking, M., & Riper, H. (2015). Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: A meta-analysis of randomized controlled outcome trials. *PloS One*, *10*(3), Article e0119895. https://doi.org/10.1371/journal.pone.0119895

Ehde, D. M., Dillworth, T. M., & Turner, J. A. (2014). Cognitive-behavioural therapy for individuals with chronic pain: Efficacy, innovations, and directions for research. *American Psychologist*, *69*(2), 153-166. https://doi.org/10.1037/a0035747

El-Gayar, O., Timsina, P., Nawar, N., & Eid, W. (2013). Mobile applications for diabetes self-management: Status and potential. *Journal of Diabetes Science and Technology*, *7*(1), 247-262. https://doi.org/10.1177%2F193229681300700130

Emanuel, L., Berwick, D., Conway, J., Combes, J., Hatlie, M., Leape, L., Reason, J., Schyve, P., Vincent, C., & Walton, M. (2008). What exactly is patient safety? *Advances in Patient Safety: New Directions and Alternative Approaches*, *1*, 1-18.

Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence-Based Mental Health*, *1*, 70-72. https://doi.org/10.1136/ebmh.1.3.70

Eysenbach, G., & CONSORT-EHEALTH Group. (2011). CONSORT-EHEALTH: Improving and standardizing evaluation reports of web-based and mobile health interventions. *Journal of Medical Internet Research*, *13*(4), Article e126. https://doi.org/10.2196/jmir.1923

Faurholt-Jepson, M., Busk, J., Porarinsdottir, H., Frost, M., Bardram, J. E., Vinberg, M., & Kessing, L. V. (2019). Objective smartphone data as a potential diagnostic marker of bipolar disorder. *Australian and New Zealand Journal of Psychiatry*, *53*(2), 119-128. https://doi.org/10.1177/00048674188089

Feng, C. (2020, March 30). *China's mental health counsellors reach out online to help those affected by coronavirus pandemic*. South China Morning Post. https://www.scmp.com/tech/apps-social/article/3077260/chinas-mental-health-counsellors-reach-out-online-help-those

Firth, J., Torous, J., Carney, R., Newby, J. M., Cosco, T. D., Christensen, H., & Sarris, J. (2018). Digital technologies in the treatment of anxiety: Recent innovations and future directions. *Current Psychiatry Reports*, *20*(6), Article 44. https://doi.org/10.1007/s11920-018-0910-2

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

Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S., & Sarris, J. (2017). Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, *218*, 15-22. https://doi.org/10.1016/j.jad.2017.04.046

Fleming, T. M., Bavin, L., Lucassen, M., Stasiak, K., Hopkins, S., & Merry, S. N. (2018). Beyond the trial: Systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *Journal of Medical Internet Research*, *20*(6), Article e199. https://doi.org/10.2196/jmir.9275

Flett, J. A. M., Hayne, H., Riordan, B. C., Thompson, L. M., & Conner, T. S. (2019). Mobile mindfulness meditation: A randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness*, *10*(5), 863-876. https://doi.org/10.1007/s12671-018-1050-9

Frank, E., Pong, J., Asher, Y., & Soares, C. N. (2018). Smart phone technologies and ecological momentary data: is this the way forward on depression management and research? *Current Opinion In Psychiatry*, *31*(1), 3-6. https://doi.org/10.1097/YCO.0000000000000382

Freedman, S. R., & Enright, R. D. (1996). Forgiveness as an intervention goal with incest survivors. *Journal of Consulting and Clinical Psychology*, *64*(5), 983-992. https://doi.org/10.1037/0022-006X.64.5.983

Froh, J. J., Sefick, W. J., & Emmons, R. A. (2008). Counting blessings in early adolescence: An experimental study of gratitude and subjective wellbeing. *Journal of School Psychology*, *46*(2), 213-233. https://doi.org/10.1016/j.jsp.2007.03.005

Gaziano, J. M., Sesso, H. D., Christen, W. G., Bubes, V., Smith, J. P., MacFadyen, J., Shvartz, M., Manson, J. E., Glynn, R. J., & Buring, J. E. (2012). Multivitamins in the prevention of cancer in men: The Physicians' Health Study II randomized controlled study. *Journal of the American Medical Association*, *308*(18), 1871-1880. https://doi.org/10.1001/jama.2012.14641

Gratzer, D., Khalid-Khan, F., & Balasingham, S. (2018). Chapter 11 - The Internet and CBT: A New Clinical Application of an Effective Therapy. In O. Senormanci (Ed.), *Cognitive behavioral therapy and clinical applications* (pp. 201-222). Intech Open. https://doi.org/10.5772/intechopen.72146

Gun, S. Y., Titov, N., & Andrews, G. (2011). Acceptability of Internet treatment of anxiety and depression. *Australasian Psychiatry*, *19*(3), 259-264. https://doi.org/10.3109/10398562.2011.562295

Gunter, R. W., & Whittal, M. L. (2010). Dissemination of cognitive-behavioral treatments for anxiety disorders: Overcoming barriers and improving patient access. *Clinical Psychology Review*, *30*(2), 194-202. https://doi.org/10.1016/j.cpr.2009.11.001

Hamilton, K. E., & Dobson, K. S. (2002). Cognitive therapy of depression: Pretreatment patient predictors of outcome. *Clinical Psychology Review*, *22*(6), 875-893. https://doi.org/10.1016/S0272-7358(02)00106-X

Harvard Medical School. (2008). Positive psychology in practice. *Harvard Mental Health Letter*. Retrieved October 18, 2020, from https://www.health.harvard.edu/mind-and-mood/positive_psychology_in_practice

Hatfield, D. R., & Ogles, B. M. (2004). The use of outcome measures by psychologists in clinical practice. *Professional Psychology: Research and Practice*, *35*(5), 485-491. https://doi.org/10.1037/0735-7028.35.5.485

Hayes, S. C., & Hofmann, S. G. (2017). The third wave of cognitive behavioral therapy and the rise of process-based care. *World Psychiatry*, *16*(3), 245-246. https://doi.org/10.1002/wps.20442

Heilweil, R. (2020, March 20). *Feeling anxious about coronavirus? There's an app for that*. Vox. https://www.vox.com/recode/2020/3/20/21185351/mental-health-apps-coronavirus-pandemic-anxiety

Hempel, C., Sezier, A., & Terry, G. (2018). What helps or hinders clinicians in their decision-making processes when using or prescribing mHealth apps in practice? An exploratory study. *New Zealand Journal of Physiotherapy*, *46*(2), 72-77. https://doi.org/10.15619/NZJP/46.2.04

Hendrikoff, L., Kambeitz-Ilankovic, L., Pryss, R., Senner, F., Falkai, P., Pogarell, O., Hasan, A., & Peters, H. (2019). Prospective acceptance of distinct mobile mental health features in psychiatric patients and mental health professionals. *Journal of Psychiatric Research*, *109*, 126-132. https://doi.org/10.1016/j.jpsychires.2018.11.025

Henry, J. D., & Crawford, J. R. (2005). The short-form version of the Depression Anxiety Stress Scales (DASS-21): Construct validity and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, *44*(2), 227-239. https://doi.org/10.1348/014466505X29657

Hernan, A., Philpot, B., Edmonds, A., & Reddy, P. (2010). Healthy minds for country youth: Help-seeking for depression among rural adolescents. *The Australian Journal of Rural Health*, *18*(3), 118-124. https://doi.org/10.1111/j.1440-1584.2010.01136.x

Hewlett, E., & Moran, V. (2014). *Making Mental Health Count: The Social and Economic Costs of Neglecting Mental Health Care*. OECD Health Policy Studies. OECD Publishing. https://read.oecd-ilibrary.org/social-issues-migration-health/making-mental-health-count_9789264208445-en#page1

Hickie, I. B., Davenport, T. A., Luscombe, G. M., Moore, M., Griffiths, K. M., & Christensen, H. (2010). Practitioner-supported delivery of internet-based cognitive behaviour therapy: evaluation of the feasibility of conducting a cluster randomised trial. *Medical Journal of Australia*, *192*(11), S31-S35. https://doi.org/10.5694/j.1326-5377.2010.tb03690.x

Hidalgo-Mazzei, D., & Young, A. H. (2019). Psychiatry foretold. *Australian and New Zealand Journal of Psychiatry*, *53*(4), 365-366. https://doi.org/10.1177/0004867418816821

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), Article e301. https://doi.org/10.2196/jmir.6342

Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, *78*(2), 169-183. https://doi.org/10.1037/a0018555

Hollis, C., Sampson, S., Simons, L., Davies, E. B., Churchill, R., Betton, V., Butler, D., Chapman, K., Easton, K., Gronlund, T. A., Kabir, T., Rawsthorne, M., Rye, E., & Tomlin, A. (2018). Identifying research priorities for digital technology in mental health care: Results of the James Lind Alliance Priority Setting Partnership. *Lancet Psychiatry*, *5*(10), 845-854. https://doi.org/10.1016/S2215-0366(18)30296-7

Holloway, D., Green, L., & Livingstone, S. (2013). *Zero to eight: Young children and their Internet use*. ECU Publications 2013. Edith Cowan University Research Online. http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1930&context=ecuworks2013

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidenced-based practice in special education. *Exceptional Children*, *71*(2), 165-179. https://doi.org/10.1177%2F001440290507100203

Houle, T. T. (2009). Statistical analyses for single-case experimental designs. In D. H. Barlow, M. K. Nock, & M. Hersen (Eds.), *Single-case experimental designs - Strategies for studying behavior change* (3rd ed., pp. 271-306). Pearson Education, Inc.

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: Causal mediation of change. *Journal of Consulting and Clinical Psychology*, *61*(4), 678-685. https://doi.org/10.1037/0022-006X.61.4.678

Howells, A., Ivtzan, I., & Eiroa-Orosa, F. J. (2016). Putting the "app" in happiness: A randomised controlled trial of a smartphone-based mindfulness intervention to enhance wellbeing. *Journal of Happiness Studies*, *17*, 163-185. https://doi.org/10.1007/s10902-014-9589-1

Huang, H.-Y., & Bashir, M. (2017). Users' adoption of mental health apps: Examining the impact of information cues. *JMIR MHealth and UHealth*, *5*(6), Article e83. https://doi.org/10.2196/mhealth.6827

Huguet, A., Rao, S., McGrath, P. J., Wozney, L., Wheaton, M., Conrod, J., & Rozario, S. (2016). A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PloS One*, *11*(5), Article e0154248. https://doi.org/10.1371/journal.pone.0154248

Hussain, M., Al-Haiqi, A., Zaidan, A. A., Zaidan, B. B., Kiah, M. L. M., Anuar, N. B., & Abdulnabi, M. (2015). The landscape of research on smartphone medical apps: Coherent taxonomy, motivations, open challenges and recommendations. *Computer Methods and Programs in Biomedicine*, *122*(3), 393-408. https://doi.org/10.1016/j.cmpb.2015.08.015

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), Article e124. https://doi.org/10.1371/journal.pmed.0020124

IQVIA Institute. (2017). *The Growing Value of Digital Health*. IQVIA Institute For Human Data Science. Retrieved January 25, 2019, from https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health

Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *Journal of Consulting and Clinical Psychology*, *67*(3), 300-307. https://doi.org/10.1037/0022-006X.67.3.300

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12-19. https://doi.org/10.1037/0022-006X.59.1.12

Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: examining and forecasting change. *Frontiers in Psychology*, *6*, Article 727. https://doi.org/10.3389/fpsyg.2015.00727

Jennings, L. (2018). Antidepressants. In G. T. Grossberg & L. J. Kinsella (Eds.), *Clinical psychopharmacology for neurologists - A practical guide* (pp. 45-71). Springer, Cham. https://doi.org/10.1007/978-3-319-74604-3

Jessen, S., Mirkovic, J., & Ruland, C. M. (2018). Creating gameful design in mHealth: A participatory co-design approach. *JMIR MHealth and UHealth*, *6*(12), Article e11579. https://doi.org/10.2196/11579

Jones, E., Lebrun-Harris, L. A., Sripipatana, A., & Ngo-Metzger, Q. (2014). Access to mental health services among patients at health centers and factors associated with unmet needs. *Journal of Health Care for the Poor and Underserved*, *25*(1), 425-436. https://doi.org/10.1353/hpu.2014.0056

Jones, N., & Moffitt, M. (2016). Ethical guidelines for mobile app development within health and mental health fields. *Professional Psychology: Research and Practice*, *47*(2), 155-162. https://doi.org/10.1037/pro0000069

Jorm, A. F. (2012). Mental health literacy: Empowering the community to take action for better mental health. *American Psychologist*, *67*(3), 231-243. https://doi.org/10.1037/a0025957

Kallivayalil, R. A. (2020). COVID-19 and Mental Health: Working together in India. *World Psychiatric Association*. Retrieved June 16, 2020, from https://www.wpanet.org/post/covid-19-and-mental-health-working-together-in-india

Kaneko, Y., & Motohashi, Y. (2007). Male gender and low education with poor mental health literacy: A population-based study. *Journal of Epidemiology*, *17*(4), 114-119. https://doi.org/10.2188/jea.17.114

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, *67*(3), 332-339. https://doi.org/10.1037/0022-006X.67.3.332

Kazdin, A. E. (2017). *Research design in clinical psychology* (5th ed.). Pearson.

Kertz, S. J., MacLaren-Kelly, J., Stevens, K. T., Schrock, M., & Danitz, S. B. (2017). A review of free iPhone applications designed to target anxiety and worry. *Journal of Technology in Behavioral Science*, *2*(2), 61-70. https://doi.org/10.1007/s41347-016-0006-y

Khalaf, S. (2016). *Media, productivity and emojis give mobile another stunning growth year*. Flurry Analytics. Retrieved November 18, 2019, from http://flurrymobile.tumblr.com/post/136677391508/stateofmobile2015

Kim, J. W., Torous, J., Chan, S., & Gipson, S. Y.-M. T. (2018). Developing a digitally informed curriculum in psychiatry education and clinical practice. *Academic Psychiatry*, *42*, 782-790. https://doi.org/10.1007/s40596-018-0895-5

Kinderman, P., Hagan, P., King, S., Bowman, J., Chahal, J., Gan, L., McKnight, R., Waldon, C., Smith, M., Gilbertson, J., & Tai, S. J. (2016). The feasibility and effectiveness of Catch It, an

innovative CBT smartphone app. *BJPsych Open*, *2*(3), 204-209. https://doi.org/10.1192/bjpo.bp.115.002436

Klapdor, M. (2020, March 23). *COVID-19 economic response - social security measures part 2: $750 lump sum payments*. Parliament of Australia. https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/FlagPost/2020/March/Coronavirus_lump-sum_payments

Knesebeck, O. V. D., Mnich, E., Daubmann, A., Wegscheider, K., Angermeyer, M. C., Lambert, M., Karow, A., Harter, M., & Kofahl, C. (2013). Socioeconomic status and beliefs about depression, schizophrenia and eating disorders. *Social Psychiatry and Psychiatric Epidemiology*, *48*(5), 775-782. https://doi.org/10.1007/s00127-012-0599-1

Koh, D. (2020, May 20). *New mental health program announced to support Australian healthcare workers during COVID-19*. Healthcare IT News. https://www.healthcareit.com.au/article/new-mental-health-program-announced-support-australian-healthcare-workers-during-covid-19

Krasny-Pacini, A., & Evans, J. (2018). Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. *Annals of Physical and Rehabilitation Medicine*, *61*(3), 164-179. https://doi.org/10.1016/j.rehab.2017.12.002

Krebs, P., & Duncan, D. T. (2015). Health app use among US mobile phone owners: A national survey. *JMIR mHealth and uHealth*, *3*(4), Article e101. https://doi.org/10.2196/mhealth.4924

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Sage.

Kuhn, E., Kanuri, N., Hoffman, J. E., Garvert, D. W., Ruzek, J. I., & Taylor, C. B. (2017). A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. *Journal of Consulting and Clinical Psychology*, *85*(3), 267-273. https://doi.org/10.1037/ccp0000163

Labdoor. (2019). *Shop Smart for ....* Retrieved January 7, 2020, from https://labdoor.com/

Lai, J., & Jury, A. (2018). *Effectiveness of e-mental health approaches: Rapid review*. Te Pou. Retrieved January 20, 2020, from https://www.tepou.co.nz/uploads/files/resource-assets/E-therapy%20report%20FINAL%20July%202018.pdf

Lambert, M. J. (2012). Helping clinicians to use and learn from research-based systems: The OQ-Analyst. *Psychotherapy*, *49*(2), 109-114. https://doi.org/10.1037/a0027110

Lambert, M. J., & Finch, A. E. (1999). The Outcome Questionnaire. In M.E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcome assessment* (2nd ed., pp. 831-870). Earlbaum.

Lambert, M. J., Morton, J. S., Hatfield, D. R., Harmon, C., Hamilton, S., Reid, R. C., & Burlingame, G. M. (2004). *Administration and scoring manual for the OQ-45.2 Outcome Questionnaire*. OQ Measures LLC.

Lan, A., Lee, A., Munroe, K., McRae, C., Kaleis, L., Keshavjee, K., & Guergachi, A. (2018). Review of cognitive behavioural therapy mobile apps using a reference architecture embedded in the patient-provider relationship. *Biomedical Engineering Online*, *17*, Article 183. https://doi.org/10.1186/s12938-018-0611-4

Lee, R. A., & Jung, M. E. (2018). Evaluation of an mHealth app (Destressify) on university students' mental health: Pilot trial. *JMIR Mental Health*, *5*(1), Article e2. https://doi.org/10.2196/mental.8324

Lewis, G., Duffy, L., Ades, A., Amos, R., Araya, R., Brabyn, S., Button, K. S., Churchill, R., Derrick, C., Dowrick, C., Gilbody, S., Fawsitt, C., Hollingworth, W., Jones, V., Kendrick, T., Kessler, D., Kounali, D., Khan, N., Lanham, P., … Lewis, G. (2019). The clinical effectiveness of sertraline in primary care and the role of depression severity and duration (PANDA): A pragmatic, double-blind, placebo-controlled randomised trial. *Lancet Psychiatry*, *6*(11), 903-914. https://doi.org/10.1016/S2215-0366(19)30366-9

Lifeline. (2020, March 18). *Lifeline will continue answering calls through COVID-19.* https://www.lifeline.org.au/static/uploads/files/20200318-lifeline-to-continue-answering-calls-through-coronavirus-final-wfmgdbkq.pdf

Liu, S., Yang, L., Zhang, C., Xiang, Y.-T., Liu, Z., Hu, S., & Zhang, B. (2020). Online mental health services in China during the COVID-19 outbreak. *Lancet Psychiatry*, *7*(4), e17-e18. https://doi.org/10.1016/S2215-0366(20)30077-8

Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry, 18*(3), 325-336. https://doi.org/10.1002/wps.20673

Lobo, M. A., Moeyaert, M., Cunha, A. B., & Babik, I. (2017). Single-case design, analysis, and quality assessment for intervention research. *Journal of Neurologic Physical Therapy*, *41*(3), 187-197. https://doi.org/10.1097/NPT.0000000000000187

Lorant, V., Croux, C., Weich, S., Deliege, D., Mackenbach, J., & Ansseau, M. (2007). Depression and socio-economic risk factors: 7-year longitudinal population study. *British Journal of Psychiatry*, *190*(4), 293-298. https://doi.org/10.1192/bjp.bp.105.020040

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335-343. https://doi.org/10.1016/0005-7967(94)00075-U

Luchini, C., Stubbs, B., Solmi, M., & Veronese, N. (2017). Assessing the quality of studies in meta-analyses: Advantages and limitations of the Newcastle Ottawa Scale. *World Journal of Meta-Analysis*, *5*(4), 80-84. https://doi.org/10.13105/wjma.v5.i4.80

Lui, J. H. L., Marcus, D. K., & Barry, C. T. (2017). Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. *Professional Psychology: Research and Practice*, *48*(3), 199-210. https://doi.org/dx.doi.org/10.1037/pro0000122

Lynch, T. R., Chapman, A. L., Rosenthal, M. Z., Kuo, J. R., & Linehan, M. M. (2006). Mechanisms of change in dialectical behavior therapy: Theoretical and empirical observations. *Journal of Clinical Psychology*, *62*(4), 459-480. https://doi.org/10.1002/jclp.20243

Lyubomirsky, S., Sheldon, K. M., & Schkade, D. (2005). Pursuing happiness: The architecture of sustainable change. *Review of General Psychology*, *9*(2), 111-131. https://doi.org/10.1037/1089-2680.9.2.111

Machalicek, W., & Horner, R. H. (2018). Special issue on advances in single-case research design and analysis. *Developmental Neurorehabilitation*, *21*(4), 209-211. https://doi.org/10.1080/17518423.2018.1468600

Magill, M., & Ray, L. A. (2009). Cognitive-behavioral treatment with adult alcohol and illicit drug users: A meta-analysis of randomized controlled trials. *Journal of Studies on Alcohol and Drugs*, *70*(4), 516-527. https://doi.org/10.15288/jsad.2009.70.516

Makovsky. (2015). *Fifth Annual Pulse of Online Search*. Retrieved October 30, 2019, from http://www.makovsky.com/news/fifth-annual-pulse-of-online-health-survey-2/

Makovsky. (2016). *Sixth Annual Pulse Of Online Search*. Retrieved December 12 2019, from http://www.makovsky.com/news/makovsky-pulse-of-online-search-survey-2016/

Manicavasagar, V., Horswood, D., Burckhardt, R., Lum, A., Hadzi-Pavlovic, D., & Parker, G. (2014). Feasibility and effectiveness of a web-based positive psychology program for youth mental health: Randomized controlled trial. *Journal of Medical Internet Research*, *16*(6), Article e140. https://doi.org/10.2196/jmir.3176

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2019). The digital psychiatrist: In search of evidence-based apps for anxiety and depression. *Frontiers in Psychiatry*, *10*, Article 831. https://doi.org/10.3389/fpsyt.2019.00831

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020a). Apps with maps - Anxiety and depression mobile apps with evidence-based frameworks: Systematic search of major app stores. *JMIR Mental Health*, *7*(6), Article e16525. https://doi.org/10.2196/16525

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020b). Clinical or gimmickal: The use and effectiveness of mobile mental health apps for treating anxiety and depression. *Australian and New Zealand Journal of Psychiatry*, *54*(1), 20-28. https://doi.org/10.1177%2F0004867419876700

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020c). Effectiveness of using mental health mobile apps as digital antidepressants for reducing anxiety and depression: Protocol for a multiple baseline across-individuals design. *JMIR Research Protocols*, *9*(7), Article e17159. https://doi.org/10.2196/17159

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020d). The role of digital mental health resources to treat trauma symptoms in Australia during COVID-19. *Psychological Trauma: Theory, Research, Practice and Policy*, *12*(S1), S269-S271. https://doi.org/10.1037/tra0000627

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

Marshall, J. M., Dunstan, D. A., & Bartik, W. (2020f). Treating psychological trauma in the midst of COVID-19: The role of smartphone apps. *Frontiers in Public Health*, *8*, Article 402. https://doi.org/10.3389/fpubh.2020.00402

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, *67*(3), 267-277. https://doi.org/10.1016/j.jclinepi.2013.08.015

McDermut, W., Miller, I. W., & Brown, R. A. (2001). The efficacy of group psychotherapy for depression: A meta-analysis and review of the empirical research. *Clinical Psychology Science and Practice*, *8*(1), 98-116. https://doi.org/10.1093/clipsy.8.1.98

Medhora, S. (2020, March 27). *Calls to Lifeline jump 20 per cent as coronavirus crisis takes hold*. ABC Triple J Hack. https://www.abc.net.au/triplej/programs/hack/calls-to-lifeline-go-up-due-to-coronavirus-covid-19/12096922

Medical Startups. (2020). *Top 84 startups developing health apps*. Retrieved September 12, 2020, from https://www.medicalstartups.org/top/apps/

Mehrotra, S., & Tripathi, R. (2018). Recent developments in the use of smartphone interventions for mental health. *Current Opinion In Psychiatry*, *31*(5), 379-388. https://doi.org/10.1097/YCO.0000000000000439

Memon, A., Taylor, K., Mohebati, L. M., Sundin, J., Cooper, M., Scanlon, T., & de Visser, R. (2016). Perceived barriers to accessing mental health services among black and minority ethnic communities: A qualitative study in Southeast England. *BMJ Open*, *6*(11), Article e012337. https://doi.org/doi:10.1136/bmjopen-2016-012337

Menon, V., Rajan, T. M., & Sarkar, S. (2017). Psychotherapeutic applications of mobile phone-based technologies: A systematic review of current research and trends. *Indian Journal of Psychological Medicine*, *39*(1), 4-11. https://doi.org/10.4103/0253-7176.198956

Merry, S. N., Stasiak, K., Shepherd, M., Frampton, C., Fleming, T., & Lucassen, M. F. G. (2012). The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. *The BMJ*, *344*, Article e2598. https://doi.org/10.1136/bmj.e2598

Miller, K. E., Kuhn, E., Yu, J., Owen, J. E., Jaworski, B. K., Taylor, K., Blonigen, D. M., & Possemato, K. (2019). Use and perceptions of mobile apps for patients among VA primary care mental and behavioral health provders. *Professional Psychology: Research and Practice*, *50*(3), 204-209. https://doi.org/10.1037/pro0000229

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

Mindfood. (2020, April 14). *NZ Government releases apps to help with mental health during COVID-19 crisis*. https://www.mindfood.com/article/nz-government-releases-apps-to-help-with-mental-health-during-covid-19-crisis/

Mohadis, H. M., & Ali, N. M. (2014, September). *A study of smartphone usage and barriers among the elderly* [Paper presentation]. 3rd International Conference on User Science and Engineering, Shah Alam, Malaysia.

Mohr, D. C., Burns, M. N., Schueller, S. M., Clarke, G., & Klinkman, M. (2013). Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. *General Hospital Psychiatry*, *35*(4), 332-338. https://doi.org/10.1016/j.genhosppsych.2013.03.008

Mohr, D. C., Cheung, K., Schueller, S. M., Brown, C. H., & Duan, N. (2013). Continuous evaluation of evolving behavioral intervention technologies. *American Journal of Preventative Medicine*, *45*(4), 517-523. https://doi.org/10.1016/j.amepre.2013.06.006

Mohr, D. C., CuijPers, P., & Lehmann, K. (2011). Supportive accountability: A model for providing human support to enhance adherence to eHealth interventions. *Journal of Medical Internet Research*, *13*(1), Article e30. https://doi.org/10.2196/jmir.1602

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of*

Internal Medicine, *151*(4), 264-269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

MoodMission. (2019). *MoodMission: Change the way you feel*. Retrieved October 10, 2019, from http://moodmission.com/

Moshi, M. R., Tooher, R., & Merlin, T. (2018). Suitability of current evaluation frameworks for use in the health technology assessment of mobile medical applications: A systematic review. *International Journal of Technology Assessment in Health Care*, *34*(5), 464-475. https://doi.org/10.1017/S026646231800051X

Mougin, S. (2020). *Covid-19 debrief: Paris nurses, doctors reveal heavy toll on mental health*. France 24. Retrieved June 16, 2020, from https://www.france24.com/en/20200521-covid-19-debrief-nurses-and-doctors-share-feelings-and-feedback

Mrazek, P., & Haggerty, R. (1994). *Reducing the risks for mental disorders: Frontiers for preventive intervention research*. National Academy Press.

Mueller, R. M., Lambert, M. J., & Burlingame, G. M. (1998). Construct validity of the Outcome Questionnaire: A confirmatory factor analysis. *Journal of Personality Assessment*, *70*(2), 248-262. https://doi.org/10.1207/s15327752jpa7002_5

Mukherjee, J. (2020, April 9). *We Tried These 4 Mental Health Apps to Help us Cope With Stress During Covid-19 Lockdown*. News 18. https://www.news18.com/news/buzz/we-tried-these-4-mental-health-apps-to-help-us-cope-with-stress-during-covid-19-lockdown-2553235.html

Mushtaq, R. (2011). *Augmented Dickey Fuller test*. SSRN. Retrieved January 20, 2020, from https://doi.org/10.2139/ssrn.1911068

Naslund, J. A., Aschbrenner, K. A., Kim, S. J., McHugo, G. J., Unutzer, J., Bartels, S. J., & Marsch, L. A. (2017). Health behavior models for informing digital technology interventions for individuals with mental illness. *Psychiatric Rehabilitation Journal*, *40*(3), 325-335. https://doi.org/10.1037/prj0000246

National Health and Medical Research Council. (2018). *National statement on ethical conduct in human research (2007) - Updated 2018*. Retrieved October 13, 2019, from https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2007-updated-2018

National Health Service. (2019). *Mental health apps*. Retrieved October 7, 2019, from https://www.nhs.uk/apps-library/category/mental-health/

Neary, M., & Schueller, S. M. (2018). State of the field of mental health apps. *Cognitive and Behavioral Practice*, *25*(4), 531-537. https://psycnet.apa.org/doi/10.1016/j.cbpra.2018.01.002

Nehabaluni. (2018). *Mobile medical apps: A game changing healthcare innovation*. HealthWorks Collective. Retrieved January 9, 2020, from https://www.healthworkscollective.com/mobile-medical-apps-a-game-changing-healthcare-innovation/

New Zealand Ministry of Health. (2017). *Digital health 2020*. Retrieved April 23, 2019, from https://www.health.govt.nz/our-work/ehealth/digital-health-2020

New Zealand Ministry of Health. (2019). *eHealth*. Retrieved April 23, 2019, from https://www.health.govt.nz/our-work/ehealth/digital-health-2020

Newman, M. G., Szkodny, L. E., Llera, S. J., & Przeworski, A. (2011). A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: Is human contact necessary for therapeutic efficacy? *Clinical Psychology Review*, *31*(1), 89-103. https://doi.org/10.1016/j.cpr.2010.09.008

newzoo. (2018). *Global mobile market report*. Retrieved September 15, 2019, from https://newzoo.com/solutions/standard/market-forecasts/global-mobile-market-report/

Nicholas, J., Boydell, K., & Christensen, H. (2016). mHealth in psychiatry: Time for methodological change. *Evidence-Based Mental Health*, *19*(2), 33-34. https://doi.org/10.1136/eb-2015-102278

Nielsen. (2015). *The mobile story - Nielsen mobile ratings*. IAB Australia. Retrieved December 10, 2019, from https://www.iabaustralia.com.au/research-and-resources/research-resources/item/12-research-and-resource/2013-3rd-mobile-ratings-report-september-2015

Nielsen, S. L., & Rimpilainen, S. (2018). *Report on international practice on digital apps*. Strathprints. Retrieved November 22, 2019, from https://strathprints.strath.ac.uk/66139/

One Mind. (2019). *PsyberGuide*. Retrieved October 7, 2019, from https://psyberguide.org/

OQ Measures. (2019). *OQ-Analyst*. Retrieved October 16, 2019, from http://www.oqmeasures.com/oq-analyst-3/

Orman, J., & O'Dea, B. (2018). e-Therapy in primary care mental health. *Australian Journal of General Practice*, *47*(4), 168-172. https://doi.org/10.31128/ajgp-11-17-4410

Osman, A., Wong, J. L., Bagge, C. L., Freedenthal, S., Gutierrez, P. M., & Lozano, G. (2012). The Depression Anxiety Stress Scales-21 (DASS-21): Further examination of dimensions, scale reliability and correlates. *Journal of Clinical Psychology*, *68*(12), 1322-1338. https://doi.org/10.1002/jclp.21908

Owen, J. E., Kuhn, E., Jaworski, B. K., McGee-Vincent, P., Juhasz, K., Hoffman, J. E., & Rosen, C. (2018). VA mobile apps for PTSD and related problems: Public health resources for veterans and those who care for them. *mHealth*, *4*(7), Article 28. https://doi.org/10.21037/mhealth.2018.05.07

Parker, L., Halter, V., Karliychuk, T., & Grundy, Q. (2019). How private is your mental health app data? An empirical study of mental health app privacy policies and practices. *International Journal of Law and Psychiatry*, *64*, 198-204. https://doi.org/10.1016/j.ijlp.2019.04.002

Paul, A. M., & Fleming, C. J. E. (2019). Anxiety management on campus: An evaluation of a mobile health intervention. *Journal of Technology in Behavioral Science*, *4*(1), 58-61. https://doi.org/10.1007/s41347-018-0074-2

Pew Research Center. (2019). *Who owns cellphones and smartphones*. Retrieved June 29, 2019, from https://www.pewinternet.org/fact-sheet/mobile/

Pham, Q., Wiljer, D., & Cafazzo, J. A. (2016). Beyond the randomized controlled trial: A review of alternatives in mHealth clinical trial methods. *JMIR MHealth and UHealth*, *4*(3), Article e107. https://doi.org/10.2196/mhealth.5720

Powell, J., Hamborg, T., Stallard, N., Burls, A., McSorley, J., Bennett, K., Griffiths, K. M., & Christensen, H. (2013). Effectiveness of a web-based cognitive-behavioral tool to improve mental wellbeing in the general population: Randomized controlled trial. *Journal of Medical Internet Research*, *15*(1), Article e2. https://doi.org/10.2196/jmir.2240

Proudfoot, J., Clarke, J., Birch, M.-R., Whitton, A. E., Parker, G., Manicavasagar, V., Harrison, V., Christensen, H., & Hadzi-Pavlovic, D. (2013). 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*, Article 312. https://doi.org/10.1186/1471-244X-13-312

Psychotropic Expert Groups. (2013). *Therapeutic guidelines: Psychotropic* (Vol. 7). Therapeutic Guidelines Limited.

Pung, A., Fletcher, S. L., & Gunn, J. M. (2018). Mobile app use by primary care patients to manage their depressive symptoms: Qualitative study. *Journal of Medical Internet Research*, *20*(9), Article e10035. https://doi.org/10.2196/10035:10.2196/10035

Qualtrics. (2019). *Make every interaction a great experience*. Retrieved October 16, 2019, from https://www.qualtrics.com/au/

Rahmati, A., Tossell, C., Shepard., C., Kortum, P., & Zhong, L. (2012, September 21–24). *Exploring iPhone usage: The influence of socioeconomic differences on smartphone adoption, usage and usability* [Paper presentation]. Mobile HCI '12, San Francisco, United States.

Reachout.com. (2019). *Tools and apps*. Retrieved October 7, 2019, from https://au.reachout.com/tools-and-apps

Reachout.com. (2020, May 8). *New data shows young people are reaching out for COVID-19 mental health support.* https://about.au.reachout.com/new-data-shows-young-people-are-reaching-out-for-covid-19-mental-health-support/

Research2Guidance. (2017). *mHealth Economics 2017 - Current status and future trends in mobile health*. Retrieved January 9, 2020, from https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/

Research New Zealand. (2015). *A report on a survey of New Zealanders' use of smartphones and other mobile communication devices 2015*. Retrieved January 21, 2020, from http://www.researchnz.com/pdf/Special%20Reports/Research%20New%20Zealand%20Special%20Report%20-%20Use%20of%20Smartphones.pdf

Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, *32*(4), 329-342. https://doi.org/10.1016/j.cpr.2012.02.004

Richardson, T., Stallard, P., & Velleman, S. (2010). Computerised cognitive behavioural therapy for the prevention and treatment of depression and anxiety in children and adolescents: A systematic review. *Clinical Child & Family Psychology Review*, *13*(3), 275-290. https://doi.org/10.1007/s10567-010-0069-9

Rickard, N., Arjmand, H.-A., Bakker, D., & Seabrook, E. (2016). Development of a mobile phone app to support self-monitoring of emotional well-being: A mental health digital innovation. *JMIR Mental Health*, *3*(4), Article e49. https://doi.org/10.2196/mental.6202

Rodriguez-Paras, C., Tippey, K., Brown, E., Sasangohar, F., Creech, S., Kum, H.-C., Lawley, M., & Benzer, J. (2017). Posttraumatic stress disorder and mobile health: App investigation and

scoping literature review. *JMIR mHealth and uHealth*, *5*(10), Article e156. https://doi.org/10.2196/mhealth.7318

Roepke, A. M., Jaffee, S. R., Riffle, O. M., McGonigal, J., Broome, R., & Maxwell, B. (2015). Randomized controlled trial of SuperBetter, a smartphone-based/Internet-based self-help tool to reduce depressive symptoms. *Games For Health Journal*, *4*(3), 235-246. https://doi.org/10.1089/g4h.2014.0046

Roy, M. J., Costanzo, M. E., Highland, K. B., Olsen, C., Clayborne, D., & Law, W. (2017). An app a day keeps the doctor away: Guided education and training via smartphones in subthreshold post traumatic stress disorder. *Cyberpsychology, Behavior and Social Networking*, *20*(8), 470-478. https://doi.org/10.1089/cyber.2017.0221

Royal New Zealand College of General Practitioners. (2019). *Health Navigator: App library*. Retrieved October 7, 2019, from https://www.healthnavigator.org.nz/apps/

Rubanovich, C. K., Mohr, D. C., & Schueller, S. M. (2017). Health app use among individuals with symptoms of depression and anxiety: A survey study with thematic coding. *JMIR Mental Health*, *4*(2), Article e22. https://doi.org/10.2196/mental.7603

Ruiz, F. J. (2012). Acceptance and commitment therapy versus traditional cognitive behavioral therapy: A systematic review and meta-analysis of current empirical evidence. *International Journal of Psychology & Psychological Therapy*, *12*(2), 333-357.

Sakai, C., Mackie, T. I., Shetgiri, R., Franzen, S., Partap, A., Flores, G., & Leslie, L. K. (2014). Mental health beliefs and barriers to accessing mental health services in youth aging out of foster care. *Academic Pediatrics*, *14*(6), 565-573. https://doi.org/10.1016/j.acap.2014.07.003

Sander, L. B., Schorndanner, J., Terhorst, Y., Spanhel, K., Pryss, R., Baumeister, H., & Messner, E.-M. (2020). "Help for trauma from the app stores?" A systematic review and standardised rating of apps for Post-Traumatic Stress Disorder (PTSD). *European Journal of*

Psychotraumatology, *11*(1), Article 1701788. https://doi.org/10.1080/20008198.2019.1701788

Santo, K., Richtering, S. S., Chalmers, J., Thiagalingam, A., Chow, C. K., & Redfern, J. (2016). Mobile phone apps to improve medication adherence: A systematic stepwise process to identify high-quality apps. *JMIR MHealth and UHealth*, *4*(4), Article e132. https://doi.org/10.2196/mhealth.6742

Schellong, J., Lorenz, P., & Weidner, K. (2019). Proposing a standardized, step-by-step model for creating post-traumatic stress disorder (PTSD) related mobile mental health apps in a framework based on technical and medical norms. *European Journal of Psychotraumatology*, *10*, Article 1611090. https://doi.org/10.1080/20008198.2019.1611090

Schueller, S. M., Neary, M., O'Loughlin, K., & Adkins, E. C. (2018). Discovery of and interest in health apps among those with mental health needs: Survey and focus group study. *Journal of Medical Internet Research*, *20*(6), Article e10141. https://doi.org/10.2196/10141

Schulz, K. F., Altman, D. G., Moher, D., & The CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Trials*, *11*, Article 32. https://doi.org/10.1186/1745-6215-11-32

Schwartz, S., Schultz, S., Reider, A., & Saunders, E. F. H. (2016). Daily mood monitoring of symptoms using smartphones in bipolar disorder: A pilot study assessing the feasibility of ecological momentary assessment. *Journal of Affective Disorders*, *191*, 88-93. https://doi.org/10.1016/j.jad.2015.11.013

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, *55*(1), 5-14. https://doi.org/10.1037//0003-066X.55.1.5

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Emprical validation of interventions. *American Psychologist*, *60*(5), 410-421. https://doi.org/10.1037/0003-066X.60.5.410

SGS. (2020). *Furniture*. Retrieved January 7, 2020, from https://www.sgs.com/en/consumer-goods-retail/hardgoods/furniture

Shah, A., Kraemer, K. R., Won, C. R., Black, S., & Hasenbein, W. (2018). Developing digital intervention games for mental disorders: A review. *Games For Health Journal*, *7*(4), 213-224. https://doi.org/10.1089/g4h.2017.0150

Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., Porter, A. C., Tugwell, P., Moher, D., & Bouter, L. M. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, *7,* Article 10. https://doi.org/10.1186/1471-2288-7-10

Shea, B. J., Reeves, B. C., Wells, G. A., Thuku, M., Hamel, C., Moran, J., Moher, D., Tugwell, P., Welch, V., Kristjansson, E., & Henry, D. A. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, *358*, Article j4008. https://doi.org/10.1136/bmj.j4008

Shen, N., Levitan, M.-J., Johnson, A., Bender, J. L., Hamilton-Page, M., Jadad, A. R., & Wiljer, D. (2015). Finding a depression app: A review and content analysis of the depression app marketplace. *JMIR MHealth and UHealth*, *3*(1), Article e16. https://doi.org/10.2196/mhealth.3713

Sheridan, S. M. (2014). Single-case designs and large-N studies: The best of both worlds. In T. R. Kratoch & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 299-308). American Psychological Association.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1-32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Sin, N. L., & Lyubomirsky, S. (2009). Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: A practice-friendly meta-analysis. *Journal of Clinical Psychology*, *65*(5), 467-487. https://doi.org/10.1002/jclp.20593

Sinclair, C., Holloway, K., Riley, G., & Auret, K. (2013). Online mental health resources in rural Australia: Clinician perceptions of acceptability. *Journal of Medical Internet Research*, *15*(9), Article e193. https://doi.org/10.2196/jmir.2772

Singal, A. G., Higgins, P. D. R., & Waljee, A. K. (2014). A primer on effectiveness and efficacy trials. *Clinical and Translational Gastroenterology*, *5*(1), Article e45. https://doi.org/http://dx.doi.org/10.1038/ctg.2013.13

Sloan, D. M., Gallagher, M. W., Feinstein, B. A., Lee, D. J., & Pruneau, G. M. (2011). Efficacy of telehealth treatments for posttraumatic stress-related symptoms: A meta-analysis. *Cognitive Behaviour Therapy*, *40*(2), 111-125. https://doi.org/10.1080/16506073.2010.550058

Smiling Mind. (2019). *Smiling Mind: Help shape our future*. Retrieved October 10, 2019, from https://www.smilingmind.com.au/

Smith, A. (2015). *Chapter one: A portrait of smartphone ownership*. Pew Research Center. Retrieved January 17, 2019, from http://www.pewinternet.org/2015/04/01/chapter-one-a-portrait-of-smartphone-ownership/

Sprenger, M., Mettler, T., & Osma, J. (2017). Health professionals' perspective on the promotion of e-mental health apps in the context of maternal depression. *PloS One*, *12*(7), Article e0180867. https://doi.org/doi.org/10.1371/journal.pone.0180867

Statista. (2019a). *Number of apps available in leading app stores as of 1st quarter 2019*. Retrieved July 12, 2019, from https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/

Statista. (2019b). *Number of smartphone users worldwide from 2016 to 2021 (in billions)*. Retrieved October 19, 2019, from www.statista.com/statistics/330695/numberof-smartphone-users-worldwide/

Statista. (2020). *Leading health and fitness apps in the Google Play store worldwide in March 2020, by revenue*. Retrieved April 11, 2020, from https://www.statista.com/statistics/695697/top-android-health-apps-in-google-play-by-revenue/

Stats NZ. (2017). *New Zealand internet is going mobile*. Retrieved April 23, 2019, from http://archive.stats.govt.nz/browse_for_stats/industry_sectors/information_technology_and_communications/isp-2017-mobile-connections-story.aspx

Stawarz, K., Preist, C., Tallon, D., Wiles, N., & Coyle, D. (2018). User experience of cognitive behavioral therapy apps for depression: An analysis of app functionality and user reviews. *Journal of Medical Internet Research*, *20*(6), Article e10120. https://doi.org/10.2196/10120

Stiles-Shields, C., Montague, E., Kwasny, M. J., & Mohr, D. C. (2019). Behavioral and cognitive intervention strategies delivered via coached apps for depression: Pilot trial. *Psychological Services*, *16*(2), 233-238. https://doi.org/10.1037/ser0000261

Stoyanov, S. R., Hides, L., Kavanagh, D. J., & Wilson, H. (2016). Development and validation of the user version of the Mobile Application Rating Scale (uMARS). *JMIR MHealth and UHealth*, *4*(2), Article e72. https://doi.org/10.2196/mhealth.5849

Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile app rating scale: A new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth*, *3*(1), Article e27. https://doi.org/10.2196/mhealth.3422

Stress Refuge Inc. (2015). *Introducing DeStressify*. Retrieved October 10, 2019, from https://www.destressify.com/

Sunyaev, A., Dehling, T., Taylor, P. L., & Mandl, K. D. (2015). Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*, *22*(e1), e28-e33. https://doi.org/10.1136/amiajnl-2013-002605

SuperBetter LLC. (2019). *SuperBetter: Everyone has heroic potential*. Retrieved October 10, 2019, from https://www.superbetter.com

Tang, J. (2015). *Family socioeconomic status and personal media technology use* [Master's thesis, Middle Tennessee State University]. https://jewlscholar.mtsu.edu/bitstream/handle/mtsu/4470/Tang_mtsu_0170N_10373.pdf?sequence=1&isAllowed=y

Tang, T. Z., & DeRubeis, R. J. (1999). Sudden gains and critical sessions in cognitive-behavioral therapy for depression. *Journal of Consulting and Clinical Psychology*, *67*(6), 894-904. https://doi.org/10.1037/0022-006X.67.6.894

Tanner, B. (2012). Validity of global physical and emotional SUDS. *Applied Psychophysiology and Biofeedback*, *37*(1), 31-34. https://doi.org/10.1007/s10484-011-9174-x

Terry, N. P., & Gunter, T. D. (2018). Regulating mobile mental health apps. *Behavioral Sciences and the Law*, *36*(2), 136-144. https://doi.org/10.1002/bsl.2339

Thoma, N., Pilecki, B., & McKay, D. (2015). Contemporary cognitive behavior therapy: A review of theory, history, and evidence. *Psychodynamic Psychiatry*, *43*(3), 423-462. https://doi.org/10.1521/pdps.2015.43.3.423

Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A. F., Keshavan, M., Linardon, J., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, chat bots, and virtual reality. *World Psychiatry, 20*(3), 318-335. https://doi.org/10.1002/wps.20883

Torous, J., Chan, S., Luo, J., Boland, R., & Hilty, D. (2018). Clinical informatics in psychiatric training: Preparing today's trainees for the already present future. *Academic Psychiatry*, *42*, 694-697. https://doi.org/10.1007/s40596-017-0811-4

Torous, J., & Firth, J. (2016). The digital placebo effect: mobile mental health meets clinical psychiatry. *Lancet Psychiatry*, *3*(2), 100-102. https://doi.org/10.1016/S2215-0366(15)00565-9

Torous, J., Firth, J., Huckvale, K., Larsen, M. E., Cosco, T. D., Carney, R., Chan, S. R., Pratap, A., Yellowlees, P., Wykes, T., Keshaven, M., & Christensen, H. (2018). The emerging imperative for a consensus approach toward the rating and clinical recommendation of mental health apps. *The Journal of Mental and Nervous Disease*, *206*(8), 662-666. https://doi.org/10.1097/NMD.0000000000000864

Torous, J., Levin, M. E., Ahern, D. K., & Oser, M. L. (2017). Cognitive behavioral mobile applications: Clinical studies, marketplace overview, and research agenda. *Cognitive and Behavioral Practice*, *24*(2), 215-225. https://doi.org/10.1016/j.cbpra.2016.05.007

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

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

University of New England. (2019). *Management and storage of research data and materials policy*. Retrieved October 13, 2019, from https://policies.une.edu.au/view.current.php?id=00208

Urbaniak, G. C., & Plous, S. (2019). *Research Randomizer*. Retrieved October 14, 2019, from https://www.randomizer.org/

Uryutin, O. (2018, September 13). *A brief history of web app.* Oleg Uryutin. https://aplextor.medium.com/a-brief-history-of-web-app-50d188f30d

U.S. Food and Drug Administration. (2019). *Device software functions including mobile medical applications*. Retrieved April 11, 2020, from https://www.fda.gov/medical-devices/digital-health/device-software-functions-including-mobile-medical-applications

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215-247. https://doi.org/10.3102/1076998609346961

Vigo, D., Thornicroft, G., & Atun, R. (2016). Estimating the true global burden of mental illness. *Lancet Psychiatry*, *3*(2), 171-178. https://doi.org/10.1016/S2215-0366(15)00505-2

Vollestad, J., Nielsen, M. B., & Nielsen, G. H. (2012). Mindfulness- and acceptance-based interventions for anxiety disorders: A systematic review and meta-analysis. *British Journal of Clinical Psychology*, *51*(53), 239-260. https://doi.org/10.1111/j.2044-8260.2011.02024.x

Waller, R., & Gilbody, S. (2009). Barriers to the uptake of computerized cognitive behavioural therapy: A systematic review of the quantitative and qualitative evidence. *Psychological Medicine*, *39*(5), 705-712. https://doi.org/10.1017/S0033291708004224

Wang, K., Varma, D. S., & Prosperi, M. (2018). A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. *Journal of Psychiatric Research*, *107*, 73-78. https://doi.org/10.1016/j.jpsychires.2018.10.006

Wangberg, S. C., Gammon, D., & Spitznogle, K. (2007). In the eye of the beholder: Exploring psychologists' attitudes towards and use of e-therapy in Norway. *CyberPsychology & Behavior*, *10*(3), 418-423. https://doi.org/10.1089/cpb.2006.9937

Watson, B., Proctor, S., & Cochrane, W. (2004). Using randomised controlled trials (RCTs) to test service interventions: Issues of standardisation, selection and generalisability. *Nurse Researcher*, *11*(3), 28-42. https://doi.org/10.7748/nr2004.04.11.3.28.c6203

Webb, M. S., Rodriguez de Ybarra, D., Baker, E. A., Reis, I. M., & Carey, M. P. (2010). Cognitive-behavioral therapy to promote smoking cessation among African American smokers: A

randomized clinical trial. *Journal of Consulting and Clinical Psychology*, *78*(1), 24-33. https://doi.org/10.1037/a0017669

Wilhelm, S., Weingarden, H., Ladis, I., Braddick, V., Shin, J., & Jacobson, N. C. (2020). Cognitive-behavioral therapy in the digital age: Presidential address. *Behavior Therapy*, *51*(1), 1-14. https://doi.org/10.1016/j.beth.2019.08.001

Wolpe, J. (1969). *The Practice of Behavior Therapy*. Pergamon Press.

Wolpe, J., & Lazarus, A. A. (1966). *Behavior therapy techniques: A guide to the treatment of neuroses*. Pergamon Press.

World Health Organization. (2012). *National eHealth strategy toolkit*. Retrieved November 2, 2019, from https://apps.who.int/iris/bitstream/handle/10665/75211/9789241548465_eng.pdf;jsessionid=17E8CA86DDC642233E56E8AEB441A2B0?sequence=1

World Health Organization. (2013). *Mental health action plan 2013-2020*. Retrieved November 2, 2019, from https://www.who.int/mental_health/action_plan_2013/bw_version.pdf?ua=1

World Health Organization. (2016). *eHealth*. Retrieved November 30, 2019, from http://www.who.int/topics/ehealth/en/

World Health Organization. (2018). *Depression*. Retrieved August 10, 2019, from https://www.who.int/news-room/fact-sheets/detail/depression

Worthen-Chaudhari, L., McGonigal, J., Logan, K., Bockbrader, M., A, Yeats, K. O., & Mysiw, W. J. (2017). Reducing concussion symptoms among teenage youth: Evaluation of a mobile health app. *Brain Injury*, *31*(10), 1279-1286. https://doi.org/10.1080/02699052.2017.1332388

Wright, S. (2020, June 2). *"Panic index" shows Australians were the world's best panic buyers*. The Sydney Morning Herald. https://www.smh.com.au/politics/federal/panic-index-shows-australians-were-the-world-s-best-panic-buyers-20200602-p54ync.html

Xie, Z., & Zhu, S. (2015, June). *Appwatcher: Unveiling the underground market of trading mobile app reviews* [Conference session]. 8th ACM Conference on Security and Privacy in Wireless and Mobile Networks, New York, NY, United States. https://dl.acm.org/citation.cfm?id=2766510

Young, N. S., Ioannidis, J. P. A., & Al-Ubaydli, O. (2008). Why current publication practices may distort science. *PLoS Medicine*, *5*(10), Article e201. https://doi.org/10.1371/journal.pmed.0050201

Young, A. S., Klap, R., Sherbourne, C. D., & Wells, K. B. (2001). The quality of care for depressive and anxiety disorders in the United States. *Archives of General Psychiatry*, *58*, 55-61. https://doi.org/10.1001/archpsyc.58.1.55

Zelmer, J., van Hoof, K., Notarianni, M. A., van Mierlo, T., Schellenberg, M., & Tannenbaum, C. (2018). An assessment framework for e-Mental health apps in Canada: Results of a modified Delphi process. *JMIR MHealth and UHealth*, *6*(7), Article e10016. https://doi.org/10.2196/10016

Zhou, S.-G., Hou, Y.-F., Liu, D., & Zhang, X.-Y. (2017). Effect of cognitive behavioral therapy versus interpersonal psychotherapy in patients with major depressive disorder: A meta-analysis of randomized controlled trials. *Chinese Medical Journal (Engl.)*, *130*(23), 2844-2851. https://doi.org/10.4103/0366-6999.219149

Zhou, X., Harding, L. E., Edirippulige, S., & Smith, A. C. (2020). The role of telehealth in reducing the mental health burden from COVID-19. *Telemedicine and E-Health*, *26*(4), 1-3. https://doi.org/10.1089/tmj.2020.0068

Appendix A: Supplementary Data from Main Study (Chapter 10 – Study 4)

Table A.1

Classification of Participant Improvements Across Daily Distress, Symptomatology, and Life Functioning

Participant	SUDS	DASS-42			OQ-45.2			
		Depression	Anxiety	Stress	Total Score	Symptom Distress	Interpersonal Relations	Social Role
A1	Mod	High	High	Mod	High	High	Less	Less
A2	High	Mod	Mod	Mod	High	Mod	Mod	High
A3	Less	High	High	Less	High	High	Less	Less
A4	Less	Less	Mod	Less	Less	Less	Mod	Less
A5	Mod	Mod	Less	Mod	Mod	Less	Mod	Less
B1	High	High	High	High	High	High	Less	Mod
B2	High	High	Mod	High	High	High	Less	Less
B3	Mod	High	High	High	High	High	Mod	Mod
B4	High	Mod	Mod	Less	Mod	Less	Less	Mod
B5	High	High	High	High	High	High	High	High
B6	Less	Less	Mod	Mod	Mod	Less	Less	Less
B7	High	Mod	High	High	High	Mod	Less	Less
C1	High	Mod	Less	Mod	High	High	Mod	Mod
C2	Mod	Less	Mod	Less	Mod	Mod	Mod	Less
C3	Less	Less	Less	Less	Less	Less	Less	Less
C4	Mod	Less	Less	Less	High	Mod	Less	Less

C5	High	Mod	Mod	Mod	High	High	Mod	Mod
C6	Less	Less	Mod	Less	High	High	Less	Less
D1	Less	Less	Less	Less	Less	Less	Less	Less
D2	Mod	Less	Less	Less	High	Less	Less	Mod
D3	High	Mod	Mod	Mod	High	High	Mod	Mod
D4	Less	Less	Mod	Less	Less	Mod	Less	Less
D5	Less	Less	Less	Less	Mod	Mod	Less	Less
E1	Less	Less	Mod	Mod	Less	Mod	Less	Less
E2	Less	Mod	Less	Less	Mod	Mod	Less	Less
E3	Less	Less	Mod	Mod	Mod	Mod	Less	Less
E4	High	Less	Mod	High	Mod	Mod	Mod	Mod
E5	Less	Less	Mod	Less	Mod	Less	Less	Less
E6	High	Less	Mod	Mod	High	High	Less	Less

Note. High = highly effective, Mod = moderately effective, and Less = less effective. Classification of each participant into an effectiveness group is based on criteria set out in Table A.1 above and author consensus after visually studying the various Tables and Figures of daily distress (SUDS), symptomatology (DASS), life functioning (OQ-45), and clinical significance and statistically reliable improvements.

Table A.2

Participant App Ratings From the uMARS Questionnaire

Participant	Section A – Engagement (out of 25)	Section B – Functionality (out of 20)	Section C – Aesthetics (out of 15)	Section D – Information (out of 20)	Section E – App subjective quality (out of 20)	Section F – Perceived impact (out of 30)	TOTAL (out of 130)	STAR RATING (out of 5)
A1	18	17	9	9	6	14	73	2
A2	15	13	14	13	12	30	97	3
A3	16	14	8	17	11	21	87	3
A4	17	12	11	12	7	15	74	3
A5	11	12	9	13	7	15	67	2
SuperBetter M	15.40	13.60	10.20	12.80	8.60	19.00	79.60	2.60
SuperBetter SD	2.70	2.07	2.39	2.86	2.70	6.75	12.16	0.55
B1	21	18	13	17	15	22	106	4
B2	19	17	12	17	18	28	111	5
B3	16	17	12	16	13	23	97	4
B4	10	13	8	11	5	12	59	1

B5	19	16	12	18	14	22	101	3
B6	17	16	12	15	17	25	102	5
B7	22	19	15	19	20	29	124	5
Smiling Mind M	17.71	16.57	12.00	16.14	14.57	23.00	100.00	3.86
Smiling Mind SD	3.99	1.90	2.08	2.61	4.86	5.60	20.12	1.46
C1	19	16	14	18	13	24	104	4
C2	20	17	13	17	16	25	108	4
C3	14	12	9	17	7	22	81	3
C4	22	16	14	18	13	25	108	4
C5	22	16	13	17	13	25	106	3
C6	14	13	10	15	10	20	82	3
MoodMission M	18.50	15.00	12.17	17.00	12.00	23.50	98.17	3.50
MoodMission SD	3.67	2.00	2.14	1.10	3.10	2.07	13.00	0.55
D1	17	16	13	18	16	25	105	4
D2	11	12	9	14	6	18	70	2
D3	16	13	11	17	11	24	92	3

D4	16	15	11	18	8	23	91	3
D5	13	11	8	13	5	10	60	1
MindShift M	14.60	13.40	10.40	16.00	9.20	20.00	83.60	2.60
MindShift SD	2.51	2.07	1.95	2.35	4.44	6.21	18.20	1.14
E1	16	16	10	15	15	22	94	4
E2	16	14	9	15	8	22	84	3
E3	15	16	8	13	8	17	77	3
E4	15	15	9	16	7	18	80	3
E5	10	12	7	12	5	10	56	2
E6	19	15	13	17	14	24	102	3
Destressify M	15.17	14.67	9.33	14.67	9.50	18.83	82.17	3.00
Destressify SD	2.93	1.51	2.07	1.86	4.04	5.08	15.83	0.63

Note. Participants A1 – A5 = *SuperBetter*; B1 – B7 = *Smiling Mind*; C1 – C6 = *MoodMission*; D1 – D5 = *MindShift*; and E1 – E6 = *Destressify*.

Appendix B: Information and Supplementary Data from Pilot Study (Chapter 7 – Study 3) and Other Chapters

Textbox B.1

Search Terms Used for Systematic Review (Chapter 4 – Study 1)

1. mental health app
2. mental health apps
3. depression app
4. depression apps
5. anxiety app
6. anxiety apps
7. wellbeing app
8. wellbeing apps
9. happiness app
10. happiness apps
11. psychological distress app
12. psychological distress apps
13. positive psychology app
14. positive psychology apps
15. suicide app
16. suicide apps
17. mental illness app
18. mental illness apps
19. CBT app
20. CBT apps
21. cognitive behavioural therapy app
22. cognitive behavioural therapy apps
23. cognitive behavioral therapy app
24. cognitive behavioral therapy apps
25. ACT app
26. ACT apps
27. acceptance and commitment therapy app
28. acceptance and commitment therapy apps
29. DBT app
30. DBT apps
31. dialectical behaviour therapy app
32. dialectical behaviour therapy apps
33. dialectical behavior therapy app
34. dialectical behavior therapy apps
35. IPT app
36. IPT apps
37. interpersonal therapy app
38. interpersonal therapy apps
39. e-mental health app
40. e-mental health apps
41. eMental health app
42. eMental health apps
43. mobile mental health
44. mobile mental health app
45. mobile mental health apps
46. smartphone mental health
47. smartphone mental health app
48. smartphone mental health apps
49. mHealth app
50. mHealth apps

Document B.1

Protocol for App Store Systematic Reviews (PASSR) (Chapter 5 – Study 2)

Adapted from the PRISMA Checklist by Moher et al. (2009) and AMSTAR 2 Checklist by Shae et al. (2017)

Section 1: Conducting an App Store Systematic Review
(adapted from PRISMA)

TITLE

1. Identify the report as an app store systematic review (and meta-analysis if applicable) in the title.

ABSTRACT

2. Provide a structured Abstract confirming that the review utilized the PASSR framework, and, as applicable: background, objectives, app store sources, study eligibility criteria, app appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, and systematic review registration number.

INTRODUCTION

3. In the Introduction, describe the rationale for the app store review in the context of what is already known.

4. Provide an explicit statement of questions being addressed with reference to an adapted PICOS model for app store reviews (all of these may not be relevant to every app store review):
a) Population of the group of apps being examined; what do the apps have in common?
b) What type of interventions / activities are being used by the app?
c) What types of comparison apps are being analyzed (if appropriate)?
d) What are the stated outcomes of the apps being analyzed?
e) What type of systematic, theoretical framework is being used by the app?

METHODS

5. Indicate if a review protocol has been published or is otherwise publicly accessible, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.

6. Specify app characteristics; refer to adapted PICOS model for app store reviews again if appropriate; and report other characteristics (e.g., app availability, app cost, app user ratings, app language, app development information) used as criteria for eligibility, giving rationale.

7. Describe all information sources (e.g., app store types with dates of coverage, contact with app developers to identify research [particularly published research in peer-reviewed forums] and confirmation of developers' associations and credentials) in the search, and dates of the search, including start and end dates.

8. Present full electronic search strategy for at least one app store, including any limits used (as far as the app store search options will allow), such that it could be repeated. Note any differences in the search options available between different app stores.

9. State the process for selecting apps (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis). This should state whether apps were reviewed based on their store descriptions, or if apps were individually downloaded and verified through actual use.

10. Describe method of data extraction from apps (e.g., more than one investigator duplicating the process) and any processes for obtaining and confirming data from investigators.

11. List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.

12. Describe methods used for assessing risk of bias of individual apps (e.g., has the research on efficacy and effectiveness been completed by authors with an association to the app), and how this information is to be used in any data synthesis or critical analysis.

13. State the principal summary measures and descriptive statistics, including the limits of such statistics.

14. Describe the methods of handling data and combining results of studies, if relevant. If conducting a meta-analysis, include an assessment of consistency.

15. Specify any assessment of risk of bias that may affect the cumulative research evidence that is located after conducting the app store review (e.g., publication bias, selective reporting within studies, independence / non-independence of research authors etc.).

16. Describe methods or aspects of additional analyses that may be unique to the present app store review, if appropriate, indicating why they are appropriate and stating their purpose.

RESULTS

17. Give numbers of apps screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.

18. For each app, present characteristics for which data were extracted (e.g., targeted purpose of each app, developer characteristics, theoretical framework of the app, PICOS etc.). Provide a list, in a Table within the main body of the text if the list is small, or in an Appendix or accessible online format if the list is large.

19. Present data on risk of bias of each app and, if available, any outcome level assessment (see item 12).

20. In identifying literature for evidence claimed by an app, provide summary data for intervention groups proportionate to the aims of the app store review (e.g., demographic statistics, effect estimates and confidence intervals, etc.). If appropriate, consider using visual summary techniques, such as forest plots.

21. If relevant, present results of each meta-analysis done, including confidence intervals and measures of consistency.

22. Present results of any assessment of risk of bias across apps, and/or across any research that has been found relating to the final list of apps in the review (see Item 15).

23. If appropriate, give results of additional analyses (see Item 16).

DISCUSSION

24. Summarize the main findings including the strength of evidence for each app. Consider the relevance of findings in relation to the industry / sector that this group of apps belongs, being as precise and focused as possible. If appropriate, rather than stating the implications of findings in a general sense (e.g., for the “health industry”), narrow the focus (e.g., for the “treatment of diabetes”).

25. Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., limitations of search options in app stores, reporting bias). Authors of the review should provide a disclosure of their role in any developmental capacity or association with any apps within the group of apps being reviewed.

26. Provide a general interpretation of the results in the context of other evidence, and implications for future research.

FUNDING

27. Describe sources of funding for the app store review and other support (e.g., supply of data); role of funders for the systematic review (e.g., did the funders specify a preferred publication route).

Section 2: Evaluating an App Store Systematic Review

(adapted from AMSTAR 2)

1. Did the research questions and inclusion criteria for the app store review include the components of (adapted PICOS model for app store reviews):
a) Population of the group of apps being examined; what do the apps have in common?
b) What type of interventions / activities are being used by the app?
c) What types of comparison apps are being analyzed (if appropriate)?
d) What are the stated outcomes of the apps being analyzed?
e) What type of systematic, theoretical framework is being used by the app?

2. Did the report of the app store review contain an explicit statement that the review methods were established prior to the conduct of the review (for example, by having its protocol published), and did the report justify any significant deviations from the protocol?

3. Did the app store review authors explain their selection of the app designs for inclusion in the review?

4. Did the app store review authors use a comprehensive search strategy for their app store search?

5. Did the app store review use multiple reviewers to select relevant apps, and use an adequate consensus strategy in shortlisting apps?

6. Was data extraction adequate for meeting the app store review's aims, and did more than one reviewer duplicate the data extraction?

7. Did the app store review authors provide a list of excluded apps, or make a statement that such a list was available upon request? Were exclusions justified?

8. Did the app store review authors describe the included apps in adequate detail?

9. Did the app store review authors define the parameters of their search, including dates of their search, which app stores were searched, and in which country they performed their search?

10. Did the app store review authors report on the sources of funding for their review studies?

11. Did the app store review authors use appropriate methods for statistical analysis of results in line with the aims of their study?

12. Did the app store review authors adequately report on or make a statement about the sources of each health app's development (if relevant to the aims of the study).

13. Did the app store review authors list the key words that were used for each search so that the same search can be duplicated by readers of the study, as well as providing a rationale for using these key words?

14. Did the app store review include the names of the shortlisted apps (in the case where a small number of apps were shortlisted), or a statement from the authors agreeing to provide this information upon request (in the case where many apps were shortlisted whose names could not be included due to limitations of journal space)?

15. Did the app store review authors make a statement about claims of published research evidence by shortlisted apps? Did the authors confirm the existence of such claimed research by doing a literature search?

16. Did the app store review authors report any potential sources of conflict of interest, including if they are involved in the development of health apps, or if they stand to gain financially or otherwise from the proceeds or success of a health app?

Figure B.1

Advertisement for Participants

Figure B.2

Participant App Information Sheet for SuperBetter, *Page 1*

APP INFORMATION

SuperBetter

Background

- SuperBetter was developed in the United States by game designer Jane McGonigal, with assistance from researchers at Stanford University and the University of Pennsylvania. She gives a talk that provides some background on her approach at: https://www.youtube.com/watch?v=irsTFdCtcuQ.
- More information about SuperBetter can be found on their website: https://www.superbetter.com/.
- Research on this app has been published: Roepke, A. M., Jaffee, S. R., Riffle, O. M., McGonigal, J., Broome, R., & Maxwell, B. (2015). Randomized controlled trial of SuperBetter, a smartphone-based/Internet-based self-help tool to reduce depressive symptoms. *Games for Health Journal,* 4(3), 235-246.
- SuperBetter uses evidence-based techniques from positive psychology and neuroscience (neuroplasticity).
- More information about positive psychology can be found here: https://www.blackdoginstitute.org.au/docs/default-source/factsheets/positivepsychology.pdf.
- More information about neuroscience, neuroplasticity, and mental health can be found here: https://positivepsychology.com/neuroplasticity/.

How To Use SuperBetter

- Download the app:
  from either the **Apple App Store** if you have an iPhone
  https://apps.apple.com/au/app/superbetter/id536634968.

  or **Google Play** if you have an Android phone
  https://play.google.com/store/apps/details?id=com.superbetter.paid&hl=en.
- You will need to create an account by providing an e-mail address and password. Follow the instructions when you open the app for the first time after downloading it, and take the time to answer the questions – it helps the app choose appropriate activities for you. If you're concerned about your data, read the SuperBetter privacy policy at https://www.superbetter.com/terms (note that you have to scroll down about halfway to get to the privacy section).
- A good place to start with SuperBetter is to tap on the "More" button at the bottom right of screen, then tap on "FAQ", and read the entire FAQ section. It provides an excellent background to how SuperBetter works and will get you ready to start "playing" immediately.

Figure B.3

Participant App Information Sheet for SuperBetter, *Page 2*

- SuperBetter uses the idea of playing a video game and works by getting you to complete "quests". Quests are activities designed to get you closer to achieving your specific goals.
- "Power-ups" are smaller activities designed to make you feel better and improve your mood quickly. Part of the game of SuperBetter is to complete as many power-ups and quests as you can.
- SuperBetter also has "bad guys" and "allies". Bad guys are things that get in the way of achieving your goals, such as anxiety or depression. Allies are people who support you on your journey, such as a therapist, family member, or friend. You can send details about "epic wins" and other achievements to your allies, and they can also send you messages of support.
- After you have read the FAQ section, we recommend that you tap on the "*SB To do*" button at the bottom centre of screen, then tap the "*Do 3 Quests*" banner, then tap the "*Hero, Start Here!*" button. Follow the instructions about setting up your first quest.
- We would like you to use this app for at least 10 minutes per day, 5 days per week for at least 10 weeks. (We will tell you when the "official" use period has finished, but you will be able to keep using the app after this time if you want to.) If you want to use it for longer than 10 minutes per day and more than 5 days per week, this is OK – it's up to you. As long as you're using it for ***at least*** 10 minutes per day, 5 days per week.
- We will be asking you during the 10-week period how long you have approximately been using the app for.
- We will occasionally e-mail you links to some simple online questionnaires to fill out. These will mostly take less than 10 minutes. It would be appreciated if you could complete these as soon as possible after they arrive in your inbox.
- And don't forget to reply to our daily text messages. We will ask you how you're feeling that day, and you simply have to reply with a number from 0 to 10, with **0 indicating you are completely happy** and **10 indicating that you are in the worst mood possible**. At the end of the study, we will add up how many text message replies you have given us and we will give you $0.50 for each one we received.

If You Need Help

- In a mental health emergency, phone 000 or present at your local hospital emergency department.
- Lifeline: https://www.lifeline.org.au/ or **13 11 14**
- Beyond Blue: https://beyondblue.org.au/ or **1300 224 636**
- If you have any queries or concerns about the research project, please contact Jamie Marshall, Clinical Psychologist, by e-mail (jmarsh21@myune.edu.au) or phone [redacted]

Figure B.4

Screenshots of the SuperBetter *App*

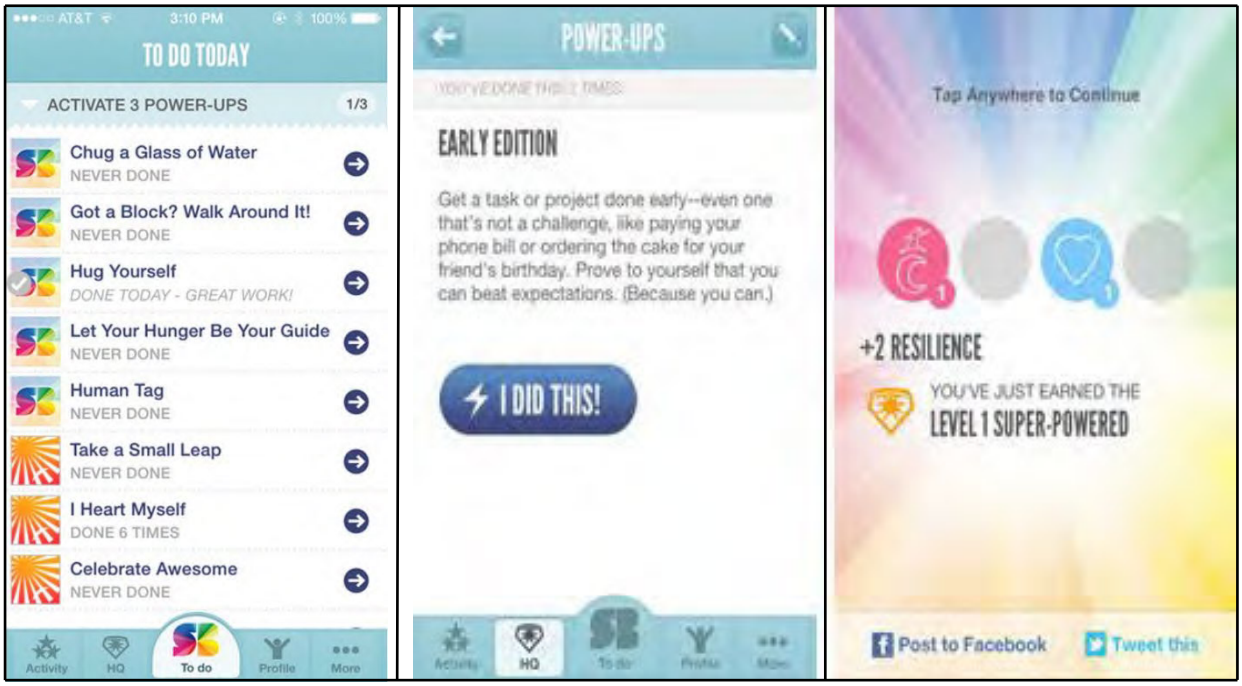
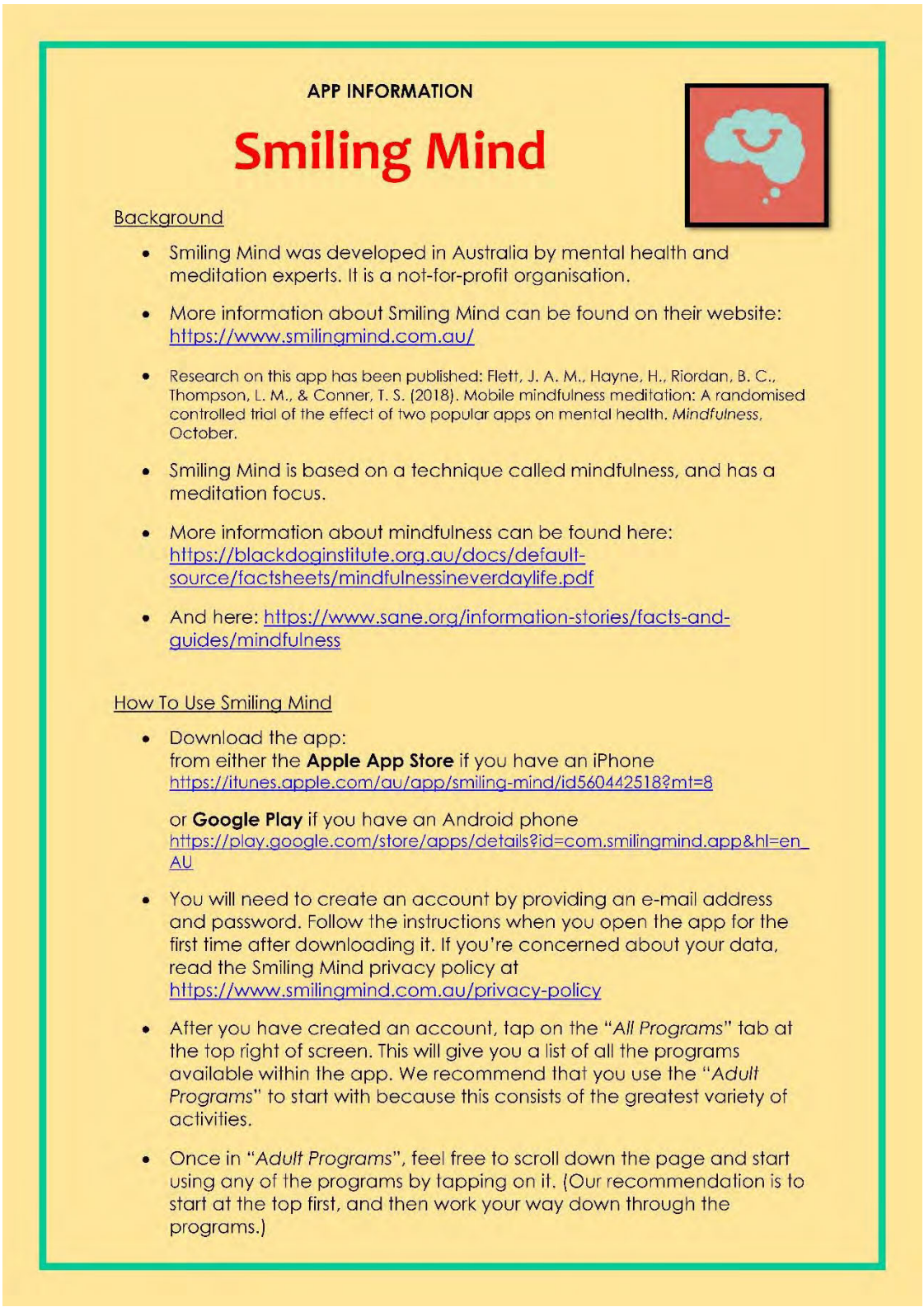


Figure B.5

Participant App Information Sheet for Smiling Mind*, Page 1*



APP INFORMATION

Smiling Mind

Background

- Smiling Mind was developed in Australia by mental health and meditation experts. It is a not-for-profit organisation.
- More information about Smiling Mind can be found on their website: https://www.smilingmind.com.au/
- Research on this app has been published: Flett, J. A. M., Hayne, H., Riordan, B. C., Thompson, L. M., & Conner, T. S. (2018). Mobile mindfulness meditation: A randomised controlled trial of the effect of two popular apps on mental health. *Mindfulness*, October.
- Smiling Mind is based on a technique called mindfulness, and has a meditation focus.
- More information about mindfulness can be found here: https://blackdoginstitute.org.au/docs/default-source/factsheets/mindfulnessineverdaylife.pdf
- And here: https://www.sane.org/information-stories/facts-and-guides/mindfulness

How To Use Smiling Mind

- Download the app:
  from either the **Apple App Store** if you have an iPhone
  https://itunes.apple.com/au/app/smiling-mind/id560442518?mt=8

  or **Google Play** if you have an Android phone
  https://play.google.com/store/apps/details?id=com.smilingmind.app&hl=en_AU
- You will need to create an account by providing an e-mail address and password. Follow the instructions when you open the app for the first time after downloading it. If you're concerned about your data, read the Smiling Mind privacy policy at https://www.smilingmind.com.au/privacy-policy
- After you have created an account, tap on the "*All Programs*" tab at the top right of screen. This will give you a list of all the programs available within the app. We recommend that you use the "*Adult Programs*" to start with because this consists of the greatest variety of activities.
- Once in "*Adult Programs*", feel free to scroll down the page and start using any of the programs by tapping on it. (Our recommendation is to start at the top first, and then work your way down through the programs.)

Figure B.6

Participant App Information Sheet for Smiling Mind, *Page 2*

- You can create a list of your favourite activities by adding them to your favourites folder. You simply tap on the activity, and when it opens, tap on the love heart in the top right hand corner. If you want to access something you have saved in favourites, look along the bottom of the screen and you should see the "*Favourites*" folder towards the right.
- There is a simple dashboard that allows you to see how many sessions you have completed and how much time you have spent meditating. Just tap on the "*Dashboard*" tab at the bottom of screen.
- To select a new program to do, first tap on "*My Programs*" at the bottom left of screen, and then tap on any of the three tabs at the top of the screen that say: "*My Programs*", "*My Sessions*", or "*All Programs*".
- We would like you to use this app for at least 10 minutes per day, 5 days per week for at least 10 weeks. (We will tell you when the "official" use period has finished, but you will be able to keep using the app after this time if you want to.) If you want to use it for longer than 10 minutes per day and more than 5 days per week, this is OK – it's up to you. As long as you're using it for ***at least*** 10 minutes per day, 5 days per week.
- We will be asking you during the 10-week period how long you have approximately been using the app for.
- We will occasionally e-mail you links to some simple online questionnaires to fill out. These will mostly take less than 10 minutes. It would be appreciated if you could complete these as soon as possible after they arrive in your inbox.
- And don't forget to reply to our daily text messages. We will ask you how you're feeling that day, and you simply have to reply with a number from 0 to 10, with **0 indicating you are completely happy** and **10 indicating that you are in the worst mood possible**. At the end of the study, we will add up how many text message replies you have given us and we will give you $0.50 for each one we received.

If You Need Help

- In a mental health emergency, phone 000 or present at your local hospital emergency department.
- Lifeline: https://www.lifeline.org.au/ or **13 11 14**
- Beyond Blue: https://beyondblue.org.au/ or **1300 224 636**
- If you have any queries or concerns about the research project, please contact Jamie Marshall, Clinical Psychologist, by e-mail (jmarsh21@myune.edu.au) or phone [redacted]

Figure B.7

Screenshots of the Smiling Mind *App*

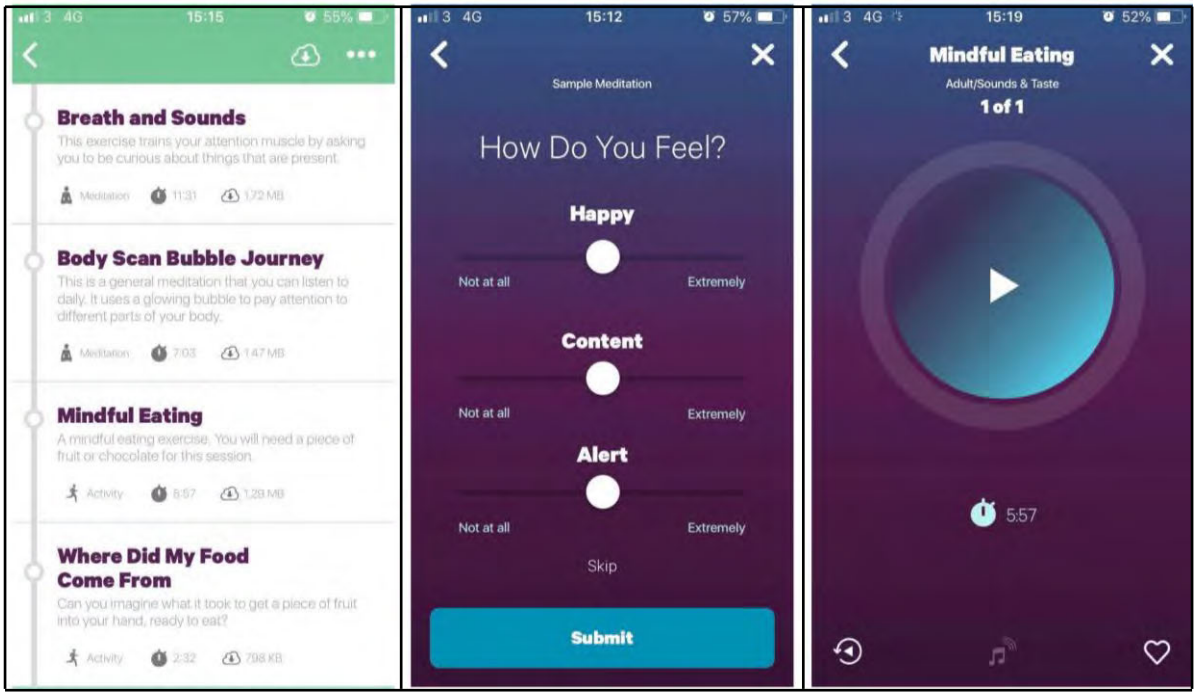


Figure B.8

Participant App Information Sheet for MoodMission*, Page 1*

APP INFORMATION

MoodMission

Background

- MoodMission was developed in Australia by a team of psychologists and mental health researchers at Monash University.
- More information about MoodMission can be found on their website: http://moodmission.com/.
- Research on this app has been published: Bakker, D., Kazantzis, N., Rickwood, D., & Rickard, N. (2018). A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behavior Research and Therapy, 109*, 75-83.
- MoodMission uses an evidence-based technique called cognitive-behaviour therapy (CBT).
- More information about CBT can be found here: https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/cognitive-behaviour-therapy.
- And here: https://au.reachout.com/articles/cognitive-behavioural-therapy.

How To Use MoodMission

- Download the app:
  from either the **Apple App Store** if you have an iPhone
  https://apps.apple.com/au/app/moodmission/id1140332763.

  or **Google Play** if you have an Android phone
  https://play.google.com/store/apps/details?id=com.moodmission.moodmissionapp.
- You will need to create an account by providing an e-mail address and password. Follow the instructions when you open the app for the first time after downloading it, and take the time to answer the questionnaire – it helps the app choose appropriate activities for you. If you're concerned about your data, read the MoodMission privacy policy at http://moodmission.com/privacy.
- MoodMission works by asking you questions about how you are feeling, and suggests all types of activities in response to how you answer the questions.
- MoodMission does not make you do just one thing, but always provides different options. If you do not like any of the listed options, you have the opportunity to ask MoodMission to suggest others.

Figure B.9

Participant App Information Sheet for MoodMission*, Page 2*

- If you are feeling acutely anxious or depressed, you can tap on "Other support options" at the bottom of the screen, and MoodMission provides emergency mental health support options you can immediately link in with.
- At the bottom left of screen is a "*Mission Log*" button. Tap on this to find a record of all your achievements, including the number of missions you have completed, and your rank. The more missions you complete, the higher your rank. You can also view other statistics here.
- At the bottom right of screen is the "*Expeditions*" button. Tap on this to find suggested programs for overcoming specific fears e.g. fear of public speaking, fear of flying etc.
- We would like you to use this app for at least 10 minutes per day, 5 days per week for at least 10 weeks. (We will tell you when the "official" use period has finished, but you will be able to keep using the app after this time if you want to.) If you want to use it for longer than 10 minutes per day and more than 5 days per week, this is OK – it's up to you. As long as you're using it for ***at least*** 10 minutes per day, 5 days per week.
- We will be asking you during the 10-week period how long you have approximately been using the app for.
- We will occasionally e-mail you links to some simple online questionnaires to fill out. These will mostly take less than 10 minutes. It would be appreciated if you could complete these as soon as possible after they arrive in your inbox.
- And don't forget to reply to our daily text messages. We will ask you how you're feeling that day, and you simply have to reply with a number from 0 to 10, with **0 indicating you are completely happy** and **10 indicating that you are in the worst mood possible**. At the end of the study, we will add up how many text message replies you have given us and we will give you $0.50 for each one we received.

If You Need Help

- In a mental health emergency, phone 000 or present at your local hospital emergency department.
- Lifeline: https://www.lifeline.org.au/ or **13 11 14**
- Beyond Blue: https://beyondblue.org.au/ or **1300 224 636**
- If you have any queries or concerns about the research project, please contact Jamie Marshall, Clinical Psychologist, by e-mail (jmarsh21@myune.edu.au) or phone [redacted]

Figure B.10

Screenshots of the MoodMission *App*

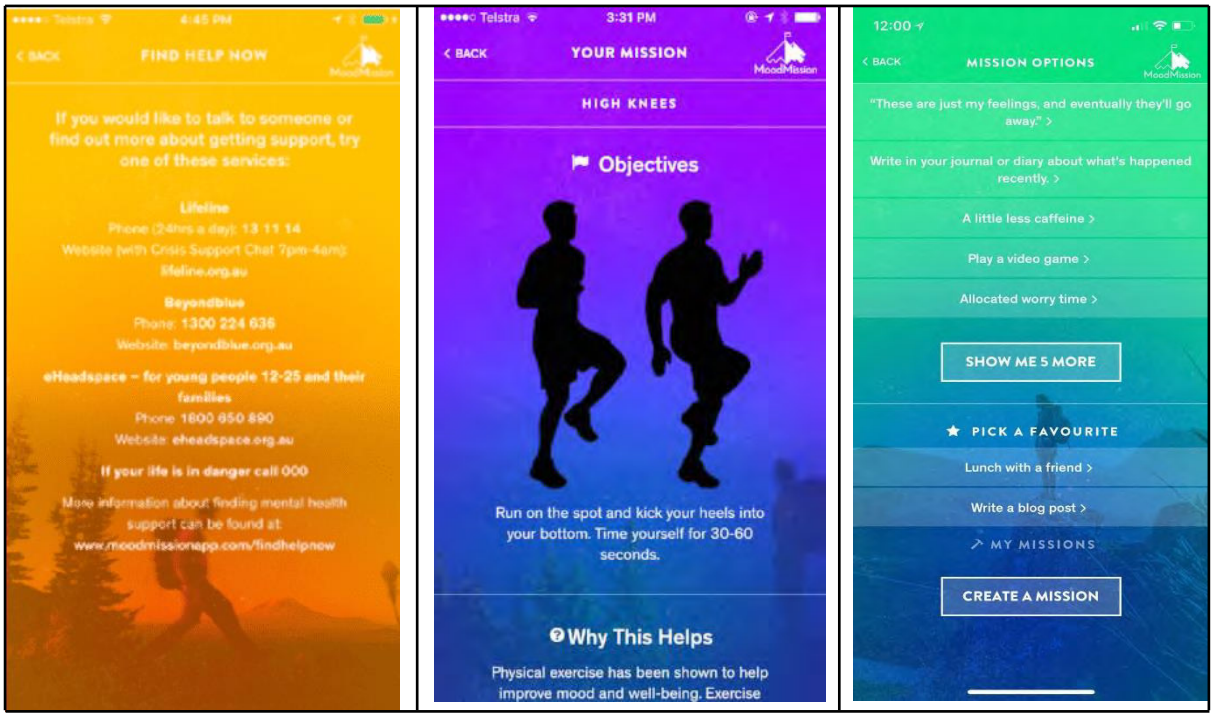


Figure B.11

Participant App Information Sheet for MindShift, *Page 1*

APP INFORMATION

MindShift CBT

Background

- MindShift was developed in Canada by a not-for-profit mental health support organisation.
- More information about MindShift can be found on their website: https://anxietycanada.com/resources/mindshift-cbt/
- Research on this app has been published: Paul, A. M., & Fleming, C. J. E. (2019). Anxiety management on campus: An evaluation of a mobile health intervention. *Journal of Technology in Behavioral Science*, 4(1), 58-61.
- MindShift uses an evidence-based technique called cognitive-behaviour therapy (CBT). Although its focus is anxiety, the techniques can be applied to depression as well.
- More information about CBT can be found here: https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/cognitive-behaviour-therapy.
- And here: https://au.reachout.com/articles/cognitive-behavioural-therapy.

How To Use MindShift

- Download the app:
  from either the **Apple App Store** if you have an iPhone
  https://apps.apple.com/au/app/mindshift-cbt-anxiety-canada/id634684825

  or **Google Play** if you have an Android phone
  https://play.google.com/store/apps/details?id=com.bstro.MindShift&hl=en
- You will need to create an account by providing an e-mail address, password, and other information. If you're concerned about your data, read the MindShift privacy policy at:
  https://anxietycanada.com/mindshift-cbt-privacy-policy/
- After you have created an account, the app opens to its home screen (which you can access at any time by tapping on the "*Home*" button at the bottom left of screen). The Home screen allows you to choose how you are feeling today, and what type of anxiety you suffer from.
- Tapping on a type of anxiety takes you to information about that type of worry. When you scroll down the anxiety screen, the app allows you to select categories within that type of anxiety which lists some of the more common unhelpful thoughts associated with it.

Figure B.12

Participant App Information Sheet for MindShift, *Page 2*

- We recommend that you start by scrolling down the home page until you come to "*Tools*". Tap on the button, "*Healthy Thinking*", then tap on "*Thought Journal*", and then tap on "*Add New +*". This will get you started using one of the main functions of the app: coming up with more helpful thoughts when you are feeling anxious and/or depressed due to thinking negatively.

- You can tap on "*Learn*" at the bottom of screen any time to read more information about anxiety and CBT. You can tap on "*Goals*" at the bottom of screen any time to set goals you would like to achieve.

- If you would like some suggestions from the app about things you can do quickly to help ease acute anxiety or depression symptoms, tap on the blue smiley face at the bottom centre of screen.

- We would like you to use this app for at least 10 minutes per day, 5 days per week for at least 10 weeks. (We will tell you when the "official" use period has finished, but you will be able to keep using the app after this time if you want to.) If you want to use it for longer than 10 minutes per day and more than 5 days per week, this is OK – it's up to you. As long as you're using it for ***at least*** 10 minutes per day, 5 days per week.

- We will be asking you during the 10-week period how long you have approximately been using the app for.

- We will occasionally e-mail you links to some simple online questionnaires to fill out. These will mostly take less than 10 minutes. It would be appreciated if you could complete these as soon as possible after they arrive in your inbox.

- And don't forget to reply to our daily text messages. We will ask you how you're feeling that day, and you simply have to reply with a number from 0 to 10, with **0 indicating you are completely happy** and **10 indicating that you are in the worst mood possible**. At the end of the study, we will add up how many text message replies you have given us and we will give you $0.50 for each one we received.

If You Need Help

- In a mental health emergency, phone 000 or present at your local hospital emergency department.

- Lifeline: https://www.lifeline.org.au/ or **13 11 14**

- Beyond Blue: https://beyondblue.org.au/ or **1300 224 636**

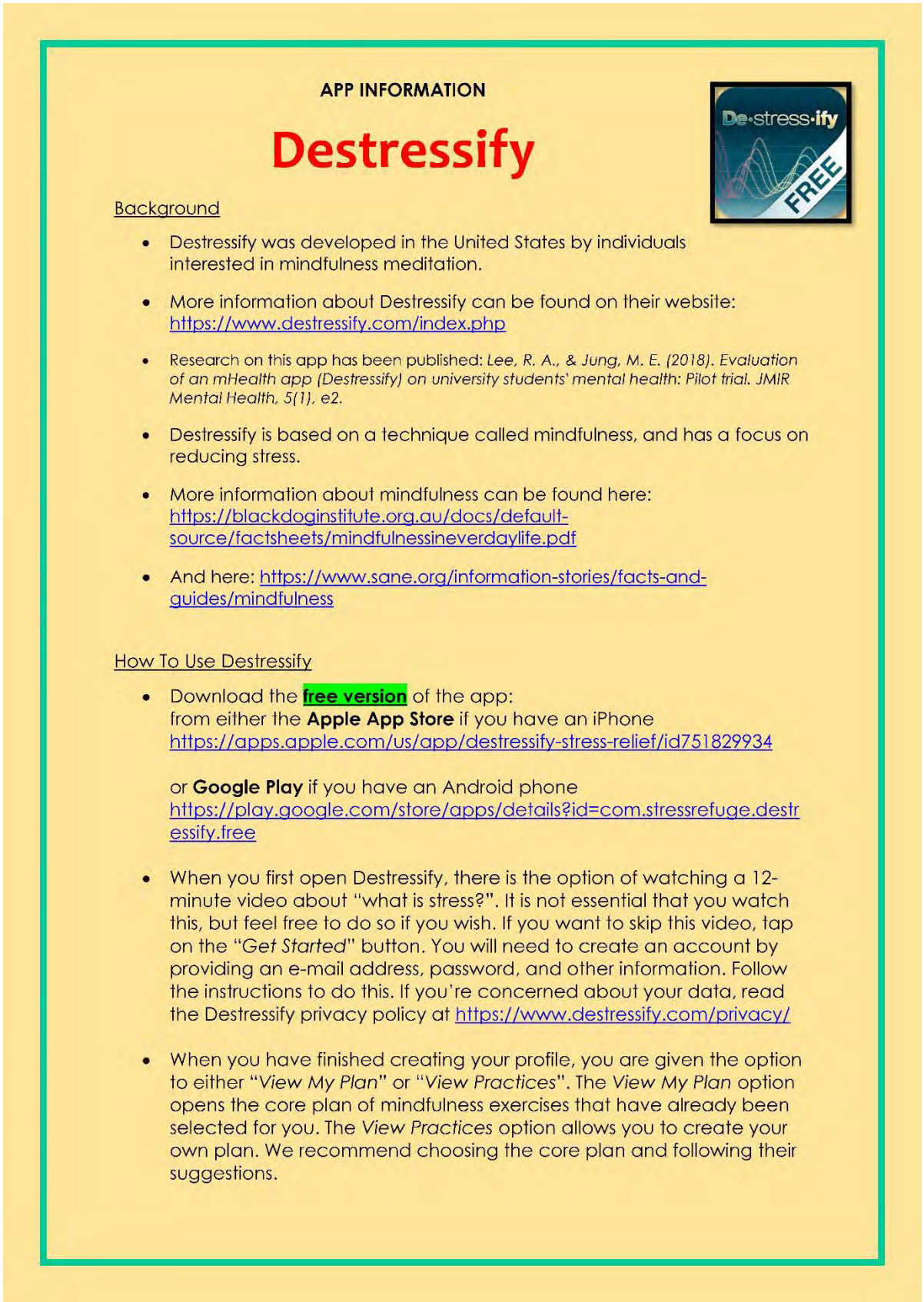
- If you have any queries or concerns about the research project, please contact Jamie Marshall, Clinical Psychologist, by e-mail (jmarsh21@myune.edu.au) or phone [redacted]

Figure B.13

Screenshots of the MindShift *App*

Figure B.14

Participant App Information Sheet for Destressify, *Page 1*



APP INFORMATION

Destressify

Background

- Destressify was developed in the United States by individuals interested in mindfulness meditation.
- More information about Destressify can be found on their website: https://www.destressify.com/index.php
- Research on this app has been published: *Lee, R. A., & Jung, M. E. (2018). Evaluation of an mHealth app (Destressify) on university students' mental health: Pilot trial. JMIR Mental Health, 5(1), e2.*
- Destressify is based on a technique called mindfulness, and has a focus on reducing stress.
- More information about mindfulness can be found here: https://blackdoginstitute.org.au/docs/default-source/factsheets/mindfulnessineverdaylife.pdf
- And here: https://www.sane.org/information-stories/facts-and-guides/mindfulness

How To Use Destressify

- Download the **free version** of the app:
  from either the **Apple App Store** if you have an iPhone
  https://apps.apple.com/us/app/destressify-stress-relief/id751829934

  or **Google Play** if you have an Android phone
  https://play.google.com/store/apps/details?id=com.stressrefuge.destressify.free
- When you first open Destressify, there is the option of watching a 12-minute video about "what is stress?". It is not essential that you watch this, but feel free to do so if you wish. If you want to skip this video, tap on the "*Get Started*" button. You will need to create an account by providing an e-mail address, password, and other information. Follow the instructions to do this. If you're concerned about your data, read the Destressify privacy policy at https://www.destressify.com/privacy/
- When you have finished creating your profile, you are given the option to either "*View My Plan*" or "*View Practices*". The *View My Plan* option opens the core plan of mindfulness exercises that have already been selected for you. The *View Practices* option allows you to create your own plan. We recommend choosing the core plan and following their suggestions.

Figure B.15

Participant App Information Sheet for Destressify, *Page 2*

- After tapping on *View My Plan*, you can immediately begin your first mindfulness activity. You can also tap on "*View/Edit Schedule*" to see the names of future activities.

- In the top right corner of screen is the Menu folder. Tap on this to access the features of the app. Not all the features are available in this free version. If you want to access all the features, you need to download Destressify Pro at a cost, but this is not required for our research – just the free version of the app is needed.

- In the Menu, tap on the button "*Learn*" to read information on mindfulness and stress.

- In the Menu, tap on the button "*Practices*" to read information on the different types of activities available in the app.

- We would like you to use this app for at least 10 minutes per day, 5 days per week for at least 10 weeks. (We will tell you when the "official" use period has finished, but you will be able to keep using the app after this time if you want to.) If you want to use it for longer than 10 minutes per day and more than 5 days per week, this is OK – it's up to you. As long as you're using it for ***at least*** 10 minutes per day, 5 days per week.

- We will be asking you during the 10-week period how long you have approximately been using the app for.

- We will occasionally e-mail you links to some simple online questionnaires to fill out. These will mostly take less than 10 minutes. It would be appreciated if you could complete these as soon as possible after they arrive in your inbox.

- And don't forget to reply to our daily text messages. We will ask you how you're feeling that day, and you simply have to reply with a number from 0 to 10, with **0 indicating you are completely happy** and **10 indicating that you are in the worst mood possible**. At the end of the study, we will add up how many text message replies you have given us and we will give you $0.50 for each one we received.

If You Need Help

- In a mental health emergency, phone 000 or present at your local hospital emergency department.

- Lifeline: https://www.lifeline.org.au/ or **13 11 14**

- Beyond Blue: https://beyondblue.org.au/ or **1300 224 636**

- If you have any queries or concerns about the research project, please contact Jamie Marshall, Clinical Psychologist, by e-mail (jmarsh21@myune.edu.au) or phone [redacted].

Figure B.16

Screenshots of the Destressify *App*

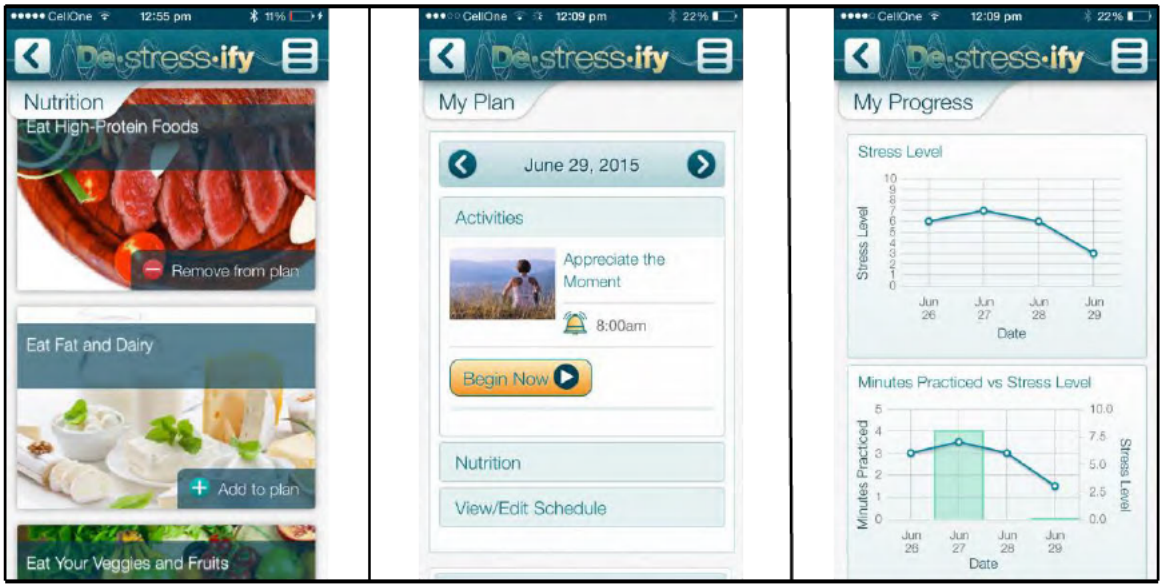


Table B.1

Characteristics Summary of Each Mental Health App in the Present Study

App name	Theoretical framework(s)	Corporate governance	Previous research	Country of origin
Destressify	Mindfulness	Private corporation	Lee & Jung (2018)	United States of America
MoodMission	CBT	Private corporation	Bakker & Rickard (2019)	Australia
Smiling Mind	Mindfulness	Not-for-profit	Flett et al. (2019)	Australia
MindShift	CBT	Not-for-profit	Paul & Fleming (2019)	Canada
SuperBetter	CBT / Positive Psychology / Neuroplasticity	Private corporation	Roepke et al. (2015)	United States of America

Table B.2

Links to Professional Independent Reviews for the Apps in the Present Study

App name	Link to PsyberGuide review	Link to other independent review
Destressify	None found	None found
MoodMission	https://psyberguide.org/apps/moodmission/	https://au.reachout.com/tools-and-apps/moodmission
Smiling Mind	https://psyberguide.org/apps/smiling-mind/	https://au.reachout.com/tools-and-apps/smiling-mind
MindShift	https://psyberguide.org/apps/mindshift/	https://www.healthnavigator.org.nz/apps/m/mindshift-app/
SuperBetter	https://psyberguide.org/apps/superbetter/	https://au.reachout.com/tools-and-apps/superbetter

Table B.3

Summary Details of Previous Research on the Apps Used in the Present Study

Article authors (a*pp name*)	Design / methodology	Sample type	Study focus	Outcome measures used
Lee & Jung (2018) (*Destressify*)	4 weeks app use – 5 days per week. Participants told to use the "Core Plan" on the app and told not to engage with other features of the app. Intervention app is based on mindfulness. Control group was waitlisted	American undergraduate students	Anxiety, depression, stress	State-Trait Anxiety Inventory; Quick Inventory of Depressive Symptomatology Self-Report; Perceived Stress Scale (PSS)
Bakker & Rickard (2019) (*MoodMission*)	30 days total app use with four push notifications sent to participants during this period reminding them to engage with the app. Intervention app is based on CBT	Australian community sample of individuals who first downloaded the app and then agreed to be in the research	Anxiety, depression	Generalized Anxiety Disorder Scale 7-Item; Patient Health Questionnaire 9-Item; Warwick Edinburgh Mental Well-Being Scale
Flett et al. (2019) (*Smiling Mind*)	40 days total app use – 10 mins per day for the first 10 days, then use the app at participant discretion for remaining 30 days. The app intervention is based on mindfulness. RCT design with control group using an app named Evernote to test for the digital placebo effect	New Zealand undergraduate students	Anxiety, depression, stress, flourishing	Hospital Anxiety and Depression Scale-Anxiety Subscale; Center for Epidemiological Studies Depression Scale (CES-D); PSS; Flourishing Scale
Paul & Fleming (2019) (*MindShift*)	3 weeks app use – minimum of 15 mins per day, 5 days per week. The app has an anxiety focus and is underpinned by CBT. Participants were only included if they were not receiving any other treatments for anxiety. Participants were given a tutorial on using the app	Canadian undergraduate students	Focus was on anxiety, but also examined depression	Patient Health Questionnaire (PHQ)-15; PHQ-Panic; Generalized Anxiety Disorder Scale-7 (GAD-7); PHQ-9

Roepke et al. (2015) (*SuperBetter*)	App use for 10 mins per day for 1 month using an RCT design with a waitlist control group. Sample was described as "motivated" individuals who already had an interest in managing their mental health with an app because they had first downloaded the app before being recruited to the trial. App is based on CBT / positive psychology / neuroplasticity	Americans who downloaded the app and self-report measure indicated significant depression symptoms	Anxiety, depression, life satisfaction	GAD-7; CES-D; Satisfaction With Life Scale

Table B.4

Summary of Statistics of Previous Research on the Apps Used in the Present Study

Article authors (*app name*)	Intervention group N	Control group N	Mean age (years)	*P*-value	Effect size
Lee & Jung (2018) (*Destressify*)	77	86	20.6	Anxiety = .02* Depression = .09 Stress = .06	η^2_p = .05 (anxiety) η^2_p = .02 (depression) η^2_p = .02 (stress)
Bakker & Rickard (2019) (*MoodMission*)	617	None	26.9	Anxiety (severe clinical) = .05* Depression (severe clinical) = .01*	β (Anxiety – severe clinical – coping self-efficacy mediator) = - .378 β (Depression – severe clinical – coping self-efficacy mediator) = - .386
Flett et al. (2019) (*Smiling Mind*)	58	67	20.1	Anxiety = < .05* Depression = > .05 Stress = > .05 Flourishing = > .05	Cohen's d = .29 (anxiety), Cohen's d = .13 (depression), Cohen's d = .18 (stress), Cohen's d = .01 (flourishing)
Paul & Fleming (2019) (*MindShift*)	16	None	19.8	Somatic anxiety = < .01* General anxiety = < .01* Panic = .99 Depression = < .01*	η = .34 (somatic anxiety) η = .51 (general anxiety) η = .00 (panic) η = .30 (depression)
Roepke et al. (2015) (*SuperBetter*)	190 (97 in regular app group, and 93 in "enhanced" app group)	93	Not given	Regular app anxiety = < .001* Enhanced app anxiety = .01* Regular app depression = < .001* Enhanced app depression = .01* Regular app life satisfaction = .007* Enhanced app life satisfaction = < .001*	Cohen's d regular app = 0.92 (depression) Cohen's d enhanced app = 0.43 (depression) (Note: Individual effect sizes for other outcomes not given in article, but it was stated that they were between 0.43 and 1.36.)

* Indicated as statistically significant in the article.

Textbox B.2

Summary of Previous Research on the Apps Used in The Present Study

Lee and Jung (2018) found statistically significant reductions in anxiety from pre- to post-treatment, but found no significant change in depression or stress; Bakker and Rickard (2019) recorded statistically significant improvements in anxiety and depression in "nonclinical" and "severe" participants (as determined by their self-report measures at study commencement), with no such improvements in "moderate" participants; Flett et al. (2019) found statistically significant reductions in anxiety symptoms, but not depression; Paul and Fleming (2019) found statistically significant reductions in somatic and generalized anxiety, and depression, but not for panic; and Roepke et al. (2015) obtained statistically significant reductions in anxiety and depression compared to waitlist.

Document B.2

Demographics Questionnaire

In a mental health emergency,
phone Lifeline on 13 11 14 or the Suicide Callback Service on 1300 659 467,
or present at the emergency department of your nearest hospital

Information for Participants

INFORMATION FOR PARTICIPANTS

I wish to invite you to participate in my research project, described below.

My name is Jamie Marshall and I am conducting this research as part of my PhD in the School of Psychology at the University of New England. My supervisors are Prof Debra Dunstan and Dr Warren Bartik.

Research Project Title: I bet there's an app for that: Using mobile mental health apps for reducing anxiety and depression
Aim of the Research: The research aims to explore the effectiveness of certain mobile mental health apps in reducing symptoms of anxiety and depression.
Survey: Following this information sheet, you will be asked to take a survey that will take less than 30 minutes to complete. Some time after you have completed and submitted the survey, you will be sent an e-mail or text message with further information about the study.
Confidentiality: Any personal details gathered in the course of the study will remain

confidential. No individual will be identified by name in any publication of the results. All names will be replaced by pseudonyms or numbers; this will ensure your anonymity. If you agree, I would like to quote some of your responses. This will also be done in a way to ensure that you are not identified.

Participation is Voluntary: Please understand that your involvement in this study is voluntary and I respect your right to stop participating in the study at any time without consequence and without needing to provide an explanation. If you choose to remain in the study for the duration of the data gathering phase, this will likely take approximately 18 weeks. During this time, you will be asked to use a mobile app as instructed. Other than completing some questionnaires at four different time points, you will also be asked to provide a daily reply text message with a number out of 10 indicating your mood for that day. For each daily text message you send in, you will receive a payment of $0.50 cents, payable at the conclusion of the study. We would also like to contact you six months after your involvement has finished with another questionnaire.

Questions: Questions contained in the surveys will concern mental health, specifically anxiety and depression, and your use of mobile apps to reduce symptoms of these issues.

Use of Information: I will use information from the surveys as part of my doctoral thesis, which I expect to complete in December 2020. Information from the surveys may also be used in academic journal articles and conference presentations before and after this date. At all times, I will safeguard your identity by presenting the information in a way that will not allow you to be identified.

Upsetting Issues: If this research raises any personal or upsetting issues, please contact Lifeline on 13 11 14, or your local Community Health Centre.

Storage of Information: All hardcopy notes or information generated by this study will be scanned, uploaded, and saved with all other electronic data and will be kept on Cloud.UNE, UNE's centrally managed cloud server accessible only by the research team. It

will also be kept on a password protected computer in the same location that only the research team will have access to.

Disposal of Information: All the data collected in this research will be kept indefinitely on UNE's centrally managed research data storage facility, Cloud.UNE. All information stored in local computer hard drives will be deleted after the conclusion of the research.

Requirements: To participate in this research, you are required to have access to a smartphone and/or tablet device, be 18 years of age or older, and have previously been diagnosed with an anxiety and/or depression condition by a qualified health professional. If your mental health condition and/or history includes psychotic episodes, this research project is not suitable for you.

Approval: This project has been approved by the Human Research Ethics Committee of the University of New England (Approval No. HE19-186, Valid to 01/11/2020).

Contact Details for Researchers:

Feel free to contact me with any questions about this research by email at **jmarsh21@myune.edu.au** or by phone during business hours on **0412 575 185**.

You may also contact my supervisors.

My Principal supervisor's name is Prof Debra Dunstan and she can be contacted by email at **hospsych@une.edu.au** or by phone on **(02) 6773 3012**.

My Co-supervisor's name is Dr Warren Bartik and he can be contacted by email at **wbartik@une.edu.au** or by phone on **(02) 6773 3743**.

Complaints: Should you have any complaints concerning the manner in which this research is conducted, please contact:

Mrs Jo-Ann Sozou

Research Ethics Officer

Research Services

University of New England, Armidale, NSW 2351

Tel: (02) 6773 3449 Email: humanethics@une.edu.au

Thank you for considering this request and I look forward to further contact with you.

Regards,

JAMIE MARSHALL

Clinical Psychologist & PhD Candidate

M.Psych. (clinical), B.A. (Hons.) (Psych.), Grad. Dip. Soc. Sci., MAPS FCCLP

AHPRA Registration No. PSY0001685145

Participant Consent

CONSENT:

I have read the Information For Participants;

I agree to participate in this activity, knowing that I may withdraw at any time;

I agree that research data gathered for this study may be quoted and published without identifying me; and

I am 18 years of age or older.

- ◯ Yes, I consent to participate in this research project.
- ◯ No, I do not wish to participate in this research project.

Personal Details / Demographic Information 1

We would like to ask some personal questions that may improve our ability to understand the findings from this research. Remember, your name will never be matched to this information.

What is your age (in years only)?

What is your gender?

- ○ Male
- ○ Female
- ○ Other

Where do you live?

- ○ In a major city
- ○ In a large regional city/town
- ○ In a rural / remote area

Please indicate the highest level of education you have completed:

- ○ Primary (Year 6 or below)
- ○ Secondary (Years 7 - 10)
- ○ Secondary (Years 11 or 12)
- ○ University degree
- ○ Other tertiary (e.g. TAFE)

Personal Details / Demographic Information 2

Continued ...

Your occupation (if not being paid to work, please write "not in paid work"):

If in paid employment, how many hours do you work a week?

- ○ Less than 30 hours
- ○ 30 hours or more
- ○ Not in paid work

Approximately how much money do you earn a week after tax? (Including any government payments you receive)

- ○ Less than $300
- ○ $301 - $600
- ○ $601 - $1000
- ○ More than $1000
- ○ Prefer not to say

Mental Health Information

We would now like to ask some questions related to your health, and specifically your mental health. These questions are very important to this research and we ask that you answer as honestly as you can even though much of this information is sensitive. Again, remember that your name will not be matched to any of this information.

In the past, you have previously been diagnosed with (select more than one if applicable):

- ☐ Depression
- ☐ Anxiety (including panic attacks, posttraumatic stress disorder, obsessive-compulsive disorder, social anxiety, and phobias) - Please state what type of anxiety you suffer from.
- ☐ Other (please type in any other mental health condition you have been diagnosed with)

Approximately how long have you had anxiety and/or depression?

- ○ Less than a year
- ○ 1-5 years
- ○ 6-10 years
- ○ 11 years or longer

Are you currently seeing a psychiatrist, psychologist, counsellor, or other mental health professional for counselling related to your anxiety and/or depression?

- ☐ Psychiatrist
- ☐ Psychologist
- ☐ Counsellor
- ☐ Other mental health professional
- ☐ Not currently receiving counselling

Are you currently on any type of medication for your anxiety and/or depression?

- ○ Yes
- ○ No

If yes, what type of medication is it? (If you know the name of your medication but you are not sure which category it belongs to, write the name of it next to "Other")

- ☐ Don't know
- ☐ I've never been on medication for my anxiety and/or depression
- ☐ Antidepressant
- ☐ Benzodiazepine / tranquiliser

- [ ] Antipsychotic
- [ ] Mood stabiliser
- [ ] Other (if known, please type in the name of the medication if it does not fit into any of the above categories)

What self-help strategies do you currently practice to help manage your symptoms of anxiety and/or depression? (Select all that apply)

- [ ] Meditation / mindfulness
- [ ] Physical exercise
- [ ] Dietary considerations
- [ ] Abstain from drinking alcohol
- [ ] Increase my social contact and activities with other people
- [ ] Engage in a hobby
- [ ] Other (please tell us)

General Health & Other Information

Continued

Do you currently suffer from any other chronic medical condition (e.g. asthma, diabetes etc.)?

- ○ No
- ○ Yes (please tell us what other medical conditions you have)

On average, how many hours of sleep would you get each night?

- ○ Less than 4 hours
- ○ 4-6 hours
- ○ 7-8 hours

- ○ More than 8 hours

Overall, how would you rate your physical health?

- ○ Excellent
- ○ Good
- ○ Average
- ○ Poor
- ○ Terrible

In your opinion, which of the following can cause anxiety and/or depression? (You can select more than one)

- ☐ Don't know
- ☐ Genetics (family history)
- ☐ Weak character
- ☐ Abnormal brain chemistry (chemicals in the brain not working properly)
- ☐ The weather (people in cold, dark places get more depressed than people in warmer, bright places)
- ☐ Major catastrophic event (e.g. death of a loved one, natural disaster, being the victim of a crime etc.)
- ☐ Other (please type in anything else that may cause anxiety and/or depression)

How would you rate the following treatment options for anxiety and/or depression?

	Helpful	Harmful	Don't Know
GP / Family Doctor	○	○	○
Pharmacist / Chemist	○	○	○
Counsellor or Social Worker	○	○	○

	Helpful	Harmful	Don't Know
Telephone counselling service (such as Lifeline)	○	○	○
Webchat / website counselling service	○	○	○
Psychiatrist	○	○	○
Psychologist	○	○	○
Close family and friends	○	○	○
Naturopath / herbalist	○	○	○
The clergy, a minister, or a priest	○	○	○

Please respond to the following statements by checking the appropriate response:

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewh agree
My psychological health is likely to improve as a result of participating in this research	○	○	○	○
I am ready to work on changing my psychological health for the better	○	○	○	○
I am motivated to do what the mobile app suggests is good for my psychological health	○	○	○	○

What brand of smartphone do you use?

- ○ Apple iPhone
- ○ Samsung
- ○ Google Pixel
- ○ Oppo
- ○ Huawei

- ○ Nokia
- ○ LG
- ○ Other (please tell us what type of phone you have if it's not in the above list)

How would you rate your abilities to use the functions of your smartphone?

Terrible	Poor	Average	Good	Excellent
○	○	○	○	○

How would you rate your abilities to engage with and use other types of technology generally?

Terrible	Poor	Average	Good	Excellent
○	○	○	○	○

Have you previously downloaded a mobile app to help your anxiety and/or depression?

- ○ No
- ○ Yes (please tell us the name of any apps that you remember downloading)

If you have previously downloaded an app to help your anxiety and/or depression, was it helpful?

- ○ No
- ○ Yes
- ○ I've never downloaded an app to help my anxiety and/or depression

Technology has the potential to help people manage their anxiety and/or depression.

Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
○	○	○	○	○

Document B.3

Daily SUDS Rating

HOW DO I FEEL TODAY?

- ○ **0** ..... No distress
- ○ **1**
- ○ **2** ..... Minimal distress
- ○ **3**
- ○ **4**
- ○ **5** ..... Moderate distress
- ○ **6**
- ○ **7**
- ○ **8** ..... Very distressed
- ○ **9**
- ○ **10** ..... Worst distress

Document B.4

DASS-21 Questionnaire

DASS21

Name: Date:

Please read each statement and circle a number 0, 1, 2 or 3 which indicates how much the statement applied to you **over the past week**. There are no right or wrong answers. Do not spend too much time on any statement.

The rating scale is as follows:

0 Did not apply to me at all
1 Applied to me to some degree, or some of the time
2 Applied to me to a considerable degree or a good part of time
3 Applied to me very much or most of the time

1 (s)	I found it hard to wind down	0	1	2	3
2 (a)	I was aware of dryness of my mouth	0	1	2	3
3 (d)	I couldn't seem to experience any positive feeling at all	0	1	2	3
4 (a)	I experienced breathing difficulty (e.g. excessively rapid breathing, breathlessness in the absence of physical exertion)	0	1	2	3
5 (d)	I found it difficult to work up the initiative to do things	0	1	2	3
6 (s)	I tended to over-react to situations	0	1	2	3
7 (a)	I experienced trembling (e.g. in the hands)	0	1	2	3
8 (s)	I felt that I was using a lot of nervous energy	0	1	2	3
9 (a)	I was worried about situations in which I might panic and make a fool of myself	0	1	2	3
10 (d)	I felt that I had nothing to look forward to	0	1	2	3
11 (s)	I found myself getting agitated	0	1	2	3
12 (s)	I found it difficult to relax	0	1	2	3
13 (d)	I felt down-hearted and blue	0	1	2	3
14 (s)	I was intolerant of anything that kept me from getting on with what I was doing	0	1	2	3
15 (a)	I felt I was close to panic	0	1	2	3
16 (d)	I was unable to become enthusiastic about anything	0	1	2	3
17 (d)	I felt I wasn't worth much as a person	0	1	2	3
18 (s)	I felt that I was rather touchy	0	1	2	3
19 (a)	I was aware of the action of my heart in the absence of physical exertion (e.g. sense of heart rate increase, heart missing a beat)	0	1	2	3
20 (a)	I felt scared without any good reason	0	1	2	3
21 (d)	I felt that life was meaningless	0	1	2	3

Document B.5

OQ-45.2 Questionnaire

Outcome Questionnaire (OQ®-45.2) **Name:**_____ **Date:**___/___/_____

Instructions: Looking back over the last week, including today, help us understand how you have been feeling. Read each item carefully and fill the circle completely under the category which best describes your current situation. For this questionnaire, work is defined as employment, school, housework, volunteer work, and so forth.

	Never	Rarely	Sometimes	Frequently	Almost Always
1. I get along well with others	O	O	O	O	O
2. I tire quickly	O	O	O	O	O
3. I feel no interest in things	O	O	O	O	O
4. I feel stressed at work/school	O	O	O	O	O
5. I blame myself for things	O	O	O	O	O
6. I feel irritated	O	O	O	O	O
7. I feel unhappy in my marriage/significant relationship	O	O	O	O	O
8. I have thoughts of ending my life	O	O	O	O	O
9. I feel weak	O	O	O	O	O
10. I feel fearful	O	O	O	O	O
11. After heavy drinking, I need a drink the next morning to get going. (If you do not drink, mark "never")	O	O	O	O	O
12. I find my work/school satisfying	O	O	O	O	O
13. I am a happy person	O	O	O	O	O
14. I work/study too much	O	O	O	O	O
15. I feel worthless	O	O	O	O	O
16. I am concerned about family troubles	O	O	O	O	O
17. I have an unfulfilling sex life	O	O	O	O	O
18. I feel lonely	O	O	O	O	O
19. I have frequent arguments	O	O	O	O	O
20. I feel loved and wanted	O	O	O	O	O
21. I enjoy my spare time	O	O	O	O	O
22. I have difficulty concentrating	O	O	O	O	O
23. I feel hopeless about the future	O	O	O	O	O
24. I like myself	O	O	O	O	O
25. Disturbing thoughts come into my mind that I cannot get rid of	O	O	O	O	O
26. I feel annoyed by people criticizing my drinking (or drug use). (If not applicable, mark "never")	O	O	O	O	O
27. I have an uncomfortable	O	O	O	O	O
28. I am not working/studying as well as I used to	O	O	O	O	O
29. My health worries me too much	O	O	O	O	O
30. I have trouble getting along with friends and close acquaintances	O	O	O	O	O
31. I am satisfied with my life	O	O	O	O	O
32. I have trouble at work/school because of drinking or drug use. (If not applicable, mark "never")	O	O	O	O	O
33. I feel that something bad is going to happen	O	O	O	O	O
34. I have sore muscles	O	O	O	O	O
35. I feel afraid of open spaces, of driving, or being on buses, subways, and so forth.	O	O	O	O	O
36. I feel nervous	O	O	O	O	O
37. I feel my love relationships are full and complete	O	O	O	O	O
38. I feel that I am not doing well at work/school	O	O	O	O	O
39. I have too many disagreements at work/school	O	O	O	O	O
40. I feel something is wrong with my mind	O	O	O	O	O
41. I have trouble falling asleep or staying asleep	O	O	O	O	O
42. I feel blue	O	O	O	O	O
43. I am satisfied with my relationships with others	O	O	O	O	O
44. I feel angry enough at work/school to do something I might regret	O	O	O	O	O
45. I have headaches	O	O	O	O	O

Developed by
Michael J. Lambert, Ph.D.
and
Gary M. Burlingame, Ph.D.

For More Information Contact:

OQ Measures, LLC
P.O. Box 521047
Salt Lake City, UT 84152

Toll-Free USA:
1-888-MH-SCORE
(1-888-647-2673)

Phone: (801) 649-4392
Fax: (801) 747-6900

Email:
INFO@OQMEASURES.COM

Website:
WWW.OQMEASURES.COM

Document B.6

uMARS Questionnaire – Page 1

Instructions for use:

Raters should:

1. Use the app and trial it thoroughly for at least 10 minutes;
2. Determine how easy it is to use, how well it functions and does it do what it purports to do;
3. Review app settings, developer information, external links, security features, etc.

Scoring

A: Engagement Mean Score = _____

B: Functionality Mean Score = _____

C: Aesthetics Mean Score = _____

D: Information Mean Score* = _____

* Exclude questions rated as "N/A" from the mean score calculation.

App quality mean score _____ = A + B + C + D / 4

The *App subjective quality* scale can be reported as individual items or as a mean score, depending on the aims of the research.

The *Perceived impact* items can be adjusted and used to obtain information on the perceived impact of the app on the user's knowledge, attitudes and intentions related to the target health behaviour.

Document B.7

uMARS Questionnaire – Page 2

Mobile Application Rating Scale: user version (uMARS)

App Name: _____

Circle the number that most accurately represents the quality of the app you are rating. All items are rated on a 5-point scale from "1.Inadequate" to "5.Excellent". Select N/A if the app component is irrelevant.

App Quality Ratings

SECTION A

Engagement – fun, interesting, customisable, interactive, has prompts (e.g. sends alerts, messages, reminders, feedback, enables sharing)

1. **Entertainment: Is the app fun/entertaining to use? Does it have components that make it more fun than other similar apps?**
   1 Dull, not fun or entertaining at all
   2 Mostly boring
   3 OK, fun enough to entertain user for a brief time (< 5 minutes)
   4 Moderately fun and entertaining, would entertain user for some time (5-10 minutes total)
   5 Highly entertaining and fun, would stimulate repeat use

2. **Interest: Is the app interesting to use? Does it present its information in an interesting way compared to other similar apps?**
   1 Not interesting at all
   2 Mostly uninteresting
   3 OK, neither interesting nor uninteresting; would engage user for a brief time (< 5 minutes)
   4 Moderately interesting; would engage user for some time (5-10 minutes total)
   5 Very interesting, would engage user in repeat use

3. **Customisation: Does it allow you to customise the settings and preferences that you would like to (e.g. sound, content and notifications)?**
   1 Does not allow any customisation or requires setting to be input every time
   2 Allows little customisation and that limits app's functions
   3 Basic customisation to function adequately
   4 Allows numerous options for customisation
   5 Allows complete tailoring the user's characteristics/preferences, remembers all settings

4. **Interactivity: Does it allow user input, provide feedback, contain prompts (reminders, sharing options, notifications, etc.)?**
   1 No interactive features and/or no response to user input
   2 Some, but not enough interactive features which limits app's functions
   3 Basic interactive features to function adequately
   4 Offers a variety of interactive features, feedback and user input options
   5 Very high level of responsiveness through interactive features, feedback and user input options

Document B.8

uMARS Questionnaire – Page 3

5. **Target group: Is the app content (visuals, language, design) appropriate for the target audience?**
   1 Completely inappropriate, unclear or confusing
   2 Mostly inappropriate, unclear or confusing
   3 Acceptable but not specifically designed for the target audience. May be inappropriate/ unclear/confusing at times
   4 Designed for the target audience, with minor issues
   5 Designed specifically for the target audience, no issues found

SECTION B

Functionality – app functioning, easy to learn, navigation, flow logic, and gestural design of app

6. **Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?**
   1 App is broken; no/insufficient/inaccurate response (e.g. crashes/bugs/broken features, etc.)
   2 Some functions work, but lagging or contains major technical problems
   3 App works overall. Some technical problems need fixing, or is slow at times
   4 Mostly functional with minor/negligible problems
   5 Perfect/timely response; no technical bugs found, or contains a 'loading time left' indicator (if relevant)

7. **Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons and instructions?**
   1 No/limited instructions; menu labels, icons are confusing; complicated
   2 Takes a lot of time or effort
   3 Takes some time or effort
   4 Easy to learn (or has clear instructions)
   5 Able to use app immediately; intuitive; simple (no instructions needed)

8. **Navigation: Does moving between screens make sense; Does app have all necessary links between screens?**
   1 No logical connection between screens at all /navigation is difficult
   2 Understandable after a lot of time/effort
   3 Understandable after some time/effort
   4 Easy to understand/navigate
   5 Perfectly logical, easy, clear and intuitive screen flow throughout, and/or has shortcuts

9. **Gestural design: Do taps/swipes/pinches/scrolls make sense? Are they consistent across all components/screens?**
   1 Completely inconsistent/confusing
   2 Often inconsistent/confusing
   3 OK with some inconsistencies/confusing elements
   4 Mostly consistent/intuitive with negligible problems
   5 Perfectly consistent and intuitive

Document B.9

uMARS Questionnaire – Page 4

SECTION C

Aesthetics – graphic design, overall visual appeal, colour scheme, and stylistic consistency

10. Layout: Is arrangement and size of buttons, icons, menus and content on the screen appropriate?

1 Very bad design, cluttered, some options impossible to select, locate, see or read
2 Bad design, random, unclear, some options difficult to select/locate/see/read
3 Satisfactory, few problems with selecting/locating/seeing/reading items
4 Mostly clear, able to select/locate/see/read items
5 Professional, simple, clear, orderly, logically organised

11. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus and content?

1 Graphics appear amateur, very poor visual design - disproportionate, stylistically inconsistent
2 Low quality/low resolution graphics; low quality visual design – disproportionate
3 Moderate quality graphics and visual design (generally consistent in style)
4 High quality/resolution graphics and visual design – mostly proportionate, consistent in style
5 Very high quality/resolution graphics and visual design - proportionate, consistent in style throughout

12. Visual appeal: How good does the app look?

1 Ugly, unpleasant to look at, poorly designed, clashing, mismatched colours
2 Bad – poorly designed, bad use of colour, visually boring
3 OK – average, neither pleasant, nor unpleasant
4 Pleasant – seamless graphics – consistent and professionally designed
5 Beautiful – very attractive, memorable, stands out; use of colour enhances app features/menus

SECTION D

Information – Contains high quality information (e.g. text, feedback, measures, references) from a credible source

13. Quality of information: Is app content correct, well written, and relevant to the goal/topic of the app?

N/A There is no information within the app
1 Irrelevant/inappropriate/incoherent/incorrect
2 Poor. Barely relevant/appropriate/coherent/may be incorrect
3 Moderately relevant/appropriate/coherent/and appears correct
4 Relevant/appropriate/coherent/correct
5 Highly relevant, appropriate, coherent, and correct

14. Quantity of information: Is the information within the app comprehensive but concise?

N/A There is no information within the app
1 Minimal or overwhelming
2 Insufficient or possibly overwhelming
3 OK but not comprehensive or concise
4 Offers a broad range of information, has some gaps or unnecessary detail; or has no links to more information and resources
5 Comprehensive and concise; contains links to more information and resources

Document B.10

uMARS Questionnaire – Page 5

15. Visual information: Is visual explanation of concepts – through charts/graphs/images/videos, etc. – clear, logical, correct?

N/A There is no visual information within the app (e.g. it only contains audio, or text)
1 Completely unclear/confusing/wrong or necessary but missing
2 Mostly unclear/confusing/wrong
3 OK but often unclear/confusing/wrong
4 Mostly clear/logical/correct with negligible issues
5 Perfectly clear/logical/correct

16. Credibility of source: does the information within the app seem to come from a credible source?

N/A There is no information within the app
1 Suspicious source
2 Lacks credibility
3 Not suspicious but legitimacy of source is unclear
4 Possibly comes from a legitimate source
5 Definitely comes from a legitimate/specialised source

App subjective quality

SECTION E

17. Would you recommend this app to people who might benefit from it?

1	Not at all	I would not recommend this app to anyone
2		There are very few people I would recommend this app to
3	Maybe	There are several people I would recommend this app to
4		There are many people I would recommend this app to
5	Definitely	I would recommend this app to everyone

18. How many times do you think you would use this app in the next 12 months if it was relevant to you?

1 None
2 1-2
3 3-10
4 10-50
5 >50

19. Would you pay for this app?

1 Definitely not
2
3
4
5 Definitely yes

20. What is your overall (star) rating of the app?

1	★	One of the worst apps I've used
2	★★	
3	★★★	Average
4	★★★★	
5	★★★★★	One of the best apps I've used

Document B.11

uMARS Questionnaire – Page 6

Perceived impact

SECTION F

1. **Awareness – This app has increased my awareness of the importance of addressing the health behaviour**

Strongly disagree				Strongly Agree
1	2	3	4	5

2. **Knowledge – This app has increased my knowledge/understanding of the health behaviour**

Strongly disagree				Strongly Agree
1	2	3	4	5

3. **Attitudes – The app has changed my attitudes toward improving this health behaviour**

Strongly disagree				Strongly Agree
1	2	3	4	5

4. **Intention to change – The app has increased my intentions/motivation to address this health behaviour**

Strongly disagree				Strongly Agree
1	2	3	4	5

5. **Help seeking – This app would encourage me to seek further help to address the health behaviour (if I needed it)**

Strongly disagree				Strongly Agree
1	2	3	4	5

6. **Behaviour change – Use of this app will increase/decrease the health behaviour**

Strongly disagree				Strongly Agree
1	2	3	4	5

Further comments about the app?

THANK YOU!

Document B.12

TREND Statement Checklist Page 1

TREND Statement Checklist

Paper Section/ Topic	Item No	Descriptor	Reported? ✓	Reported? Pg #
Title and Abstract				
Title and Abstract	1	• Information on how unit were allocated to interventions	✓	
		• Structured abstract recommended		
		• Information on target population or study sample	✓	
Introduction				
Background	2	• Scientific background and explanation of rationale	✓	
		• Theories used in designing behavioral interventions		
Methods				
Participants	3	• Eligibility criteria for participants, including criteria at different levels in recruitment/sampling plan (e.g., cities, clinics, subjects)	✓	
		• Method of recruitment (e.g., referral, self-selection), including the sampling method if a systematic sampling plan was implemented	✓	
		• Recruitment setting	✓	
		Settings and locations where the data were collected	N/A	
Interventions	4	• Details of the interventions intended for each study condition and how and when they were actually administered, specifically including:	✓	
		○ Content: what was given?	✓	
		○ Delivery method: how was the content given?	✓	
		○ Unit of delivery: how were the subjects grouped during delivery?	N/A	
		○ Deliverer: who delivered the intervention?	N/A	
		○ Setting: where was the intervention delivered?	N/A	
		○ Exposure quantity and duration: how many sessions or episodes or events were intended to be delivered? How long were they intended to last?	✓	
		○ Time span: how long was it intended to take to deliver the intervention to each unit?	✓	
		○ Activities to increase compliance or adherence (e.g., incentives)	✓	
Objectives	5	• Specific objectives and hypotheses	✓	
Outcomes	6	• Clearly defined primary and secondary outcome measures	✓	
		• Methods used to collect data and any methods used to enhance the quality of measurements	✓	
		• Information on validated instruments such as psychometric and biometric properties	✓	
Sample Size	7	How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules	✓	
Assignment Method	8	Unit of assignment (the unit being assigned to study condition, e.g., individual, group, community)	✓	
		Method used to assign units to study conditions, including details of any restriction (e.g., blocking, stratification, minimization)	✓	
		Inclusion of aspects employed to help minimize potential bias induced due to non-randomization (e.g., matching)	✓	

Document B.13

TREND Statement Checklist Page 2

TREND Statement Checklist

Blinding (masking)	9	• Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to study condition assignment; if so, statement regarding how the blinding was accomplished and how it was assessed.	N/A	
Unit of Analysis	10	• Description of the smallest unit that is being analyzed to assess intervention effects (e.g., individual, group, or community)	✓	
		• If the unit of analysis differs from the unit of assignment, the analytical method used to account for this (e.g., adjusting the standard error estimates by the design effect or using multilevel analysis)		
Statistical Methods	11	• Statistical methods used to compare study groups for primary methods outcome(s), including complex methods of correlated data	N/A	
		• Statistical methods used for additional analyses, such as a subgroup analyses and adjusted analysis	✓	
		• Methods for imputing missing data, if used	N/A	
		• Statistical software or programs used	✓	
Results				
Participant flow	12	Flow of participants through each stage of the study: enrollment, assignment, allocation, and intervention exposure, follow-up, analysis (a diagram is strongly recommended)	✓	
		○ Enrollment: the numbers of participants screened for eligibility, found to be eligible or not eligible, declined to be enrolled, and enrolled in the study	✓	
		○ Assignment: the numbers of participants assigned to a study condition	✓	
		○ Allocation and intervention exposure: the number of participants assigned to each study condition and the number of participants who received each intervention	✓	
		○ Follow-up: the number of participants who completed the follow-up or did not complete the follow-up (i.e., lost to follow-up), by study condition	✓	
		○ Analysis: the number of participants included in or excluded from the main analysis, by study condition	✓	
		• Description of protocol deviations from study as planned, along with reasons	N/A	
Recruitment	13	• Dates defining the periods of recruitment and follow-up	✓	
Baseline Data	14	Baseline demographic and clinical characteristics of participants in each study condition	✓	
		Baseline characteristics for each study condition relevant to specific disease prevention research	✓	
		Baseline comparisons of those lost to follow-up and those retained, overall and by study condition		
		• Comparison between study population at baseline and target population of interest		
Baseline equivalence	15	• Data on study group equivalence at baseline and statistical methods used to control for baseline differences	✓	

Document B.14

TREND Statement Checklist Page 3

TREND Statement Checklist

Numbers analyzed	16	• Number of participants (denominator) included in each analysis for each study condition, particularly when the denominators change for different outcomes; statement of the results in absolute numbers when feasible	✓	
		• Indication of whether the analysis strategy was "intention to treat" or, if not, description of how non-compliers were treated in the analyses	N/A	
Outcomes and estimation	17	• For each primary and secondary outcome, a summary of results for each estimation study condition, and the estimated effect size and a confidence interval to indicate the precision	✓	
		• Inclusion of null and negative findings	✓	
		Inclusion of results from testing pre-specified causal pathways through which the intervention was intended to operate, if any	N/A	
Ancillary analyses	18	• Summary of other analyses performed, including subgroup or restricted analyses, indicating which are pre-specified or exploratory	✓	
Adverse events	19	• Summary of all important adverse events or unintended effects in each study condition (including summary measures, effect size estimates, and confidence intervals)	N/A	
DISCUSSION				
Interpretation	20	• Interpretation of the results, taking into account study hypotheses, sources of potential bias, imprecision of measures, multiplicative analyses, and other limitations or weaknesses of the study	✓	
		Discussion of results taking into account the mechanism by which the intervention was intended to work (causal pathways) or alternative mechanisms or explanations	✓	
		Discussion of the success of and barriers to implementing the intervention, fidelity of implementation	✓	
		• Discussion of research, programmatic, or policy implications	✓	
Generalizability	21	• Generalizability (external validity) of the trial findings, taking into account the study population, the characteristics of the intervention, length of follow-up, incentives, compliance rates, specific sites/settings involved in the study, and other contextual issues	✓	
Overall Evidence	22	• General interpretation of the results in the context of current evidence and current theory	✓	

From: Des Jarlais, D. C., Lyles, C., Crepaz, N., & the Trend Group (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health*, 94, 361-366. For more information, visit: http://www.cdc.gov/trendstatement/

Table B.5

uMARS App Ratings for Pilot Study Participants

Participant	Section A – Engagement (out of 25)	Section B – Functionality (out of 20)	Section C – Aesthetics (out of 15)	Section D – Information (out of 20)	Section E – App subjective quality (out of 20)	Section F – Perceived impact (out of 30)	TOTAL (out of 130)
1	18	13	12	14	13	20	90
2	21	18	14	18	20	27	118
3	21	16	13	16	17	26	109
4	17	13	10	13	8	17	78