
Ethical issues in the design and study of online therapy for mental health

Greg Wadley

Department of Computing and Information Systems
The University of Melbourne
Parkville 3010 Australia
greg.wadley@unimelb.edu.au

Mario Alvarez-Jimenez

Centre for Youth Mental Health
The University of Melbourne
Parkville 3010 Australia
malvarez@unimelb.edu.au

Reeva Lederman

Department of Computing and Information Systems
The University of Melbourne
Parkville 3010 Australia
reevaml@unimelb.edu.au

John Gleeson

School of Psychology
Australian Catholic University
115 Victoria Pde, Fitzroy, Australia
john.gleeson@acu.edu.au

Abstract

Research interest is burgeoning in the design and implementation of technology-based therapy for mental health and emotional wellbeing. But significant problems attend such projects. Some arise from ethical issues inherent in mental health research, while others arise when technology is used to deliver therapy. We illustrate this ethical environment via discussion of issues that have arisen in an Australian project to build web-based therapies for young people suffering from psychosis and mood disorders, and for their carers.

Author Keywords

mental health; online therapy; research ethics;

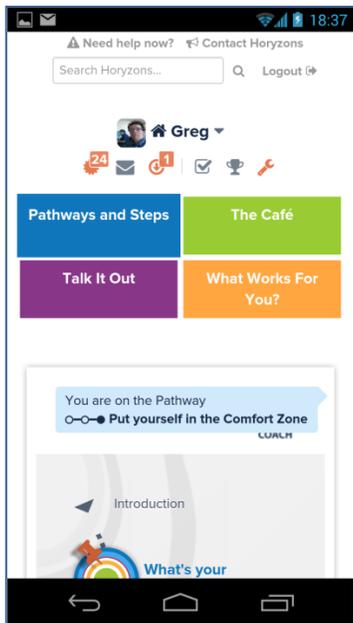
ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Mental illness and its treatment

Serious mental illnesses are devastating and stigmatising conditions which most often have their onset in during adolescence when young people are at a critical stage of their social and intellectual development. Previous work has shown that sufferers have significant potential for recovery if they remain engaged in effective treatment. However public-models of health care insufficiently resourced to provide adequate long-term “face-to-face” care in youth health. At the same time, young people increasingly expect that their interactions with institutions will be technology-based. As a result, they are receptive to the suggestion that mental health treatments will be accessible from their phone or computer. Recently significant interest has arisen in such systems.

However placing therapy online raises ethical issues involving safety and security among other concerns (Gleeson et al., 2014). The online setting cannot accommodate some specific safety and privacy controls inherent in the traditional setting.



Client's smartphone interface to our online mental health therapy showing four main features: Pathways and Steps (therapy content), the Café (social interaction), Talk It Out (structured group discussion), and What Works For You? (peer-sourced tips).

This discussion paper describes a project by the authors to build an online therapy for youth mental health, and discusses the ethical issues encountered in this project. We raise 14 issues and, where appropriate, discuss our responses to them.

About our project

The project is motivated by the need to provide efficacious and cost-effective services for young people with mental illness, to examine the effectiveness of the online alternative and to measure its efficacy as a long term treatment option; that is, as an adjunct to traditional services. The project has been running since 2009, with initial testing completed and trials underway including several pilot studies and an extended RCT involving up to 100 mental health clients conducted over 2014-8.

This is a highly multi-disciplinary collaboration involving psychologists, technologists, professional writers, graphic artists, youth workers and clients of a youth mental health centre. Consistent with contemporary models of consumer involvement in mental health services, we are employing participatory design methods so that clients and clinicians are heavily involved in designing the therapy. Over several years we have refined our design through usability expert testing; focus groups with clients and clinicians; pilot trials of prototype therapies; focus groups with moderators employed in the RCT, and user feedback.

Software features

The system is available to clients and moderators via desktop computer, tablet or phone. It delivers therapy content mainly via the modalities of text and graphics, with some audio and video. Materials are prepared by a

professional author with experience writing for young audiences. The software also offers a range of social features which allow interaction among the mental health clients using the site and the clinicians who moderate it. Most prominent of these features is the "Café" which is a newsfeed rather like Facebook's.

Our approach to site moderation is based on the principles of Supportive Accountability (Mohr et al., 2013) which are designed to promote hope and optimism and encourage clients to be accountable to moderators perceived as being trustworthy, benevolent experts. Moderators see the same pages that clients do, and take part in conversations along with clients. (Moderators also have a range of tools that clients do not see.) The site is moderated at scheduled times on weekdays and weekends so that regular and timely feedback can be provided to users to establish an atmosphere of trust.

Current trials

Following a safety and acceptability trial of our therapy with young people recovering from psychosis (Alvarez-Jimenez et al., 2013), we have begun an extended RCT, recruiting participants cumulatively over a 4 year period (2014-8).

We have also created versions for trial among young people with mood disorders, and parents and carers of clients.

Results to date

Results after one year of continuous use indicate that the technology is acceptable to clients and clinicians, with some variation. Few trial participants have quit, though engagement is an ongoing concern, and a

significant proportion of our work is devoted to maintaining engagement. We have experienced few safety issues, and the software has proved adequate to deal with those that have arisen. It is too soon to be certain about psychological outcomes; however many ethical challenges are already clear. Similar challenges will likely arise for any comparable project.

Underlying ethical context

The ethical issues that arise in this project stem in part from the underlying mental health context and would apply in any comparable therapy project. Ethical problems attached to doing research in the mental health area have been elaborated by Coyle et al. (2007). For good reason, regulations make it difficult for researchers to access mental health clients and settings. In this section we sketch the underlying ethical landscape in mental health, focusing on stigma, resource limitations, low user motivation and duty of care.

1. There is a significant stigma associated with mental health issues which can make many people, especially the young, unwilling to be identified as clients of mental health services [Corrigan, 1998]. This has consequences for online identity and the need for data security and privacy.
2. Mental health services are chronically under-resourced. An online service must help to ease shortages and should never become an extra burden on services.
3. Many people suffering mental health problems lack motivation. One trial participant explained their low usage of the system as follows: "I haven't been on the

computer for a while. I lay in bed most of the day." Yet for a therapy to work it must be used. It is a significant challenge to develop a technology which is engaging for young people, simple to use for clients who may suffer cognitive deficit, and which can reverse the low rates of adherence that characterize previous work in this area (Christensen et al., 2006).

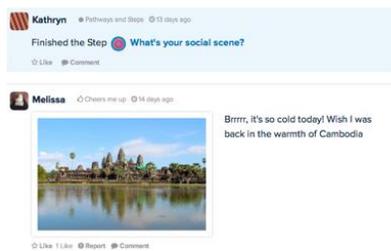
4. Young people with mental health problems can suffer periods of elevated risks. Clinics have a carefully-defined duty of care to their clients which is based partly on their ability to have reasonable knowledge of the client's circumstances. Potentially, an Internet-based service expands this duty to circumstances which are hard to manage because the client is not physically present.

Specific concerns in our project

We now outline some specific challenges that have manifested within our project. After some we describe our responses, while others remain open issues. These challenges are ethical as well as methodological and design-related; that is, ethical issues are entwined with the more general challenges that a design project in this sensitive setting will face.

1. **Privacy vs sharing:** Because of the stigma attached to mental health conditions, user privacy is important. Yet we are asking clients to share their experiences online. How can a social therapy balance the need for privacy with the need for open discussion?

Response: Our users may choose a pseudonymous username. The system allows users to delete their prior input. Users are able to 'hide' or 'unhide' themselves at



Social networking features include a "newsfeed" in which users post text, photos and video and in which system-generated messages appear.

any time to deal with temporary states of psychosis such as paranoid beliefs.

2. Asynchronicity: To meet user needs and complement existing services it is important that an Internet-based service be available 24/7. However for cost reasons moderators cannot be logged in at all times, so responses to messages and posts may not be immediate. Most critically, extreme events such as suicidal ideation may not be noticed immediately. This raises two ethical issues: firstly, are clients endangered by asynchronicity, and secondly, what psychological burden does this place on moderators who may feel a pressure to check the site and respond to clients beyond the hours for which they are employed?

Response: Users can "report" any post, which flags it for moderator attention. Each post is searched for particular keywords relating to abuse or self-harm: if found the post is blocked and flagged to moderators for immediate action; the author also receives a message with the phone number of an emergency mental health service. There is also a phone number that users can call 24/7 regarding system problems.

3. Duty of care: Clinics have adapted to dealing with clients face-to-face during business hours. An online system means that clinics are potentially in contact with clients at any time and with the client's whereabouts unknown. This creates difficult duty-of-care issues for the clinic and for individual clinicians and researchers who become aware (or could potentially have been aware) of client problems. What happens if a client indicates a problem out of hours or when no moderator is logged in? Even if the system can detect

problems by analyzing system usage, how can it know where the client is in order to direct help to them?

Response: Part of our response has been to be selective about trial participation. We employ specific inclusion and exclusion criteria and the therapy is available only to clients currently in remission. Also all clients are known to the moderation team, who have access to client health histories.

4. Moderator workload: Moderators in principle can access an online system at almost any time and place. Some have talked about coming home after a night out and checking on the system before going to sleep. This could become a burden for moderators if logging in at all hours were to become expected of them; furthermore such use cannot be reliably expected, and moderators may not always be in a good condition to work out of hours. More generally, we need to be certain that online therapy will not add to the workload of clinicians, who are overworked and under-resourced.

Response: Ultimately a best option response to this would be to have a service that was funded around the clock so moderators would not feel burdened during their time off. Failing this the site explains clearly what hours it is moderated, and contact details for alternative services that are provided out of hours. Focus groups with moderators showed that levels of out-of-hours concern was individual and personality-related. It is probably wise to address this issue in moderator training.

5. Suicidal ideation: It's foreseeable that expressions of suicidal thoughts and intentions could be distressing and trigger dangerous behaviour in others. We need to



Moderators, researchers and designers meet weekly to discuss ongoing issues arising out of the randomized control trial.

be careful that contagion does not occur within our system; on the other hand we want to encourage peer support. Many such conversations might be on the borderline of acceptable risk, making both automatic detection and human moderation difficult.

Response: Social interaction works in tandem with psycho-education modules which deal with dangerous thinking and ways to seek help. Users can report posts that worry them; these are flagged to moderators. The therapy's positive focus aims to avoid detrimental effects due to over-exposure to negative or deficit-focused content

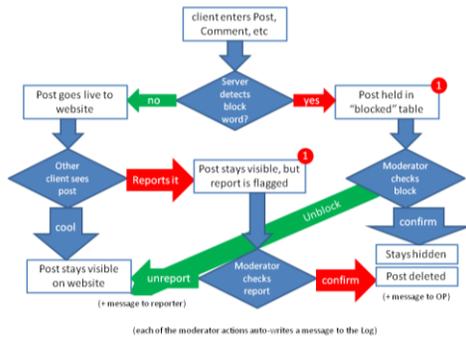


Diagram from moderator handbook explaining sociotechnical measures in place to check for and respond to the posting of problematic material by clients.

6. Censorship: Social technologies can be misused to convey abuse. Mental illness might lead some clients to be abusive, and others to be especially sensitive to abuse. The system screens posts for abusive words; however neither clients nor researchers want censorship, and it is sometimes difficult to define what is abuse and what is acceptable youth language.

Response: We examine ways to automatically screen the text input by users. However we need to bear in mind that very clear messages can be sent with video, audio and images and that these are harder to screen.

7. Data security: Security from hacking is of paramount importance, but it is impossible to be 100% secure. Can we define an acceptable level of risk? All research projects have a limited budget – how much should be devoted to security? For emergencies, clinicians want the system to store clients' contact information; however this increases of user identification in the case of a hack. What is the right balance between online and offline security?

Response: We cannot guarantee security, so we ensure users are aware of this, in accordance with the principle of informed consent.

8. Participatory design: We have engaged clients and clinicians in the design of our therapy (Wadley et al., 2013). In many contexts, including end-users in design is believed to lead to good outcomes. But in the health context there must be a trade-off between end-user and expert opinion on what the design should be.

9. Potential for disengagement from F2F services: It is feasible that providing an online therapy might lead users to visit the bricks-and-mortar clinic less often. Likewise if online therapy proves to be effective, it is plausible that governments may seek to save money by reducing bricks-and-mortar places in favour of online.

Response: It is not our intention that online therapy becomes a cheap alternative to existing services; rather it should be an adjunct. We need to maintain the message that online cannot be suitable for all clients.

10. Access: To what extent is access to hardware and Internet a problem for mental health clients in Australia? What can be done about this? Is it sufficient to assume that over time access will improve?

Response: Our experience indicates that the majority of prospective users do indeed have reasonable access. We have a quantity of tablet computers which we are providing to some users.

11. Literacy: How can a primarily text-based medium support clients whose literacy and/or competence with

English language are not sufficient? Is there an easy fix such as audio or video or automatic text translation?

12. **Research fatigue:** Young mental health clients can become “over-researched” as much is attempted with few willing participants. Getting maximum output is, in this context, not merely an economic concern but an ethical one. Which methods will best achieve this?

13. **Data retention:** How long will the data be retained by the system, and will the client be able to have it erased at a future time? This is pertinent because the data present a picture of the person at a particular phase of their life, which they may not want to remember or to be available to others in the future.

14. **Scale:** For a *social* therapy, what is the optimally-sized user cohort? Considering that an online technology is potentially accessible anywhere, where, if anywhere, should our borders lie? The individual clinic seems to be a useful organizing context, but if we draw borders, someone will miss out.

Conclusion

Creating a technology-mediated therapy in the ethically complex field of mental health is very challenging. Such pioneering projects as this one will help to elucidate the range of challenges and appropriate responses.

Our results so far are positive: when users have become unwell during the trial, the system has afforded support, and these users have been positive about their experience. However there are a range of ongoing issues to address.



Measures to maintain client engagement include posting cards to users upon their performance of particular actions within the system.

References

- [1] Alvarez-Jimenez, M., *et al.*, On the HORYZON: moderated online social therapy for long-term recovery in first episode psychosis, *Schizophrenia Research*, vol. 143, pp. 143-149, 2013.
- [2] Christensen, H., *et al.*, Free range users and one hit wonders: community users of an Internet-based cognitive behaviour therapy program. *Australian and New Zealand Journal of Psychiatry*, 2006. 40: p. 59-62.
- [3] Corrigan, P. W., The impact of stigma on severe mental illness. *Cognitive and Behavioral Practice*, 1998. 5: p. 201-222.
- [4] Coyle, D., *et al.*, Computers in talk-based mental health interventions, *Interacting with Computers*, vol. 19, pp. 545-562, 2007.
- [5] Gleeson, J. F., Lederman, R., Wadley, G., Bendall, S., McGorry, P. D., & Alvarez-Jimenez, M. (2014). Safety and privacy outcomes from a moderated online social therapy for young people with first-episode psychosis. *Safety*, 65(4).
- [6] Hagen, P., P. Collin, C. Metcalf, M. Nicholas, K. Rahilly and N. Swainston (2012). Participatory Design of evidence-based online youth mental health promotion, intervention and treatment. Melbourne, Young and Well Cooperative Research Centre.
- [7] Mohr, D., P. Cuijpers, and K. Lehman, Supportive Accountability: A Model for Providing Human Support to Enhance Adherence to eHealth Interventions. *Med Internet Res*, 2011. 13(1): p. e30.
- [8] Wadley, G., *et al.*, Participatory design of an online therapy for youth mental health, in *Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration*, 2013, pp. 517-526.