Author’s response to reviews

Title: An Evaluation of Large Group Cognitive Behaviour Therapy with Mindfulness (CBTm) Classes

Authors:

Vishal Thakur (thakurv@myumanitoba.ca)
Jacquelyne Wong (jwong8@hsc.mb.ca)
Jason Randall (jason.randall@ualberta.ca)
James Bolton (JBOLTON@hsc.mb.ca)
Sagar Parikh (parikhsa@med.umich.edu)
Natalie Mota (nmota@mymts.net)
Debbie Whitney (dwhitney@deerlodge.mb.ca)
Joshua Palay (umpalay@myumanitoba.ca)
Jolene Kinley (jolene.kinley@gmail.com)
Simran Diocee (Simran.Diocee@umanitoba.ca)
Tanya Sala (tsala@hsc.mb.ca)
Jitender Sareen (JSareen@hsc.mb.ca)

Version: 1 Date: 07 Mar 2019

Author’s response to reviews:

Editor Comments:

Having read the paper, I would like the authors to address the following comments in addition to those of the reviewers, noting that some of my comments overlap with those of other reviewers, in which case the authors can cross-refer the reader to their respective response.

1. More detail regarding the statistical analyses is required. I gather that the “mixed-effect linear regressions” were linear mixed models? If so, please report:

   a. The covariance structure used and the estimation method (maximum likelihood or restricted maximum likelihood).
This is now reported (Methods, Page 11): Maximum likelihood estimation was used in this model.

b. If it was “multi-level” (paragraph 2, page 10), please describe the hierarchical nature of the data (e.g., “the six timepoints were clustered within individual”). As Reviewer 1 stated, the question of whether group membership was an additional level in the model needs to be reported and there is potential value in including this in the model.

This is now reported (Methods, Page 11): The model allowed the data to be analyzed as a within-subject design to examine the change in scores across repeated outcome measures in the same subject (up to six time-points were clustered within individual). The individual level was the only level in the model as clients could not be clustered by groups or treatment setting given the data available.

As for the question of group membership and why we could not incorporate it into the model, we have the following explanation: Although we agree that Delgadillo et al.’s finding that "3.6% of variance in post-treatment anxiety scores was explained by variability between SC groups” is important, and should be added to the model in future studies, we unfortunately are unable to include it in the current study given the way our clients attended class, which subsequently affected data entry. Specifically, clients were welcome to attend classes in any order that they saw fit, which means they could have fallen into any class grouping if they, for whatever reason, could not neatly attend one iteration of the 4 session intervention. Thus the data was entered by ‘time-point’ for each client, in chronological order, and was not separated by class groupings. Again, as stated above, we will certainly consider accounting for this important variable by entering future data in a way that separates clients by class groupings.

c. Was time treated as continuous or categorical and what were the fixed and random effects in the model?

This is now reported (Methods, Page 11): The fixed effects were the number of classes (a continuous variable), the effect of the gap between baseline and the first attended class (a binary dummy variable that is coded 1 after the gap has occurred), and the effect of the gap between the last attended class and follow-up (a binary dummy variable that is coded 1 starting with follow-up) on GAD-7 and PHQ-9 scores. These same variables were used as random effects in the model as these effects are likely to vary between individuals.

d. Also, time between baseline assessment and session and between session 4 and follow-up was controlled for. This makes sense, but the reader is unclear how this was managed in the model (especially in the context of other time-relevant variables).

Time variables were not extracted on an individual basis in our analysis, thus the statement that time was controlled for was removed. As is stated in our response to the previous
reviewer comment, the gap variable was treated as a binary dummy variable which was coded as either 0 or 1.

e. As identified by Reviewer 3, there needs to be more detailed explanation of the other variables which were included in the analysis. Providing the key regression results tables would allow the reader to get a sense of this.

Other variables were not included in the analysis and this is now reported (Methods, Page 11): Other variables (sex, education, and mental health diagnoses) were not included or adjusted for in this analysis as they did not vary within individuals.

Ultimately, we did not think the other variables we measured would confound the regression because of the within-subject design and lack of change in those variables during the brief treatment window.

2. The authors note changes on key outcomes between pretreatment assessment and the first session of CBTm and between session 4 of CBTm and follow-up. It makes sense that this duration was statistically controlled for, but what was the mean (SD) duration of these intervals and could the authors report the magnitude of these changes (even if just mean change on GAD-7 and PHQ-9 scores).

The durations have now been included in the Methods section (Page 8): The mean duration between the intake assessment and the first attended CBTm class was 4 weeks and the median was 2 weeks. The mean duration between the last attended class and follow-up was 12 weeks and the median was 9 weeks

The magnitude of the changes in these time intervals is reported in Table 2 (e.g. mean GAD-7 score was reduced by 0.98 between intake assessment and first attended CBTm class)

3. Given the variation between the various schools of mindfulness, additional description of the nature of the mindfulness intervention is needed. Was it taught from a Vipassana or Zen perspective etc and what were the mindfulness exercises (breathing?, body scan?) was there any yoga incorporated? Were there homework exercises?

Updated the following paragraphs in Methods section (Pages 6-8): Classes were 90 minutes in length and ranged in size from 10 to 41 clients (M = 24, SD = 7) per session, not including clients’ partners, family members, or friends who were encouraged to attend. Sessions were led by two staff psychiatrists who received training in CBT and mindfulness. Occasionally other facilitators, such as medical students and residents, assisted in running the classes. There was minimal interaction between facilitators and clients, with interactions often limited to
discussions of homework and answering clients’ questions or concerns. In terms of homework, clients were encouraged to practice mindfulness meditation for five minutes, twice per day, and to partake in at least two other activities brought up in class. These homework activities included, but were not limited to, accessing self-help websites, setting goals, thought records, mood tracking, and physical exercise such as attending a yoga class. Unlike formal CBT programs, there were no individually tailored activities or specific feedback given to clients as they were progressing through homework activities and applying CBT skills. Class content was structured as follows:

Class 1: introduction and outline of the course, rules and expectations, self-help resources, mindfulness exercise, introduction to the cognitive behavioural framework, cognitive distortions, thought records, and homework.

Class 2: mindfulness exercise, review of homework, basics of behaviour therapy, exposure therapy, goal setting, and homework.

Class 3: mindfulness exercise, review of homework, discussion of healthy living, sleep hygiene, and homework.

Class 4: mindfulness exercise, review of homework, anger management strategies, assertiveness training, self-compassion, problem solving, and homework.

The four mindfulness meditations used in the CBTm classes were derived from those taught within Mindfulness-Based Stress Reduction (26) which is a transdiagnostic intervention that has proven useful with stress, depression, and anxiety management (27–29). Specifically, the meditations were body scan, awareness of the breath, awareness of the five senses and loving-kindness. These were introduced in the same order as followed within MBSR. Clients were encouraged to download the no-cost app, MindShift, which has recorded instructions for both body scan and awareness of breath. They could also seek out recorded instructions for the other meditations but no specific direction was given for this.

4. One of the key characteristics of this study is the use of a large group format. I would encourage the authors to further describe the way this was administered and what the key differences are from individual/small group approaches. For example, the skills focus of standard CBT often involves individually tailored activities and feedback from the therapist to an individual on how they’re applying skills etc. Were any of these elements sacrificed for the present intervention, and if so, do your data indicate that it mattered?

See response to comment 3 above. Detail regarding the structure, content and administration of the large-group CBTm classes is now provided in the Methods section (Pages 6-8). Although this was not explicitly contrasted to small group formats, we feel we provide the
reader with sufficient information about our large group intervention so that they may infer key
differences from small group or individual therapy which they are likely more familiar with.

In regards to individually tailored activities and feedback, we added the following
statement (Methods, Page 7): Unlike formal CBT programs, there were no individually tailored
activities or specific feedback given to clients as they were progressing through homework
activities and applying CBT skills.

In regards to whether our data indicated that it mattered, our results suggest that clients
found classes acceptable (they found them useful and >90% indicated they would attend another
class) and, although not clinically significant, we found a statistically significant improvement in
anxiety and depressive symptoms for clients who attended class. These results suggest the CBTm
classes are an acceptable strategy to engage and maintain the interest of clients seeking CBT but
it is unclear whether there is a meaningful improvement in anxiety and depressive symptoms. A
limitation of our study is that we were unable to benchmark the outcomes of our intervention
with other interventions (such as formal CBT group therapy). Future studies should directly
compare large group interventions with conventional CBT programs to offer a clearer picture on
the acceptability and effectiveness of large group CBT interventions.

5. A significant limitation which requires mentioning is that two of your key outcome measures
(acceptability and whether they would attend a further session) were single item measures of
indeterminate reliability.

Added following statement to Discussion (Page 18): Other limitations include the
indeterminate reliability of using single item measures to assess acceptability of the CBTm
classes

BMC Psychiatry operates a policy of open peer review, which means that you will be able to see
the names of the reviewers who provided the reports via the online peer review system. We
encourage you to also view the reports there, via the action links on the left-hand side of the
page, to see the names of the reviewers.

Reviewer reports:

Alice Tickell (Reviewer 1): Thank you for the opportunity to review your manuscript, titled 'An
Evaluation of Large-Group Cognitive Behaviour Therapy with Mindfulness (CBTm) Classes for
Mood and Anxiety Disorders'. Overall, the paper has many strengths such as its large sample size
and high external validity. I have a few comments that I hope you can address to clarify the study
rationale and the conclusions drawn from the data. I have also made some suggestions for
additional measures you could include to strengthen the paper.
Introduction:

(1) CBTm is a new and independently-developed intervention that readers may not be familiar with. Some more background information on the CBTm programme would be welcome. For instance, you mention that previous work from your group has used a 2-session CBT class for people with anxiety disorders, but now you are expanding by using a new 4-session CBTm class for people with anxiety and/or mood disorders, but the logic of this progression was not explained.

I suggest you state more explicitly:

* What was the decision behind developing a new intervention? I.e. why did you not use an existing programme (e.g. Stress Control)? What does the development of CBTm add to the field? What was the rationale for including mindfulness?

Added significant detail regarding the rationale behind developing CBTm and the theory behind mindfulness and CBT (Background, Pages 4-5): Large-group CBT was introduced at a tertiary care clinic in Winnipeg, Canada in 2014 to manage the problem of persistently long wait times. These transdiagnostic 2-session CBT classes were rated useful by clients, led to modest improvements in anxiety symptoms, and reduced wait-times from approximately one year to three months (20). Given these promising findings, the CBT classes were expanded to 4 sessions and introduced mindfulness principles within the core content. These 4 session transdiagnostic Cognitive Behaviour Therapy with Mindfulness (CBTm) classes were independently developed and administered at the clinic to provide clients with basic CBT principles, mindfulness skills, and various self-help resources at a time where they otherwise may not have had access to therapy.

Mindfulness is the process of being nonjudgmentally aware of the present moment, including one’s thoughts, sensations and environment, while encouraging inquisitiveness, open observation, and acceptance (21,22). Evidence suggests that mindfulness-based interventions are effective in treating anxiety and depression (22) and perform comparably to CBT (23), but limited work has looked at mindfulness within the context of CBT (24,25). It is thought one is better able to regulate their emotions and thoughts after entering a relaxed and attentive state with mindfulness, thus making it easier to apply CBT skills; this, in turn, should lead to an improvement in anxiety and depressive symptoms (24,25). To our knowledge, there is no research on brief, low intensity, large group CBT interventions that incorporate mindfulness in the literature. Thus, the current study sought to evaluate the 4-session CBTm class intervention in a Canadian population.
* Precisely what clinical population was it was developed for? The inclusion criteria for your study was the presence of a mental health diagnosis - is this a transdiagnostic intervention? Or is it, as your title suggests, specifically for people with mood/anxiety disorders?

CBTm is a transdiagnostic intervention – this was made clearer in the Background (see paragraph in response to first comment above). Although we chose to focus on anxiety and depressive symptoms in our study (as a significant majority of clients had a primary mood or anxiety disorder), the intervention is most accurately described as transdiagnostic. Table 1 was also updated to include a more detailed list of the primary diagnoses of our sample.

* What was the target of treatment (the mechanism)?

   Added the following statement (Background, Page 5): It is thought one is better able to regulate their emotions and thoughts after entering a relaxed and attentive state with mindfulness, thus making it easier to apply CBT skills; this, in turn, should lead to an improvement in anxiety and depressive symptoms (24,25).

* What are the expected outcomes?

   Our expected outcomes for the CBTm intervention were (Background, Pages 5-6): The two primary outcomes were: (a) acceptability and retention rates of CBTm classes and (b) clients’ change in anxiety and depressive symptoms as a result of attending CBTm classes. Recent UK studies demonstrated that similar large-group CBT interventions are efficient, well tolerated, and effective in treating symptoms of anxiety and depression (26–28). Thus, we hypothesized the CBTm classes would replicate these findings by being acceptable, both in terms of client feedback and retention rates, and lead to improvements in anxiety and depressive symptoms.

* You mention that it is part of a stepped care model: What are all the steps?

   The stepped care model really just involves the following two steps (Methods, Page 8): Following completion of 4 classes, clients were welcome to repeat classes as “booster sessions” or proceed to conventional CBT group therapy if more intensive treatment was required.

   So step 1 is attendance in at least 4 CBTm classes and then this is followed by step 2 which is conventional CBT group therapy (if needed)
Providing this information will give a clearer rationale for your hypotheses. E.g. If the treatment explicitly targets the cognitions that maintain mood/anxiety disorders, then one would expect a reduction in anxiety/depression symptoms… If the purpose is just to facilitate and maintain interest in pursuing CBT, then perhaps one would not expect an effect on symptoms, and the main outcome would be acceptability/engagement.

(2) Linked to the first point, please state clearly the aims of your study, and what your primary/secondary outcomes are. Currently, in the Introduction you state that you have two primary outcomes (acceptability and retention rates, and clients' change in anxiety and depression symptoms), but in the Abstract you state that changes in anxiety and depression symptoms are a secondary outcome. Therefore, it is not clear which outcomes are primary and which are secondary.

Both are primary outcomes. Updated sentence in Abstract (Page 3): Primary outcomes were (a) acceptability of the classes and retention rates and (b) changes in anxiety and depressive symptoms using Generalized Anxiety Disorder 7-item (GAD-7) and Patient Health Questionnaire 9-item (PHQ-9) scales.

(3) In the final paragraph, the rationale for the current study is outlined as to address three gaps in the extant literature. To make your study rationale stronger, instead, state the reason why your study needs to be done, i.e. how it furthers the field, instead of how it fills a gap. Sometimes gaps exist for a reason!

Updated Background (Page 5): To our knowledge, there is no research on brief, low intensity, large group CBT interventions that incorporate mindfulness in the literature. Thus, the current study sought to evaluate the 4-session CBTm class intervention in a Canadian population.

(4) When presenting your own hypotheses, use the past tense, e.g. 'We sought to expand these findings… Based on our previous work, we hypothesised…'

Hypotheses now presented in past tense

Methods:

(1) Please could you clarify certain aspects of the intervention:

* What was the level of interaction with fellow attendees/facilitators?
Following added to Methods (Page 7): There was minimal interaction between facilitators and clients, with interactions often limited to discussions of homework and answering clients’ questions or concerns.

* How much homework was allocated each week?

Following added to Methods (Page 7): In terms of homework, clients were encouraged to practice mindfulness meditation for five minutes, twice per day, and to partake in at least two other activities brought up in class. These homework activities included, but were not limited to, accessing self-help websites, setting goals, thought records, mood tracking, and physical exercise such as attending a yoga class.

* You mention that there was an average of 30-40 clients per session. Does this include the family/friends who accompanied clients? What was the average overall group size?

No this does not include family and friends, this was made clearer in the Methods (Page 6). After more detailed analysis we found the average class size to be lower: Classes were 90 minutes in length and ranged in size from 10 to 41 clients (M = 24, SD = 7) per session (not including clients’ partners, family members, or friends who were encouraged to attend).

(2) Could you benchmark the outcomes from this intervention against other interventions in the stepped care model at this service, e.g. the small group CBT? It would strengthen the manuscript if there was some comparison data.

A limitation of this study is that we do not have comparison data available from other interventions in our region, such as the small group CBT. The only study that we have from our service is Palay et al. (2017) which looked at the feasibility of a similar 2-session CBT intervention for anxiety disorders. We state the following in our background (Page 5): Large-group CBT was introduced at a tertiary care clinic in Winnipeg, Canada in 2014 to manage the problem of persistently long wait times. These transdiagnostic 2-session CBT classes were rated useful by clients, led to modest improvements in anxiety symptoms, and reduced wait-times from approximately one year to three months (20). Given these promising findings, the CBT classes were expanded to 4 sessions and introduced mindfulness principles within the core content. These 4 session transdiagnostic Cognitive Behaviour Therapy with Mindfulness (CBTm) classes were independently developed and administered at the clinic to provide clients with basic CBT principles, mindfulness skills, and various self-help resources at a time where they otherwise may not have had access to therapy.
We agree that having more comparison data would strengthen the manuscript as it would serve as a point of reference to better delineate how the CBTm classes perform in improving anxiety and depression outcomes. We do, however, contrast our findings to relevant UK studies – Delgadillo et al. (2016) & Horrell et al. (2014) – that investigated large group CBT in our Discussion (Page 16-17).

We have now included this as a major limitation of our study which future research should look at including (Discussion, Page 18): It may also be advantageous to benchmark outcomes from the CBTm intervention with other interventions at our service, such as conventional small group CBT. Having this comparison data available would strengthen the conclusions drawn about the relative effectiveness of CBTm.

(3) The PHQ-9 and GAD-7 have clinical cut-offs and criteria to evaluate whether reliable improvements/deteriorations occurred (Evans, Margison, and Barkham, 1998). Using these to report the rate of recovery, reliable improvement/deterioration, and reliable recovery would provide some additional information about the clinical significance of symptom change.

The method to calculate clinically significant change has now been included in our paper. This is reflected in the Methods and Results section (Methods, Page 11): Another method to calculate clinically significant change was applied to the GAD-7 and PHQ-9 in our study (39). According to this methodology, one can compare the mean score and standard deviation of a clinical population with that of a referential “normal” population to determine the extent to which change after treatment is clinically meaningful. We compared the scores from our sample to healthy samples from other studies (33,34). The calculation is:

\[
\frac{\left(\bar{\text{Mean}}_{\text{clin}} \times \text{SD}_{\text{norm}}\right) + \left(\bar{\text{Mean}}_{\text{norm}} \times \text{SD}_{\text{clin}}\right)}{\text{SD}_{\text{norm}} + \text{SD}_{\text{clin}}}
\]

Using this calculation, one would expect clinically significant change if the GAD-7 score change is ≥ 6.52 for anxiety, and if the PHQ-9 score change is ≥ 7.60 for depression.

(Results, Page 14): These findings were also not clinically significant according to Evans et al.’s definition (39). Specifically, the GAD-7 score improvement of 2.40 did not meet the calculated value of 6.52.

(Results, Page 15): These findings were also not clinically significant according to Evans et al.’s definition (39). Specifically, the PHQ-9 score improvement of 1.98 did not meet the calculated value of 7.60.
(4) In your multi-level model, could you take into account the class groupings? Evidence suggests that differences between groups can account for a small proportion of the variance in post-treatment outcomes (Delgadillo et al., 2016).

Although we agree that Delgadillo et al.’s finding that "3.6% of variance in post-treatment anxiety scores was explained by variability between SC groups” is important, and should be added to the model in future studies, we unfortunately are unable to include it in the current study given the way our clients attended class, which subsequently affected data entry. Specifically, clients were welcome to attend classes in any order that they saw fit, which means they could have fallen into any class grouping if they, for whatever reason, could not neatly attend one iteration of the 4 session intervention. Thus the data was entered by ‘time-point’ for each client, in chronological order, and was not separated by class groupings. Again, as stated above, we will certainly consider accounting for this important variable by entering future data in a way that separates clients by class groupings.

Results:

(1) In the sections on Anxiety Symptoms and Depressive Symptoms, include whether the symptom changes are clinically significant or not according to McMillan et al.’s (2010) criteria.

This paragraph has been moved from the discussion section to the results section

Discussion:

(1) As mentioned above, move the information about McMillan et al. (2010)'s criteria to the Results section. This is an indicator that you calculated, so it is better placed there.

This paragraph has been moved from the discussion section to the results section

(2) On line 28, you state: 'It is important to highlight that we expected dropout…' If this was a hypothesis/expectation, then move it to the Introduction section. Placed in the Discussion, it seems like a post-hoc explanation for your result. Here it is enough to state that your rate of drop-out is consistent with other similar large-group interventions found in the extant literature.

This statement was removed to offer a more balanced interpretation of the results. We limited our analysis of the dropout rate to a comparison with the literature (Discussion, Page 15): The CBTm classes had a dropout rate of 26.3%, which is consistent with similar large group CBT interventions found in the literature (23-25). More than half (53.8%) of eligible clients (those that completed the 4-session intervention) moved onto conventional small group CBT
therapy. Future research should investigate long-term dropout rates and specific reasons for dropout as these were not assessed in this study.

Specific variables that were associated with dropout were discussed separately.

(3) On line 33, you state: 'We anticipated some clients would achieve symptom improvement or remission at earlier stages and not require the 'complete' 4-session intervention'. Do you have any evidence that people who dropped out achieved symptom improvement or remission to back up this claim? If you do not have the evidence to back this up, then remove this sentence. On the contrary, your results suggested that it was participants with more severe depression symptoms at baseline who dropped out.

This statement was removed

(4) The effect sizes in your study are below efficacy benchmarks for guided self-help interventions and are found to be not clinically significant according to the McMillan’s criteria that you use. However, I am not sure the subsequent conclusions you draw from this result are adequately supported:

* Your interpretation of the lower effect sizes compared to benchmarks/extant literature is: 'In regards to length, four 90 minute sessions is brief when compared to formal CBT programs and the extant literature on large-group CBT discussed above.' (line 48). However, CBTm might actually be similar in length or longer than the intervention used by Horrell et al. (2014) if you take into account homework time. As mentioned above, it would be helpful to know how much homework was allocated.

This was updated to the following in the Discussion (Page 16): These findings may be partially explained by the length of the CBTm intervention as four 90 minute sessions is brief when compared to formal CBT programs (10) and Delgadillo et al.’s study (28). Dosing of psychotherapy is an important factor which influences efficacy, thus the brevity of our intervention makes the smaller effect sizes understandable. However, Horrell et al.’s trial (26) had higher effect sizes than CBTm, despite being similar in length, which suggests other factors may better explain our study’s lack of clinical significance.

* Another possible interpretation that you do not mention is that the lower effect sizes may reflect that CBTm is less effective than the interventions tested in the extant literature.
Added following to Discussion (Page 16): It is also possible that the CBTm intervention is simply less effective at treating anxiety and depression than the interventions tested in the extant literature.

* Another possible interpretation is that studies in the extant literature have limited their sample to those with an anxiety disorder or mood disorder, whereas your sample included those with any mental health diagnosis (10% did not have a primary diagnosis of anxiety/mood disorder).

Added following to Discussion (Page 16): One possible explanation is that our sample included a broad range of mental health diagnoses, whereas the extant literature limited their sample to those with depression (26) or a mix of anxiety and depression (28).

* You mention that local work is being done to expand the availability of CBTm classes across multiple sites and different settings. This conclusion seems premature, considering your study suggests that there may not be strong evidence for its effectiveness/clinical impact. I would question whether more work needs to be done first to test CBTm in controlled trials, before disseminating extensively. It is possible that there are more effective alternatives. Refer to Onken, Carroll, Shoham, Cuthbert & Riddle (2014) for a model of intervention development.

This statement has now been removed to make it clearer to the reader that we are still investigating the effectiveness of the CBTm intervention in our region before disseminating extensively. We are collecting valuable data from a few services running the CBTm classes and we plan to investigate relevant outcomes across these different settings to offer a clearer picture of its effectiveness. Our results from this study seem to indicate that clients find CBTm acceptable, but, as we discuss, it is still unclear as to the clinical impact. Our hope is that future research at our services will make this clearer.

(5) In relation to your finding that people with greater depression severity and low education were more likely to drop out: it would be interesting if this were discussed in relation to the format. You could speculate what it is about the large-group format that makes it less accessible to these groups: E.g. the focus on independent study/lecture format may be more challenging for those with a lower education; there may be motivational challenges for those with more severe depression symptoms. You could draw out the implications more fully here. What kinds of interventions might be more suitable for these groups?

Added following to Discussion (Page 16-17): In our study, clients with more severe depression or lower education were more likely to dropout, which may also have impacted the intervention’s effectiveness. One possible explanation for the dropout, initially speculated by
Fernandez et al. (41), is that the major symptomology of depression (namely diminished interest, poor concentration, lower energy, hopelessness, social withdrawal) may make it more difficult to engage and maintain the interest of this population. This would seem especially relevant in a large-group setting where clients require significant motivation and attention to both attend class and to adequately absorb information delivered in a didactic fashion. In regards to lower education and dropout, it is well known that these two variables are associated in psychotherapy (44), but it is less clear as to why. It is possible that lack of insight into their mental illness, differences in expectations of therapy, or difficulty understanding class content played a role. What is clear is that future practitioners delivering low intensity CBT should implement strategies to more effectively engage and treat those clients with more severe depression or lower education. Perhaps small group or individual psychotherapy, with more personalized content for specific symptom or education levels, is an appropriate approach to treatment in these vulnerable populations.

Shian-Ling Keng (Reviewer 2): This manuscript reports findings from a chart review of patients enrolling in a large, 4-session CBTm course administered in a university setting. The manuscript is well-written overall, and the findings are meaningful in that they provide data on the effectiveness and acceptability of a CBT course delivered in a real world setting. The analyses conducted were also rigorous, and I appreciate the fact that the authors analyzed the clinical significance of their findings. Below are a few suggestions to strengthen the manuscript as a whole:

* Even though the paper focuses on the evaluation of a CBT+mindfulness course, the introduction section did not provide any background information about the concept, practice, or evidence base related to mindfulness. Incorporating this information would help readers understand the rationale for delivering a course that integrates elements of CBT and mindfulness. Following information is now included in the Background (Page 5): Mindfulness is the process of being nonjudgmentally aware of the present moment, including one’s thoughts, sensations and environment, while encouraging inquisitiveness, open observation, and acceptance (21,22). Evidence suggests that mindfulness-based interventions are effective in treating anxiety and depression (22) and perform comparably to CBT (23), but limited work has looked at mindfulness within the context of CBT (24,25). It is thought one is better able to regulate their emotions and thoughts after entering a relaxed and attentive state with mindfulness, thus making it easier to apply CBT skills; this, in turn, should lead to an improvement in anxiety and depressive symptoms (24,25). To our knowledge, there is no research on brief, low intensity, large group CBT interventions that incorporate mindfulness in the literature. Thus, the current study sought to evaluate the 4-session CBTm class intervention in a Canadian population.
On this note, did other studies that have evaluated the effects of large scale CBT courses also include elements of mindfulness in the course(s)? If so, how did the data of those studies compare to the findings of this study?

To our knowledge, CBTm is the first large group CBT intervention to incorporate mindfulness

In the methods section, include information about the study's inclusion and exclusion criteria in text, as opposed to in the figures section.

Inclusion and exclusion criteria has been moved to the text

Also, more information on how diagnosis was established would be helpful. Did the clinicians use any standardized tools to establish the diagnosis of patients enrolled in the study? If not, this should probably be noted as a limitation of the study.

We noted that a standardized diagnostic tool was not used and this was also listed as a limitation

Minor note: Spell out GAD-7 and PHQ-9 fully at their first mention in text.

We have updated this

It will also be helpful to provide more information on the modality of the course delivered. Was it primarily psychoeducational and didactic? To what extent was the course experiential? What were the types of mindfulness exercises taught in the course? Were participants given any homework in between sessions?

Added the following to the Methods section (Page 6-7): Classes were 90 minutes in length and ranged in size from 10 to 41 clients (M = 24, SD = 7) per session, not including clients’ partners, family members, or friends who were encouraged to attend. Sessions were led by two staff psychiatrists who received training in CBT and mindfulness. Occasionally other facilitators, such as medical students and residents, assisted in running the classes. There was minimal interaction between facilitators and clients, with interactions often limited to discussions of homework and answering clients’ questions or concerns. In terms of homework, clients were encouraged to practice mindfulness meditation for five minutes, twice per day, and to partake in at least two other activities brought up in class. These homework activities included, but were not limited to, accessing self-help websites, setting goals, thought records, mood tracking, and physical exercise such as attending a yoga class. Unlike formal CBT programs, there were no
individually tailored activities or specific feedback given to clients as they were progressing through homework activities and applying CBT skills. Class content was structured as follows:

Class 1: introduction and outline of the course, rules and expectations, self-help resources, mindfulness exercise, introduction to the cognitive behavioural framework, cognitive distortions, thought records, and homework.

Class 2: mindfulness exercise, review of homework, basics of behaviour therapy, exposure therapy, goal setting, and homework.

Class 3: mindfulness exercise, review of homework, discussion of healthy living, sleep hygiene, and homework.

Class 4: mindfulness exercise, review of homework, anger management strategies, assertiveness training, self-compassion, problem solving, and homework.

The four mindfulness meditations used in the CBTm classes were derived from those taught within Mindfulness-Based Stress Reduction (29) which is a transdiagnostic intervention that has proven useful with stress, depression, and anxiety management (30–32). Specifically, the meditations were body scan, awareness of the breath, awareness of the five senses and loving-kindness. These were introduced in the same order as followed within MBSR. Clients were encouraged to download the no-cost app, MindShift, which has recorded instructions for both body scan and awareness of breath. They could also seek out recorded instructions for the other meditations but no specific direction was given for this.

Following completion of 4 classes, clients were welcome to repeat classes as “booster sessions” or proceed to conventional CBT group therapy if more intensive treatment was required, which is part of a larger stepped care model. Verbal or written consent from each client was not required to perform the chart review. A corresponding website (cbtm.ca) was also developed in which class materials and handouts could be accessed for free.

* As the authors noted, one of the key rationale for exploring the delivery of CBT in the format of large group courses is to reduce wait time for treatment. Can the authors provide more specific information with regards to the amount of wait time that can be reduced based on past research (or their own past research)?

Added the following to the Background section (Page 5): Large-group CBT was introduced at a tertiary care clinic in Winnipeg, Canada in 2014 to manage the problem of persistently long wait times. These transdiagnostic 2-session CBT classes were rated useful by clients, led to modest improvements in anxiety symptoms, and reduced wait-times from approximately one year to three months (20).
The current study’s wait times from the intake assessment to the first attended CBTm class was also reported (Methods, Page 8): The mean duration between the intake assessment and the first attended CBTm class was 4 weeks and the median was 2 weeks.

* In the Discussion section, in "As Delgadillo and colleagues discuss", "discuss" should be written in a past tense.

Changed to “discussed” in the Discussion section (Page 18)

* One of the notable findings of the study is that lower education and higher depression severity at baseline predict a greater likelihood of dropping out from the course. Given that this appears to be a trend from a review of past studies, can the authors provide more specific ideas / suggestions on how one might address this issue in future iteration of such courses?

Added the following to the Discussion section (Page 17-18): In our study, clients with more severe depression or lower education were more likely to dropout, which may also have impacted the intervention’s effectiveness. One possible explanation for the dropout, initially speculated by Fernandez et al. (41), is that the major symptomology of depression (namely diminished interest, poor concentration, lower energy, hopelessness, social withdrawal) may make it more difficult to engage and maintain the interest of this population. This would seem especially relevant in a large-group setting where clients require significant motivation and attention to both attend class and to adequately absorb information delivered in a didactic fashion. Interestingly, there was a significant increase in depressive symptoms in the treatment gap between the last attended CBTm class and follow-up. This demonstrates the importance of continual engagement and treatment for this population as our results seem to suggest they are prone to relapse after disengaging from the CBTm intervention for some time. In regards to lower education and dropout, it is well known that these two variables are associated in psychotherapy (44), but it is less clear as to why. It is possible that lack of insight into their mental illness, differences in expectations of therapy, or difficulty understanding class content played a role. What is clear is that future practitioners delivering low intensity CBT should implement strategies to more effectively engage and treat those clients with more severe depression or lower education. Perhaps small group or individual psychotherapy, with more personalized content for specific symptom or education levels, is an appropriate approach to treatment in these vulnerable populations.
Lena Wimmer (Reviewer 3): Thank you very much for providing me with the opportunity to review this paper. It reports an ex post facto analysis of patients who received 4-session Cognitive Behavior Therapy with mindfulness (CBTm) in a large group in order to gain insight into the acceptability and effectiveness of this approach. Considering the practical relevance of the topic, the manuscript has the potential to significantly contribute to the literature despite methodological limitations regarding the study design. In general, I find the paper to be well written. Nevertheless, in my view major changes are required before the manuscript can potentially be accepted for publication. My main concerns stem from the facts that 1) no rationale is given for using the particular therapy program under investigation, more precisely regarding the inclusion of mindfulness in addition to CBT, 2) the analytical approach is not described with sufficient clarity in order that readers could, for instance, reproduce the calculations, 3) parts of the conclusions, particularly regarding the effectiveness of CBTm, are not justified given the study design.

Detailed comments

Abstract:

p. 3 l. 16: "adult outpatients with a mood or anxiety disorder" is apparently not entirely true because 10.7% of participants were given a primary mental health diagnosis that was not a mood or anxiety disorder. Please correct.

This has been corrected. We no longer specify a particular diagnosis for our sample and instead describe this intervention as “transdiagnostic”. Table 1 has been updated to include the major diagnoses in our sample.

p. 3 l. 50: "The classes are also effective in reducing symptoms of anxiety and depression": This is an unjustified conclusion. It is true that participants who received CBTm on average showed reduced symptoms of anxiety and depression between baseline and follow-up. However, these changes cannot be traced back to CBTm due to the lack of control groups and associated lack of random assignment to treatment conditions.

p. 3 l. 51 f.: "The clinical significance of symptom improvement remains unclear": Considering the finding that the reductions of anxiety and depression were not clinically significant (e.g., p. 14 2nd paragraph) it would seem more appropriate to say that symptom improvement was not clinically significant.

For the above two points we changed it to the following in the Abstract section (Page 3): Clients attending CBTm classes experienced improvements in anxiety and depressive symptoms. Symptom improvement was not clinically significant.
Background:

It becomes clear why large-group CBT was investigated, yet no rationale is given why mindfulness was added to CBT. Furthermore, a well-established 8-week intervention program that combines CBT and mindfulness is Mindfulness-Based Cognitive Therapy (MBCT). Therefore I was surprised to see that, provided that mindfulness should, for whatever reason, be added to CBT, an apparently new approach, CBTm, was introduced rather than MBCT adapted to the desired 4-session format.

More detail regarding the rationale for developing the CBTm classes was provided in the Background: Large-group CBT was introduced at a tertiary care clinic in Winnipeg, Canada in 2014 to manage the problem of persistently long wait times. These transdiagnostic 2-session CBT classes were rated useful by clients, led to modest improvements in anxiety symptoms, and reduced wait-times from approximately one year to three months (20). Given these promising findings, the CBT classes were expanded to 4 sessions and introduced mindfulness principles within the core content. These 4 session transdiagnostic Cognitive Behaviour Therapy with Mindfulness (CBTm) classes were independently developed and administered at the clinic to provide clients with basic CBT principles, mindfulness skills, and various self-help resources at a time where they otherwise may not have had access to therapy.

Mindfulness is the process of being nonjudgmentally aware of the present moment, including one’s thoughts, sensations and environment, while encouraging inquisitiveness, open observation, and acceptance (21,22). Evidence suggests that mindfulness-based interventions are effective in treating anxiety and depression (22) and perform comparably to CBT (23), but limited work has looked at mindfulness within the context of CBT (24,25). It is thought one is better able to regulate their emotions and thoughts after entering a relaxed and attentive state with mindfulness, thus making it easier to apply CBT skills; this, in turn, should lead to an improvement in anxiety and depressive symptoms (24,25). To our knowledge, there is no research on brief, low intensity, large group CBT interventions that incorporate mindfulness in the literature. Thus, the current study sought to evaluate the 4-session CBTm class intervention in a Canadian population. We conducted a retrospective chart review of clients who attended classes between 2015 and 2016. The two primary outcomes were: (a) acceptability and retention rates of CBTm classes and (b) clients’ change in anxiety and depressive symptoms as a result of attending CBTm classes. Recent UK studies demonstrated that similar large-group CBT interventions are efficient, well tolerated, and effective in treating symptoms of anxiety and depression (26–28). Thus, we hypothesized the CBTm classes would replicate these findings by being acceptable, both in terms of client feedback and retention rates, and lead to improvements in anxiety and depressive symptoms.
Methods:

Instructors were reported to be formally trained in CBT (p. 6 l. 41). What was their qualification in mindfulness?

The statement now reads (Methods, Page 6): Sessions were led by two staff psychiatrists who received training in CBT and mindfulness.

p. 7. l. 51 f. "This treatment gap was controlled for in relevant analyses": How was this done? For instance, was the number of days between session 4 and the 1st follow-up session included as an independent variable or covariate etc.? Please provide more details.

Time variables were not extracted on an individual basis in our analysis, thus the statement that time was controlled for was removed. As is stated in our response to the previous reviewer comment, the gap variable was treated as a binary dummy variable which was coded as either 0 or 1.

p. 8 l. 24 f.: "Participants' self-reported acceptability of the CBTm classes was assessed using two items from the evaluation form they completed immediately after each session". Why were these two items selected and what were the remaining items of the evaluation form?

Added the following to the Methods (Page 9): These two items were chosen as they are easy to report and are completed by most clients. Remaining items on the evaluation form were not chosen for this study as they were often left blank by clients; these three qualitative items ask the respondent to (1) list three things they learned (2) describe what they like about the session and (3) describe what could be improved. Future studies may wish to use formal measures of satisfaction as these were not implemented in the current study.

Please provide references for the GAD-7 and the PHQ-9, and please explain abbreviations when they are mentioned in the text for the first time.

Full names of the scales included when mentioned in Methods section for the first time. References included in the “Anxiety and Depressive Symptoms” sub-section as this paragraph goes into further detail regarding the measures.

The order in which the steps of the analytical approach is presented does not match the order that is given in abstract, introduction, and results. For instance, the results section reports
acceptability before changes in anxiety and depressive symptoms, in the analytic strategy section the order is reversed. I think the article would be easier to read if the order was kept the same throughout.

We are unclear as to which section has the changes in anxiety and depressive symptoms before acceptability. We organized it so that the general flow of each section is Acceptability --> Anxiety and Depressive symptoms. The Analytic Strategy section is limited to an explanation of how we performed the analysis to determine changes in anxiety and depressive symptoms.

p. 9 l. 48: "Primary analysis used a multi-level mixed-effects linear regression model": Which levels were implemented in the model? As this was a mixed-effects model, which factors were random and which were fixed?

This is now reported (Methods, Page 11): The fixed effects were the number of classes (a continuous variable), the effect of the gap between baseline and the first attended class (a binary dummy variable that is coded 1 after the gap has occurred), and the effect of the gap between the last attended class and follow-up (a binary dummy variable that is coded 1 starting with follow-up) on GAD-7 and PHQ-9 scores. These same variables were used as random effects in the model as these effects are likely to vary between individuals.

p. 9 l. 56 f.: "The model controlled for the time between baseline and class 1, as well as treatment gap between class 4 and follow-up during the first group session": How was this implemented in the model?

Time variables were not extracted on an individual basis in our analysis, thus the statement that time was controlled for was removed. As is stated in our response to the previous reviewer comment, the gap variable was treated as a binary dummy variable which was coded as either 0 or 1.

p. 10 l. 27 f.: "The variables of interest were regressed against a binary variable indicating completion of at least 4 classes": This sounds to me as if the completion criterion was participation in at least 4 classes, however in the results section the completion criterion is reported to be participation in at least 3 classes (p. 10 l. 51). Please clarify.

For the purposes of calculating a dropout rate we had initially changed the definition of dropout to be anyone who completed < 3 classes. However, we see how it is both confusing and improper to have the definition of dropout be different in the analysis (where we had dropout as
anyone completing < 4 classes) than in our attendance calculation. Thus, we have changed the definition of dropout to be < 4 classes in both the analysis and in our calculated dropout rate. The new dropout rate is 37.5%, which is still consistent with previous literature. This has also been updated in Figure 1.

Results:

p.11 l. 19 f.: "Two significant baseline predictors of CBTm class completion were found after adjusting for other variables": Which variables were adjusted for and how was this achieved?

Other variables were not included in the analysis and this is now reported (Methods, Page 11): Other variables (sex, education, and mental health diagnoses) were not included or adjusted for in this analysis as they did not vary within individuals.

Ultimately, we did not think the other variables we measured would confound the regression because of the within-subject design and lack of change in those variables during the brief treatment window.

p. 11 l. 31 f.: "All other baseline variables (GAD-7 score, sex, type and number of mental health diagnoses) did not significantly predict class completion": Please provide p-values for insignificant tests.

This has been updated: All other baseline variables did not significantly predict class completion: GAD-7 score (p = 0.93), sex (p = 0.92), mental health diagnosis (p = 0.37).

The finding that depressive symptoms increased between class 4 and follow-up (p. 12 l. 29) would be very interesting to elaborate on in the discussion.

The following has been added to the Discussion (Page 17-18): Interestingly, there was a significant increase in depressive symptoms in the treatment gap between the last attended CBTm class and follow-up. This demonstrates the importance of continual engagement and treatment for this population as our results seem to suggest they are prone to relapse after disengaging from the CBTm intervention for some time.

We follow a few sentences later with some suggestions on how this may be remedied (Discussion, Page 18): What is clear is that future practitioners delivering low intensity CBT should implement strategies to more effectively engage and treat those clients with more severe depression or lower education. Perhaps small group or individual psychotherapy, with more
personalized content for specific symptom or education levels, is an appropriate approach to treatment in these vulnerable populations.

Discussion:

As to the conclusions regarding effectiveness and clinical significance, please refer to my comments on the abstract.

Discussion around effectiveness and clinical significance has been significantly updated given the feedback from all reviewers (Page 16-17): To further elucidate whether clients’ symptom improvement was clinically significant, we applied two separate definitions of ‘clinically significant change’ from the literature (38,39). This analysis again demonstrates that clients did not experience clinically meaningful change in either anxiety or depressive symptoms. These findings may be partially explained by the length of the CBTm intervention as four 90 minute sessions is brief when compared to formal CBT programs (10) and Delgadillo et al.’s study (28). Dosing of psychotherapy is an important factor which influences efficacy, thus the brevity of our intervention makes the smaller effect sizes understandable. However, Horrell et al.’s trial had higher effect sizes than CBTm, despite being similar in length, which suggests other factors may better explain our study’s lack of clinical significance. One possible explanation is that our sample included a broad range of mental health diagnoses, whereas the extant literature limited their sample to those with depression (26) or a mix of anxiety and depression (28). It is also possible that the CBTm intervention is simply less effective than these existing interventions at treating anxiety and depression.

The CBTm intervention’s lower effectiveness in treating anxiety and depression may be partially explained by the fact that our sample had greater baseline symptom severity compared to extant literature (26,28). Recent work demonstrates that greater symptom severity is associated with dropout (41) and poorer clinical outcomes (42,43) in psychotherapy. Consistent with this, one service in Delgadillo et al.’s study (28) with greater symptom severity attained a lower effect size (d=0.48) for anxiety relative to the other services. In our study, clients with more severe depression or lower education were more likely to dropout, which may also have impacted the intervention’s effectiveness. One possible explanation for the dropout, initially speculated by Fernandez et al. (41), is that the major symptomology of depression (namely diminished interest, poor concentration, lower energy, hopelessness, social withdrawal) may make it more difficult to engage and maintain the interest of this population. This would seem especially relevant in a large-group setting where clients require significant motivation and attention to both attend class and to adequately absorb information delivered in a didactic fashion. Interestingly, there was a significant increase in depressive symptoms in the treatment gap between the last attended CBTm class and follow-up. This demonstrates the importance of continual engagement and treatment for this population as our results seem to suggest they are prone to relapse after
disengaging from the CBTm intervention for some time. In regards to lower education and dropout, it is well known that these two variables are associated in psychotherapy (44), but it is less clear as to why. It is possible that lack of insight into their mental illness, differences in expectations of therapy, or difficulty understanding class content played a role. What is clear is that future practitioners delivering low intensity CBT should implement strategies to more effectively engage and treat those clients with more severe depression or lower education. Perhaps small group or individual psychotherapy, with more personalized content for specific symptom or education levels, is an appropriate approach to treatment in these vulnerable populations.

p. 13 l. 34f.: "We anticipated some clients would achieve symptom improvement or remission at earlier stages and not require the 'complete' 4-session intervention": This explanation for dropout does not seem consistent with the finding that "participants with more severe baseline depression or lower education were more likely to drop out" (p. 15 l. 5f.), because it is highly unlikely that persons with severe depression experience massive reduction of symptoms before completing 4 sessions of large-group CBTm. I would therefore recommend a more careful and balanced interpretation of dropout rates, also because reasons for dropout were not assessed (p. 13 l. 44f.).

Removed in depth interpretation of dropout rates as this was not a focus of the current study and, we agree, was not a completely balanced explanation. See answer to previous reviewer comment to see our detailed discussion around the finding that those with more severe depression or lower education were more likely to dropout. Future research should certainly focus on dropout as a major focus of study.