Author’s response to reviews

Title: Impact of a mobile phone and web psychological program on depression, anxiety, stress, work and social functioning: A randomised controlled trial

Authors:

Judith Proudfoot (j.proudfoot@unsw.edu.au)
Janine Clarke (janine.clarke@unsw.edu.au)
Mary-Rose Birch (mbirch@internode.on.net)
Alexis Whitton (awhitton@psy.unsw.edu.au)
Gordon Parker (g.parker@unsw.edu.au)
Vijaya Manicavasagar (v.manicavasagar@unsw.edu.au)
Virginia Harrison (gini.harrison@gmail.com)
Helen Christensen (h.christensen@blackdog.org.au)
Dusan Hadzi-pavlovic (d.hadzi-pavlovic@blackdog.org.au)

Version: 2
Date: 8 November 2013

Author’s response to reviews: see over
Dear Professor Andersson,

We thank you and the reviewers for your constructive comments. We have addressed each of the suggested revisions, and believe that this has substantially improved the manuscript.

Our responses to each comment are outlined below (numbering as per the reviewers’ comments).

We look forward to your response to these revisions.

Yours sincerely,

Judy Proudfoot (on behalf of the authors)

REVIEWER 1

Major compulsory revisions

Comment 1
Please be more explicit about the short-term efficacy of the program (i.e. by clearly stating that no differences between AC and MyCompass were found at the 3-month follow-up. Please do so in the abstract (in the results section) as well as in the discussion (first paragraph).

Response:
We have amended the Results section of the Abstract to now read:

“This myCompass group showed significantly greater improvement in symptoms of depression, anxiety and stress and in work and social functioning relative to both the control conditions at the end of the 7-week intervention phase (between-group effect sizes ranged between .22 and .55 based on the observed means). Symptom scores remained at near normal levels at 3-month follow-up. Participants in the attention control condition showed gradual symptom improvement during the post-intervention phase and their scores did not differ from the myCompass group at 3 month follow-up.”

The first paragraph of the Discussion has been amended and now reads:

“The myCompass intervention brought about rapid improvements in mental health symptoms and in work and social functioning in a large community sample. At post-intervention, the myCompass group showed...
significantly reduced symptoms of depression, anxiety and stress, and significantly improved levels of work and social functioning. Mostly moderate within-group and small-to-moderate between-group effects on the primary and secondary outcome measures were found for myCompass participants. Scores reduced to the normal range by post-intervention and treatment gains were maintained at 3-month follow-up. Participants in the AC condition showed gradual improvement over the post-intervention period and no differences were observed between myCompass and AC participants at 3 month follow-up. The within-group improvements seen in the myCompass group were later replicated in the WL group when they went through the myCompass intervention at the conclusion of their 7-week wait period.”

Comment 2
The effect sizes in Table 5 appear to be somewhat inflated, based on the results presented in Table 2 and 3. For instance, according to Table 2, participants in the myCompass group reported an increase of about 4 scale points on the WSAS. With a pooled standard deviation of - roughly - 8 [ (8.07 + 8.11 + 8.77) / 3 ], I would expect a d of around .5, while the reported d is .84. Similar inflations appear to be present for the other measures. Also, between-group effect sizes are higher than one would expect based on the parameter results in Table 3. Please check. Did estimated means reflect more favourable outcomes than the raw observed means? Were pooled standard deviations calculated correctly?

Response:
There are various approaches to effect size estimation. In revising the paper, we have recalculated the effect sizes using both observed and adjusted means (Table 5), and have adopted the approach recommended by the reviewer. The Method section has been amended to reflect this change and now reads:

“Cohen’s d was calculated following the procedure outlined in Ruwaard et al. [27] and using both observed and estimated marginal means. For all outcome measures, within- and between-group differences were standardised to Cohen’s d using the pooled standard deviation of the observed scores obtained at baseline.”

The observed and adjusted means are presented in Table 2 and the revised effect size estimates have been incorporated into Table 5 and the text.

Minor essential revisions

Comment 3
Please consider to add targeted symptom severity to the title (i.e., mild-to-moderate depression, anxiety, stress, and work and social functioning).

Response:
Thank you for your suggestion. The title has been amended and simplified as follows:

Impact of a mobile and web program on symptom and functional outcomes for people with mild- to- moderate depression, anxiety and stress: A randomised controlled trial.

Comment 4
Please add effect sizes (point estimates and 95% confidence intervals) to the results section of the abstract, as well as adherence rates. These are key outcomes that readers would expect to find in the abstract.

Response:
We have incorporated retention rates, between group effect sizes and outcomes of significance tests of the group by time interaction at post-intervention into the Results section of the Abstract.

Comment 5
The last sentence of the conclusion ("Mobile phones offer ...") is not necessarily grounded in the design and reported results of the study. In my opinion, this statement is too generic and should be best removed.

Response:
We have removed the statement to which you refer.

Comment 5(2)
Please provide references to any previous published research or descriptions of myCompass and its theoretical basis here (e.g., Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: Review of the emerging field and proof of concept study. J Ment Health 2011; 20: 509-24.). The introduction provides a little too less information on the specifics of the program (is it CBT? solution-focused? eclectic, etc), which makes it hard to follow the authors in the hypothesis that the program will have beneficial effects. While this information is provided in the Methods section, I feel the manuscript would benefit from more background information at this point.

Response:
Thank you for this suggestion. We have included a paragraph describing the aims and outcomes of the proof of concept study, as well as a description of the therapeutic approach and content of the myCompass program.

Comment 6
Please provide a short explanation for not reporting the outcomes on three secondary outcome measures that were included in the trial register (Subjective Units of Distress Scale, Self-efficacy and health service utilization). What results were obtained with these measures? If these measures did not show any effect, please report so as well. Please also explain briefly why the number of measurements in the trial register entry and the final report differ (the trial register mentions repeated measurements at baseline and at 3, 6, 9, 12, 15 and 18 weeks after intervention commencement).

Response:
The three secondary outcome measures we planned to measure were Subjective Units of Distress Scale, Mental Health Self-efficacy and health service utilisation. However, following pilot testing, we dropped the Subjective Units of Distress Scale, and reduced the number of follow-up points from 6 to 2. We found that participants were neither able to tolerate the full questionnaire battery, nor were
they responsive to all repeated measurements. We also concluded that measurement points at post-intervention and 3-months follow-up were sufficient to evaluate effectiveness.

The Mental Health Self-efficacy and health service utilisation measures will be reported in secondary outcomes papers. Mental Health Self efficacy was significantly reduced at post-test, while data from the health service utilisation measure have not been examined to date.

Comment 7
In the description of the intervention, please provide the exact number of psychological modules available to participants (especially since usage is reported in the results section). Any reference to a more extensive description of the treatment elements would also be helpful to readers who consider a replication study. Please also provide somewhat more detailed information with regard to the tailored suggestions that are made the myCompass program (and by what rule sets). For instance, what profile questionnaire was used?

Response:
We have revised the section headed ‘Intervention’ to: (a) more clearly describe the key therapeutic elements of the myCompass program (self-monitoring and therapeutic modules) and their relationship, (b) highlight the program’s mobile functionalities, and (c) clarify how feedback is provided. An additional paragraph describing the tailored recommendations of the program has also been added.

Comment 8
Treatment satisfaction was assessed for myCompass, but not for the attention control. Please briefly explain why not.

Response:
Treatment satisfaction was also assessed for the Attention Control participants and outcomes of this assessment have now been incorporated into the manuscript.

Comment 9
Please provide score ranges for the WSAS as well.

Response:
The ranges have now been included.

Comment 10
Please consider to correct statistical results for chance findings resulting from multiple testing (i.e., by using Bonferroni corrections or similar procedures).

Response:
The Bonferroni adjustment was applied to the set of interaction contrasts computed for each of the 5 study outcomes at post-intervention. Reference to this procedure has been incorporated in the Method and Results sections.
**Comment 11**
Satterthwaite's approximation appears to have been used to estimate the degrees of freedom in the mixed model analyses? Please provide the method. In addition, it may not be necessary to report degrees of freedom up to 2 decimal points precision in the Tables (as there is little practical difference in \( df = 710 \) or \( df = 710,44 \)).

*Response:*
We have added the statement below to the analysis section to clarify that Satterthwaite’s approximation was used to estimate error degrees of freedom. Decimal points have been removed from the tables.

“In the present study, restricted maximum likelihood (REML) was used to estimate model parameters and error degrees of freedom were calculated using Satterthwaite’s approximation [25].”

**Comment 12**
The definition of Adherence (last paragraph in the statistical analyses subsection, just before the results section) should perhaps be moved to the Secondary Outcomes section, as it describes three measures, not an analysis.

*Response:*
We have moved this statement as suggested.

**Comment 13**
Fifteen participants withdrew after randomization (n = 11 from the myCompass group). Their data was removed from analysis. Please explain why. I would have expected these data to be included in the intention-to-treat analyses.

*Response:*
We agree that such data would usually be included in intention-to-treat analyses. However, we did not have consent from these participants to use their data after they had withdrawn. We approached these participants to gain consent, but were unsuccessful. The following phrase has been added to explain the removal of these participants from the data analysis:

“Fifteen people subsequently withdrew from the study and, in the absence of consent to use their baseline data, we excluded them from the analyses.”

**Comment 14**
Table 1: Would it be better not to use abbreviations in the table headers, or provide a list of terms in Table notes (e.g., AC: attention control)?

*Response:*
We have written the group names in full in the Table header, and also provided an explanation in the Table notes.

**Comment 15**
Table 1/2/3/4/5: Would it be wise to explain abbreviations used in each Table using a Table note? This would help readers a lot to quickly assess the data.

Response:
Agreed. We have added notes to all tables explaining the abbreviations.

Comment 16
Table 2: Please provide N's for each group and assessment moment.

Response:
We have added group N’s at each assessment point to Table 2.

Comment 17
Table 2: “mCompass” -> myCompass

Response:
This typo has now been corrected.

Comment 18
Table 3 and Table 4: Please add p-values and parameter confidence intervals for all tests in the Tables, including those that were found to be non-significant.

Response:
P values and confidence intervals have been added for all non-significant interaction contrasts.

Comment 19
Table 3/4: Please check denominator degrees of freedom. It seems odd that these exceed the sample size (n = 720). I would expect the df’s to reflect the number of participants?

Response:
In mixed models repeated measures analyses, the maximum df (for some effects) is equal to the number of observations (i.e., the number of subjects multiplied by the number of repeats minus the number of missing data points) and, depending on the method used to calculate the df and the number of parameters in the covariance matrix, this can be reduced down to the number of subjects (or less).

Comment 20
Table 4: please add explicit reference to the two groups that were compared in the analyses in the title (myCompass vs AC).

Response:
The title of Table 4 now makes explicit reference to the myCompass and Attention Control groups.
Comment 21
Table 4: the contrasts comparing AC to myCompass during the intervention phase are essentially identically to those presented in Table 3, if I understand correctly? If so, would it be better to remove these from Table 4, to prevent redundancy and confusion?

Response:
We agree that the post-intervention contrast in Table 4 is essentially identical to that in Table 3. To reduce the redundancy, we have tried reporting the outcomes in a variety of ways, but each creates a different set of problems or questions. For example, re-running the analyses restricting to post-intervention and follow-up data would raise the question for some readers of why we didn’t include the pre-intervention data in the model. We have therefore added a footnote to the existing Table 4 explaining that the pre-post contrast is essentially the same as in Table 3, so that the Group X Time outcomes are clear.

Comment 22
Table 5: In the method section, between-group and within-group Cohen’s d appears to be calculated using different standard deviations. As a result, there is a slightly confusing inconsistency between the within-group and between-group effects. For instance, if myCompass vs. WL is d = -.50 (in favour of myCompass), then readers could expect to find that if the within-effect in the myCompass group is .84, that the effect in the WL group would be .34. However, in the table, this effect is .38. Using the same pooled standard deviation for both between and within-group effect size calculations would perhaps provide a more clear and more consistent presentation?

Response:
This has been addressed under Comment 2 above.

Comment 23
Results/last paragraph before waitlist group results ("Moderate- to- large (d range = .44 to 1.04) within-group effects were seen in the myCompass group"). Do you mean d = .84 instead of 1.04 as the maximum effect size (as listed in Table 5)?

Response:
We have corrected the text to reflect the revised estimates of within-group effects.

Comment 23(2)
Please discuss possible negative effects of drop-out more. Mixed Modelling provides estimates that are based on observed values. It statistically corrects for attrition by increasing the uncertainty of the regression estimates. However, it can not be ruled out that study drop-outs would have reported more negative outcomes. With a 50% attrition rate in the MyCompass group (significantly higher than in the control groups), this is a major concern.

Response:
We have now made specific reference in the Discussion to the possibility that dropouts from the intervention group may have been less satisfied and experienced less positive symptom outcomes than non-dropouts.
“Consistent with previous internet trials [23,33], dropout attrition was high, especially for the myCompass group, and rates of engagement for myCompass participants with the program content were highly variable (and in some instances minimal). Inspection of possible biases due to attrition showed that dropouts were more likely to be male and employed, thus reducing our confidence in generalising to these groups. While our statistical methods accounted for dropout attrition and non-completion, we cannot discount the possibility that dropouts from the intervention group were less satisfied with the program and/or experienced less positive outcomes. Further research is required to examine the predictors of usage of the myCompass program, and relations between program usage and symptom and functional gains.”

Comment 25
Statistical results were replicated in the wait list group, although within-group effect sizes were considerably smaller in this group after treatment. These results seem to contradict the statement that "The WL group achieved outcomes comparable to those observed for the myCompass group after using the program for seven weeks" (discussion/first paragraph). Please explain.

Response:
We acknowledge the apparent inconsistency in the original paragraph, and have reworded it to clarify that the pattern of significant improvement observed for the myCompass group over the intervention period was observed for the Waitlist group as follows:

“The myCompass intervention brought about rapid improvements in mental health symptoms and in work and social functioning in a large community sample. At post-intervention, the myCompass group showed significantly reduced symptoms of depression, anxiety and stress, and significantly improved levels of work and social functioning. Moderate within-group effects and small- to- moderate between-group effects on the primary and secondary outcome measures were found for myCompass participants, with treatment gains maintained at 3-month follow-up. Participants in the AC condition showed gradual improvement over the post-intervention period and no differences were observed between myCompass and AC participants at 3-month follow-up. The immediate within-group improvements seen in the myCompass group were later replicated in the WL group when they went through the myCompass intervention at the conclusion of their 7-week wait period.”

Comment 24
While the effects are quite large for a self-help program, I do not necessarily agree with the suggestion that the effects of MyCompass are as good as guided programs (second paragraph). Although there certainly are guided programs with similar effect sizes, there are also guided programs with which effect sizes have been obtained that are substantially greater (see, for example, the individual study results in the review by Andersson and Cuijpers, 2009). Therapist input can make a difference, and readers should be informed about this.

Response:
We have removed this statement from the manuscript.
REVIEWER 2

Minor essential revisions

Comment 126
"Based on our pilot data [10]" --> I think this should be mentioned in the Introduction. How was it designed? What were the outcomes and conclusions from that pilot?"

Response:
We have amended the Introduction to include a paragraph describing the aims and outcomes of the proof of concept study, including a brief description of the therapeutic approach and content of the myCompass program.

Comment 165
"...stored on the users’ phone but instead transferred via the internet using secure sockets layer" --> Were the SMS and emails secured on the phone in any way? How? That kind of data is normally always stored locally and not on servers if the user don’t remove them manually. How did you solve the security issues here?"

Response:
We have clarified this statement to make it clear that it refers to user generated self-monitoring data.

“User privacy is managed by a password protected log-on, and by ensuring that user generated data (i.e., self-monitoring ratings) are not stored on the users’ phone but instead transferred via the internet using secure sockets layer protocols (which encrypt transmitted data by rendering it unreadable to anyone other than the intended recipient) and by storing the data in secure servers.”

Comment 177
"...factual information about depression, anxiety and stress, sent to their email address." --> How much information did the AC group received (how many characters long were the information)? How much (in a quantitative matter) did these texts differ from the texts that the treatment group received?

Response:
We have expanded this paragraph to more clearly describe the nature of the attention control intervention. In particular, we have provided additional details of the e-mail and SMS messages sent to AC participants, to allow comparison with myCompass group.

“Each week for seven weeks, they received a fact sheet containing information about depression, anxiety or stress sent to their email address. The information was designed to be read on computer in approximately 10 minutes, and to be credible but void of management advice or treatment strategies. They also received on their mobile phones weekly SMS messages containing brief factual statements about depression, anxiety and stress. The mobile phone statements were also therapeutically inactive, but chosen to ensure that the control program had face validity. Messages varied in length from 160 to 275 characters so as to match, as far as possible, the minimum amount of SMS content that myCompass participants received each week, namely one self-
monitoring reminder (160 characters) plus an optional motivational statement, mental health care tip or fact (average length 95 characters).”

Comment 186
"...including the novelty of using mobile phones" --> It think this statement is difficult to control. It could be the content that the users received on their phone, as well as the format.

Response:
We agree and have amended this sentence to read as follows:

“By including both the AC group and the WL group, we were able to control for non-specific effects of the intervention”.

Comment 189
I would like more information here. How did you ensure that participants met the inclusion criteria? Was it done only from the online screening? What did you inform participants that scored > 63 on PSS? Was the whole procedure automated, including the email sent out to the included participants?

Response:
This section has been amended and a paragraph added to provide a clearer description of the study recruitment and data collection procedures.

“Data collection took place between October 2011 and October 2012. All study consent, screening and questionnaire data were collected online using online survey software. Individuals not meeting screening criteria received automated feedback explaining the reason/s for their ineligibility, and referring them to other online resources of the BDI. Individuals with symptoms in the severe range were also advised to seek face-to-face support from a health professional and were provided the contact details of crisis support services.

Eligible participants completed a baseline questionnaire prior to randomisation, a post-intervention questionnaire administered at eight weeks, and a follow-up questionnaire administered 12 weeks later for participants in the myCompass and AC groups, and 19 weeks later for the WL group. At each assessment point, participants accessed the appropriate study questionnaire via a link sent to them in an email message.”

Comment 283
A few more sub-headers would make it easier to understand and follow the results in the Outcomes at post-intervention and follow-up section.

Response:
Additional subheadings have been provided to simplify interpretation of study outcomes. Post-intervention and follow-up outcome data are now presented in separate sections and headed appropriately. An additional sub-heading for effect size data has been included.

Comment 390
"...myCompass is an unsupported and completely automated public health intervention" --> Was there a therapist sending out email to getting the included participants started? If so, I am not sure if I would call the program completely automated.

Response:
The reviewer is correct in noting that myCompass participants received an email from the research team providing details about where to find and how to register with the program. Participants in the other two groups received instructions about their “program” in the same way. The myCompass program was, however, delivered without any therapist input. The statement has been amended to reduce potential confusion as follows:

“myCompass is a fully automated public health intervention with no therapist support that people can use via the internet on their mobile phones and computers.”

Comment 42
"Mobile phone-based psychological interventions enable real time self-monitoring and..." --> I think that this framing of the intervention/program is a bit wrong. For me, this is not a mobile phone-based psychological intervention. When I read this, my thoughts are that the program is web-based and supposed to be used on a computer. That's the core of the intervention. The mobil phone component is only to remind the user and give him/her small tips. If this is not the case, I think it is important to explain the intervention more in depth and how you expect the different components to interact with each other, and what purpose they have in the intervention.

Response:
We have taken on board your suggestion and explained the intervention in more depth. The core features of the myCompass intervention consist of both the mobile-based self-monitoring of key dimensions and the computer-based therapeutic modules. The mobile phone is not just used for reminders and tips, rather it has been programmed to carry out the important function of Ecological Momentary Assessment.

The first paragraph in the section headed ‘Intervention’ has been revised into two paragraphs to: (a) more clearly describe the key therapeutic elements of the myCompass program (both the self-monitoring and the therapeutic modules) and their relationship, (b) highlight the program’s mobile functionalities, and (c) clarify how feedback is provided.

Comment 148
"...self-monitor moods, symptoms and/or behaviours on their mobile phone or computer." --> What is the core of the treatment program? For me I think that the web based modules distributed via computer are the core of the treatment program. This should be made clear to the reader. Also, when and how did the users rate their mood?

Response:
See Response 42 above.
Comment 149
"...choose dimensions they wish to rate" --> Are dimensions the same as modules?

Response:
The monitoring dimensions are different from the therapeutic modules. See Response 42 above.

Comment 154
154. "...receive and print graphical feedback about their" --> How and when were this feedback received? Did the users received the feedback on their phones? What was the feedback on?

Response:
We have clarified how feedback is provided.

“Users can schedule short message service (SMS) or email reminders to facilitate self-monitoring (frequency of reminders determined by the user); receive and print graphical feedback about their monitoring, including contextual information, on their phone or computer (to monitor change and assist identification of triggers); and elect to receive helpful facts, mental health-care tips or motivational statements by SMS or email.”

Comment 157
"...interactive psychological modules" --> It would be great if you could explain in what way are they interactive.

Response
We have provided additional information about the modules as follows:

“Each module comprises three 10-minute sessions and includes activities that users complete on the computer. Modules also recommend home practice tasks for completion between the weekly sessions to promote skill generalisation”.

Comment 395
395. "...suggest that mobile phones provide an ideal platform for" --> not sure if the mobile phone component was the active program element.

Response
This statement has been amended as follows:

“These results are encouraging, and when the practical benefits of accessibility, anonymity and ease of widespread dissemination are considered, suggest that the combination of mobile phone and internet technology provides an ideal platform for flexible and convenient delivery of mental health care”.

Figure 1
I noticed that there were a lot more attrition lost in myCompass than AC and WL. I think it would be great if this was addressed somewhere.

Response:
The issue of attrition from the intervention group is now reviewed in the Discussion section under the heading ‘Limitations’

References
The paper could use more references

Response:
We have added further references to the paper.

Discretionary revisions

Comment 178
“information was designed to be interesting” --> I am curious how the information was designed to be interesting? And how did you control that people experience the information to be interesting?

Response:
We have amended the statement to read:

“They also received on their mobile phones weekly SMS messages containing brief factual statements about depression, anxiety and stress. The mobile phone statements were also therapeutically inactive, but chosen to ensure that the control program had face validity.”