PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Feasibility of a blended group intervention (bGT) for major depression: Uncontrolled interventional study in a university setting
AUTHORS	Schuster, Raphael; Fichtenbauer, Isabelle; Sparr, Verena Maria; Berger, Thomas; Laireiter, Anton-Rupert

VERSION 1 – REVIEW

REVIEWER	Philip Batterham The Australian National University, Australia
REVIEW RETURNED	21-Aug-2017

GENERAL COMMENTS	The study reports on an uncontrolled pilot study of blended group therapy with an internet-based component. The sample was small, mostly female, with depression and/or anxiety. I have some concerns about the analyses and interpretation of findings. Major comments: 1. Page 8: LOCF provides biased estimates of outcome, assuming data were missing completely at random (unlikely to be accurate assumption). Why did the authors not used mixed models repeated measures or some other less biased method for the analyses (see e.g., Verbeke & Molenberghs: Linear Mixed Models for Longitudinal Data, 2000). 2. Page 9: If participants were excluded from analyses based on changes in medication/ psychotherapy, the analyses were not ITT. 3. Discussion: The discussion often goes beyond the findings, neglecting the limitation that no control group was included in the study. For example: “...the investigated group treatment might be a suitable and short treatment option for mild to moderate depression.” (p11) - We might see the same patterns of symptom remission in an attentional control condition. “...might best be understood as a catalytic potential of computer and media support in fostering other established treatment factors.” (p12) - We don't know if blending had a catalytic effect on treatment, as there was no comparison to unblended treatment. The authors need to be very careful in discussing the findings appropriately throughout, given the lack of control condition (it is not enough to simply note this methodological shortcoming in the limitations section). 4. Table 5: why are p-values greater than 0.05 and even values greater than 0.1 being interpreted? In addition, it is not clear what comparisons the the p-values represent or what statistics were used.

	Minor comments: 5. It is not clear why a CONSORT checklist (rather than perhaps STROBE) was attached, as the study was not an RCT or even CT. 6. Missing word: “final [year] of their Master studies” (page 6) 7. What was the target sample size and how was this determined? 8. Missing word: “currently undergoing [CBT??]. Psychiatric medication...” (p7) 9. Page 13: “...proves also feasible for less educated or older patients” – males should be added to this list as there were only 5 males in the study. 10. There are minor wording issues throughout – the paper would benefit from thorough proofing. 11. Figure 1 does not show how many participants completed the post-test or 3-month follow-up assessment.

REVIEWER	Mayke Mol GGZ inGeest, the Netherlands
REVIEW RETURNED	31-Oct-2017

GENERAL COMMENTS	The study describes the results of a patient-centred uncontrolled clinical trial for a blended group intervention for people with depression. Overall, I think this is an article that addresses an interesting and upcoming topic. This study also shows that it is possible to recruit participants within a short timeframe. However, the authors could strengthen the paper by explaining the blended rationale behind the treatment in more detail. More specifically, I have the following suggestions: • The research question (aim of the study) does not seem to cover the whole study. The question should be more comprehensive and specific at the same time. There should be a logical flow from research aim, to methodology used, results and discussion. More specific questions would also improve the discussion section, I would advise the authors to relate the questions to the found results in the discussion instead of focussing on other research on blended interventions or subjects such a therapist time. The findings in this study have great value on its own, considering the study design and sample size the conclusions can be toned down a little. • Consider a different/shorter title. • To increase the readability, I think some sentences and words need to be checked by a native speaker (e.g. working factors, battery), checked on correct use of uppercase letters in the text and tables (e.g. E-COMPARED/Minddistrict) and consistent use of words (e.g. seminar, lectures). • The treatment and study drop-out numbers throughout the paper are not consistent/clear, please address this.

	• Abstract: page 2 line 5, the authors probably mean effectiveness instead of effectivity. However, considering that the intervention was provided in a university setting to a community sample external validity of the study is low, efficacy seems therefore more fitting. Given the small sample size I would advise to refrain from using the terms efficacy or effectiveness and focus on feasibility or proof of concept. Line 13: it is unclear what the authors mean with university outpatient setting in a general community setting. This reads as though the university offers psychological care in an outpatient clinic. Please specify in the method section, where treatment was provided, the size of the treatment group (n=26?)and how this relates to standard care. The conclusion that the blended format can improve current group therapy interventions cannot be drawn from this study. • In the objective the authors refer to depression and comorbid anxiety. It remains unclear whether patients with comorbid symptoms were specifically selected and in which way the treatment addresses anxiety. Also there was no outcome measure on anxiety. The authors refer to “Treatment of depression” (page 3, line 46), it should be made clear if the focus is on a reduction of depressive symptoms or on remission of major depressive disorder. • The authors claim that this the first clinical study on blended group therapy for depression, but wasn’t that actually the other paper by Schuster et al (2017)? It is also unclear what the authors mean by the statement of a high fit of the two intervention strategies. • In the introduction some of the statements need more explanation: page 4 line 7 (because…), page 4 line 27 (needs, such as…), page 5 line 17 (because…). • The authors refer to the NICE recommendation for group CBT, however it remains unclear how this relates to the psychoeducation group intervention that is studied. Please specify which elements of the exiting psychoeducational treatment were provided online in the blended format. • Page 6 line 43, I think the authors mean “after group session” instead of “ after treatment”. On which platform was the intervention offered? • page 7 line 9, inclusion and exclusion criteria: specify how severe depression was defined. • Page 8 line 25, CSQ: was satisfaction measured on the blended treatment as a whole, or focused on just the online part? • Page 8 line 53, please explain the use of LOCF instead of a more sophisticated alternative. • Result section, I advise the authors to report the results in a more objective manner, for example ´one participant´ instead of ‘only one participant’ or ´only a small proportion´. • Page 10 line 37 and discussion section, please explain why the unguided discussion forum received a lower rating; did the participants wanted it to be guided? • Page 10 line 50, please explain why blended would fit the needs of treating alike group treatments best. • Discussion, it is unclear what the main message is of the second paragraph and some statements also need references, e.g. more positive appraisals for psychological trainings. • In Table 2 not all the numbers add up to 26. In Table 4 the number of participants is missing.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

LOCF provides biased estimates of outcome, assuming data were missing completely at random (unlikely to be accurate assumption). Why did the authors not used mixed models repeated measures
--> Applied LMM without further use of LOCF: Results improved a bit.

If participants were excluded from analyses based on changes in medication/ psychotherapy, the analyses were not ITT
-->Adopted the text and recalculated the analyses: Two patients reported changes in medication. According to ITT-principles, those patients remained in the analyses.

The discussion often goes beyond the findings:
a) No control condition: For example: "...the investigated group treatment might be a suitable and short treatment option for mild to moderate depression." We might see the same patterns of symptom remission in an attentional control condition.
b) "...might best be understood as a catalytic potential of computer and media support in fostering other established treatment factors." (p12) - We don't know if blending had a catalytic effect on treatment, as there was no comparison to unblended treatment.
--> We think: Even though the design lacks an active control condition, there are good reasons to assume, that an attentional control condition would not have resulted in similar patterns. Still, we acknowledge that other factors restrict observed effects and external validity (researcher allegiance, non-routine setting).
-->We adopted the wording: … From the viewpoint of addressed patients, modern technology and communication techniques add value to existing group treatments and the blended format should be further tested in routine care.
--> We restructured the paragraph and adopted the wording: Regarding the interpretation, patients' appraisals might best be conceptualized as the description of a catalytic effect, possibly fostering other established treatment factors, such as imparting information (Yalom & Leszcz, 2005) or motivational clarification (Grawe, 2004). Even though first findings from comparative studies are promising (c.f. Zwerenz et al., 2017; Berger et al., in press), future research has to determine, if patients' positive appraisals will translate into superior effects of blended treatments for depression.

Table 5: why are p-values greater than 0.05 and even values greater than 0.1 being interpreted? In addition, it is not clear what comparisons the the p-values represent or what statistics were used.
--> Added required information to statistical analyses section: Appraisals of computer components (Table 4) were calculated by t-tests (comparing against gran average), and for the analyses of applicability of blended treatment and process aspects (Table 5) paired t-tests were applied.
--> Toned down wording in article: … assumption is supported by tentatively more positive appraisals …
Why would we like to interpret (one) value above 0.1?:
a) The value (p= 0.103) only exceeds the criterion for tentatively significant results very marginally on the third decimal place,
b) And points in the same direction as the result in line 3 of Table 5. So, additionally to the very marginal exceeding we have support for our interpretation from a second criterion (congruency of data is given)
c) Therefore (from the viewpoint of data) not drawing this conclusion given the small sample size (and reduced power) could also be a too firm / strict interpretation (accepting H0, when there actually is a difference)

d) We do not profit from this result, as we wish to provide blended therapy for a broad clientele. So the finding itself is the more conservative interpretation
e) It fits the results of a forthcoming qualitative article
f) The exact p-value (0.103) is provided and the critical reader is provided with all information to come to a personal conclusion on this issue

It is not clear why a CONSORT checklist (rather than perhaps STROBE) was attached, as the study was not an RCT or even CT
--> We included CONSORT as a requirement from BMJ Open.

What was the target sample size and how was this determined?
-->We provided this information in the text: Power analysis was carried out using G*Power (Faul et al., 2007), resulting in an estimated sample size of N = 22, for a conservative medium within-subjects effect size of d = 0.65 (alpha-error α = .05, power β = .90).

Missing word: “currently undergoing [CBT??]. -->Added: psychotherapy

… proves also feasible for less educated or older patients” – males should be added to this list
--> If you will, we would like to keep this without males, because we also had male participants in prior and later studies and from qualitative interviews we have the impression they liked it (structure, plan and technology) even more.

Figure 1 does not show how many participants completed the post-test or 3-month follow-up assessment.
-->Included missing information in Figure 1

Reviewer 2

Given the small sample size I would advise to refrain from using the terms efficacy or effectiveness and focus on feasibility or proof of concept.
--> Addapted title, wording and paragraphs ("Efficacy")

Line 13: it is unclear what the authors mean with university outpatient setting in a general community setting.
-->Adopted: refrained to effectiveness and more emphasis on proof of concept character in title, discussion and conclusions
-->Specified: Trainings took place in an adopted seminar room of University of Vienna, Faculty of Psychology (Department of Applied Psychology: “Health, Development and Promotion”).
--> Even though this is not a fully equipped outpatient clinic, we think validity is not too low, because the intervention was developed (and later also applied) in Salzburg in a university outpatient clinic. In this context we prepare two more forthcoming articles. (blended ACT based group therapy and a qualitative article on patients’ and therapists’ appraisals)

Please specify in the method section, where treatment was provided, the size of the treatment group (n=26?) and how this relates to standard care.
--> Specified in the method section: Both trainers had prior experience with conducting group therapy in clinical settings. Sessions took place in a seminar room of University of Vienna, Faculty of Psychology (Department of Applied Psychology: “Health, Development and Promotion”).
--> Added the number of patients: N = 26

--> Added one limitation regarding standard care: Fourth, compared to University of Salzburg (where the training was developed), the institute for applied psychology lacks a fully equipped routine outpatient clinic. As a consequence, further research in routine care is warranted.

The conclusion that the blended format can improve current group therapy interventions cannot be drawn from this study.
--> Toned down conclusion in discussion + conclusion section: From the viewpoint of addressed patients, modern technology and communication techniques add value to existing group treatments and the blended format should be further tested in routine care. +
The present study indicates that the blended format is feasible for the outpatient group treatment of clinical depression. […] Future research should investigate the blended group treatment in standard care.

considering the study design and sample size the conclusions can be toned down a little.
--> Toned down conclusion in abstract: The application of in- and between-session computer-support for the group treatment of depression seems feasible.

To increase the readability, I think some sentences and words need to be checked by a native speaker (e.g. working factors, battery), checked on correct use of uppercase letters in the text and tables (e.g. E-COMPARED/Minddistrict) and consistent use of words (e.g. seminar, lectures).
->Adapted wording and use of uppercase:
->Working factor -> active treatment factor
->Battery -> questionnaire
->Minddistrict, E-COMPARED
->Seminar -> intervention
->lectures -> sessions
->Modern media -> multimedia
->Computer components -> computer elements
->Effectivity -> Efficacy
…

In the introduction some of the statements need more explanation: page 4 line 7 (because…), page 4 line 27 (needs, such as…), page 5 line 17 (because…).
--> Added two of three suggested explanations:
(page 4,27) At the same time, online interventions do not fit all patients' needs (e.g. need for more personal con-tact or personal preferences).
(page 5) …, because the structure provided by blended therapy might prevent therapists from so called therapist drift (Waller, 2009).

The authors refer to the NICE recommendation for group CBT, however it remains unclear how this relates to the psychoeducation group intervention that is studied. Please specify which elements of the exiting psychoeducational treatment were provided online in the blended format.
--> Added explanation: Thus, the integration of computer-support and face-to-face group therapy might result in an optimized treatment, in which personal contact is preserved to a wide extend.
--> Added specification in section "intervention": Online modules were made accessible via a secure web-based non-profit environment, featuring videos, online work sheets, an unguided group-chat, as well as remote therapist-patient communication.

Page 6 line 43, I think the authors mean "after group session" instead of " after treatment". On which platform was the intervention offered?
--> Added specification: Participants were able to logon the platform after the entire treatment had ended.

--> Added: Online modules were made accessible via a secure web-based non-profit environment, featuring …

page 7 line 9, inclusion and exclusion criteria: specify how severe depression was defined.
--> Corrected error in description of the clinical interview (Mini-DIPS instead of M.I.N.I): Subsequently, the complete Mini-DIPS was applied for study in- or exclusion (Magraf, 2013). The Mini-DIPS is a 30-minute short interview, based on the German DIPS (Diagnostic Interview for Psychological Disorders; Magraf et al., 2013) and features depression criteria for DSM-4 and ICD-10.
-->Specified criteria for severe depression: According to clinical judgement, participants were excluded if they suffered from severe depression (≥ 6 criteria, including main symptoms), …

Page 8 line 53, please explain the use of LOCF instead of a more sophisticated alternative.
--> Applied LMM without further use of LOCF

Page 8 line 25, CSQ: was satisfaction measured on the blended treatment as a whole, or focused on just the online part?
--> Adapted: Participants' overall satisfaction with treatment was measured by the German version …

The conclusion that the blended format can improve current group therapy interventions cannot be drawn from this study.
--> Adapted section: From the viewpoint of addressed patients, modern technology and communication techniques add value to the treatment and the blended group format might have potential to become a future treatment option for mild to moderate depression interventions.

… comorbid anxiety. It remains unclear whether patients with comorbid symptoms were specifically selected and in which way the treatment addresses anxiety. The authors refer to “Treatment of depression” (page 3, line 46), it should be made clear if the focus is on a reduction of depressive symptoms or on remission of major depressive disorder.
-->Added: During the entire procedure, no special recruitment strategy was applied for comorbid anxiety.
-->Added: The intervention focuses on reduction of depressive symptoms and aims at increasing personal resources, but was not tailored to treat comorbid anxiety.

The authors claim that this the first clinical study on blended group therapy for depression, but wasn’t that actually the other paper by Schuster et al (2017)? It is also unclear what the authors mean by the statement of a high fit of the two intervention strategies.

--> Explanation: The very first study (Schuster et al, 2017) lacked any clinical diagnostics and included participants with subclinical levels of self-reported CES-D depressiveness levels. The title acknowledged this “…intervention for adults with depressive symptoms”. Also, the treatment was not tailored specifically to depression. Therefore we regard this study to be the first clinical study for mild to moderate major depression.
--> Adapted: Corresponding results suggest a very high fit or beneficial interplay of the two intervention strategies (psychological group + online support).

Result section, I advise the authors to report the results in a more objective manner, for example ´one participant´ instead of ‘only one participant’ or ´only a small proportion´.
--> Adapted: Wording in results section

Page 10 line 37 and discussion section, please explain why the unguided discussion forum received a lower rating; did the participants wanted it to be guided?

--> Confirmation: … we conclude that online group interaction in blended treatments should either be guided by a therapist (Schulz et al., 2016), or include other incentives to increase usage and perceived relevancy.
--> Added one more explanation: During debriefing, some patients also explained, that their need for group interaction was satisfied by the weekly reunions.

Page 10 line 50, please explain why blended would fit the needs of treating alike group treatments best.
--> Added: Here, the approval rate for classical group psychotherapy was slightly lower (80 % and 72 %).

Discussion, it is unclear what the main message is of the second paragraph and some statements also need references, e.g. more positive appraisals for psychological trainings.
--> Restructured introduction of § 2
--> Inserted references
--> Changed wording: This assumption is also reflected by slightly more positive appraisals for training-alike interventions, compared to classical group psychotherapy (cf. section 3.3.2).

In Table 2 not all the numbers add up to 26. In Table 4 the number of participants is missing.
--> Added missing information and corrected numbers

VERSION 2 – REVIEW

REVIEWER	Philip Batterham The Australian National University, Australia
REVIEW RETURNED	11-Dec-2017

GENERAL COMMENTS	The authors have addressed several of my concerns, particularly the use of linear mixed models instead of LOCF. However, I have a few concerns remaining. 1. “From the viewpoint of the addressed patients, modern technology and communication techniques add value to existing group treatments.” – I’m not sure what this means (who are the “addressed patients”) or what data this conclusion is based on. Perhaps it means the same thing as the next sentence, “Regarding patients’ general appraisals, the evaluation of the blended group treatment was positive” (although this sentence also has redundancy – should be “Participants provided positive evaluations of the blended group treatment.” or something like that). 2. What does “tentatively significant results” mean? I am not familiar with this concept and it appears to defy statistical convention. The authors need to clearly state what their alpha value is, justify it (e.g., whether the work is merely exploratory, whether multiple comparisons were made), and adhere to it. It is incorrect to indicate that 0.103 is equal to or less than 0.1. It is also inaccurate to say that “We do not profit from this result... so the finding itself is more conservative interpretation”, as the authors use this liberal criterion to suggest that the findings suggest “slightly more positive appraisals” were provided. Furthermore, the consistency of the results with other work doesn’t justify ignoring statistical convention. 3. “Gran average” should be “grand average”.

	4. According to BMJ author guidelines, CONSORT statements should only be provided for RCTs. This is not an RCT so should not use CONSORT. Perhaps use STROBE instead as stated previously. 5. There are some word choices which are somewhat unclear, e.g., “omit technology”, “converse experience”, “preconditions”.

REVIEWER	Mayke Mol GGZ inGeest
REVIEW RETURNED	11-Dec-2017

GENERAL COMMENTS	General comments Thank you for revising the previous manuscript. I think the article has improved in terms of readability and that rephrasing of some of the study elements are better fitting now. I have some additional comments regarding the description of the treatment and the diagnosis of the target group. And below you will find some minor comments with the corresponding page numbers. Intervention - The authors refer throughout the paper to a blended group treatment. However, given the size of the intervention group (n=26) and the psycho-educational focus the label treatment does not seem appropriate. I would strongly advise the authors to rephrase, for example: training or intervention instead of treatment. - If I understand correctly, the intervention was provided at the University (no formal infrastructure related to patient care), not in a clinic or at a university clinic. Current description is misleading. - There is a discrepancy between the content description in the introduction and method section, please address this. Major depressive disorder It is commendable that patients were diagnosed at baseline with a structured interview. However, as this was not repeated after the intervention and the primary outcome was CES-D, the intervention appeared to focus on reduction of depressive symptoms, rather than treatment of major depressive disorder. Diagnostic interview: In the previous manuscript the authors mentioned the MINI interview, now it was changed to MINI-dips? It remains unclear why comorbid anxiety is mentioned, the intervention did not specifically focus on anxiety and it was not specifically assessed. My advice would be to focus on the depressive symptoms. - Please check the manuscript for language errors, for example 'expecations' instead of expectations and ‘depressiveness’ in the abstract.

	Abstract (page 2) Please clarify results section in manuscript. Unclear what is meant by 'an active treatment factor'. For whom is treatment intensification an important advantage and why? And what do the authors want to say with ‘almost half of patients would have preferred more time for personal exchange?’. Introduction (page 5) Paragraph on C(B)T group therapy does not seem relevant as the studied intervention is an eclectic psycho-educational self-management intervention. Consider leaving this paragraph out and discussing more relevant information instead, such as the concept of resource-oriented treatment. Methods (page 6) Please specify which platform was used to deliver the online sessions (e.g. Minddistrict, Deprexis, etc.) and provide a reference to the platform. Methods of collecting data on system usage and log data are not mentioned and operationalized, but it is explicitly mentioned in the research question. Discussion (page 12) It is unclear what is meant by 'personal communication' on page 12. Tables (page 21). Table 3 The authors report observed means over n=26. However, based on study adherence this can only be n=23. Perhaps the authors are referring to LS/estimated means instead? It is unclear how descriptive information on usage in Table 4 should be interpreted. How many logins, downloads etc where expected? Table 5 It is unclear how the p-values were calculated, and why only three values are reported.

VERSION 2 – AUTHOR RESPONSE

Suggestions by Mr. Batterham: We have been able to regard all concerns

1. “From the viewpoint of the addressed patients, modern technology and communication techniques add value to existing group treatments.” – I’m not sure what this means (who are the “addressed patients”) or what data this conclusion is based on. Perhaps it means the same thing as the next sentence, “Regarding patients’ general appraisals, the evaluation of the blended group treatment was positive” (although this sentence also has redundancy – should be “Participants provided positive evaluations of the blended group treatment.” or something like that).

Response: Adapted expression: Finally, participants provided positive evaluations of the investigated blended group format.

2. What does “tentatively significant results” mean? I am not familiar with this concept and it appears to defy statistical convention. The authors need to clearly state what their alpha value is, justify it (e.g., whether the work is merely exploratory, whether multiple comparisons were made), and adhere to it. It is incorrect to indicate that 0.103 is equal to or less than 0.1. It is also inaccurate to say that “We do not profit from this result... so the finding itself is more conservative interpretation”, as the authors use this liberal criterion to suggest that the findings suggest “slightly more positive appraisals” were provided. Furthermore, the consistency of the results with other work doesn’t justify ignoring statistical convention.

Response: With “tentatively significant results” we refer to “trend towards significance”. We agree with the reviewer in that the result p=0.103 in table 3 (question 3) lies 3 ‰ above the convention of p = 0.1 and that we definitely should adhere closely to this statistical convention.
Therefore we deleted the corresponding dagger / obelisk from the table and specified the corresponding passage in the discussion: This assumption is reflected by more positive appraisals for the improvement of training-alike interventions, compared to the improvement of classical group therapy (cf. section 3.3.2). However, regarding treatment intensification, participants’ positive appraisals between those forms of delivery did not differ significantly.

3. “Gran average” should be “grand average”.

Response: Corrected wording.

4. According to BMJ author guidelines, CONSORT statements should only be provided for RCTs. This is not an RCT so should not use CONSORT. Perhaps use STROBE instead as stated previously.

Response: We now used STROBE as suggested!

5. There are some word choices which are somewhat unclear, e.g., “omit technology”, “converse experience”, “preconditions”.

Response: Adapted wording: omit technology -> leave out technology; converse experience -> opposite experience; preconditions -> requirements

Suggestions by Ms. Mol: We have been able to regard all concerns except one cluster

If I understand correctly, the intervention was provided at the University (no formal infrastructure related to patient care), not in a clinic or at a university clinic. Current description is misleading.
Adapted: university setting instead of university outpatient setting

Comment:There is a discrepancy between the content description in the introduction and method section, please address this.

Response: Addressed this discrepancy by adapting content

Comment: Major depressive disorder: It is commendable that patients were diagnosed at baseline with a structured interview. However, as this was not repeated after the intervention and the primary outcome was CES-D, the intervention appeared to focus on reduction of depressive symptoms, rather than treatment of major depressive disorder.

Response: Added “improvement in symptoms in” major depression in relevant text passages.

Diagnostic interview: In the previous manuscript the authors mentioned the MINI interview, now it was changed to MINI-dips?

Response: Yes, this is now the correct description. I apologize for this mistake, which happened due to the similarity of both survey names.

Comment: It remains unclear why comorbid anxiety is mentioned, the intervention did not specifically focus on anxiety and it was not specifically assessed. My advice would be to focus on the depressive symptoms.

Response: We deleted all sections about comorbid anxiety. It is now only mentioned as a possible inclusion criterion. However, in the introduction section we would like to mention blended group interventions for anxiety, as these are the only existing blended group interventions/most related to our work.

Comment: Please check the manuscript for language errors, for example 'expecations' instead of expectations and ‘depressiveness’ in the abstract.

Response: We revised the manuscript for more errors.

Abstract (page 2)
Please clarify results section in manuscript. Unclear what is meant by 'an active treatment factor'. For whom is treatment intensification an important advantage and why? And what do the authors want to say with ‘almost half of patients would have preferred more time for personal exchange?’.

Response: Changed wording from “active treatment factor” to “therapeutic factor”
Specified: Patients described intensification as an important advantage of the blended format.
Specified: Half of patients (48%) would …

Methods (page 6)
Please specify which platform was used to deliver the online sessions (e.g. Minddistrict, Deprexis, etc.) and provide a reference to the platform.

Specified the platform: Moodle with SSL-VPN access

Methods of collecting data on system usage and log data are not mentioned and operationalized, but it is explicitly mentioned in the research question.

Added to the description: The platform automatically tracked personal log data for each participant and week.

Discussion (page 12)
It is unclear what is meant by 'personal communication' on page 12.

Response: Added a reference for the personal communication

Tables (page 21).

Table 3
The authors report observed means over n=26. However, based on study adherence this can only be n=23. Perhaps the authors are referring to LS/estimated means instead?

Response: Corrected this error in table 3. It should have meant estimated means! We missed to adapt this after the recalculation with linear mixed models.

Comment: It is unclear how descriptive information on usage in Table 4 should be interpreted. How many logins, downloads etc where expected?

Response: Added an example in discussion (page 12): For example, we provided one work sheet per week for download, but the actual number of downloads was reasonably higher.

Table 5
It is unclear how the p-values were calculated, and why only three values are reported.

Response: We provided a more detailed description in Section 2.5: … and for intervention applicability (Questions 2 and 3 in Table 5) and process aspects (Table 5 in sections 3.3.2 and 3.3.3) paired t-tests were applied. We did not provide statistics for questions 1) 4) and 5) of table as no paired T-Tests can be applied (Questions 1) 4) and 5) only have one level).

Cluster of comments we have not been able to regard to the full extent:

1. Intervention: - The authors refer throughout the paper to a blended group treatment. However, given the size of the intervention group (n=26) and the psycho-educational focus the label treatment does not seem appropriate. I would strongly advise the authors to rephrase, for example: training or intervention instead of treatment.

2. Introduction (page 5): Paragraph on C(B)T group therapy does not seem relevant as the studied intervention is an eclectic psycho-educational self-management intervention. Consider leaving this paragraph out and discussing more relevant information instead, such as the concept of resource-oriented treatment.

Response:
Ms. Mol argues, that sample size and the psychoeducational focus disqualifies our intervention from the label treatment or CBT. Consequently corresponding paragraphs of the introduction should be deleted or strongly reorganised.

In this context, it is not clear to us, how sample size possibly can determine the classification of a given treatment. Regarding the content of our intervention we carefully regarded best practice guidelines for empirically supported CBT group treatments of depression (e.g. psychoeducation, behavioural goal setting, cognitive restructuring, relapse prevention, and double trainer setting) during the design of the intervention – cf.: DeLucia-Waack, J. L., Kalodner, C. R., & Riva, M. (2014). Handbook of group counselling and psychotherapy. Thousand Oaks, CA: Sage. (Table 29.1, p. 372: “Best practice for CBT group treatments of depression”). Except the duration of the intervention (7 instead of 14 sessions), we closely followed all listed principles. Therefore it is difficult for us to fully comprehend, how the applied content can disqualify the intervention from being CBT. According to the Sage Handbook of Group Counselling & Psychotherapy (p. 327, table 29.1) applied strategies clearly can be identified as CBT group therapy and our intervention also included further CBT elements. Additionally, self-management therapy has a long tradition in the treatment of depression, and derived techniques such as behavioural goal setting or activity monitoring are frequently applied techniques in blended therapy. Thus, our rationale only deviates from guidelines in terms of short treatment duration. However, this is a commonly applied core principle of blended interventions and we do critically reflect on possible negative consequences of this practice in our manuscript.

Given this situation, we would kindly suggests the following adoptions:

1. We already avoided expressions such as psychotherapy, group psychotherapy or group counselling. We now also avoided the label “blended therapy” and replaced it with blended intervention or blended format in corresponding passages. Still, we regard the investigated intervention as a “psychological treatment”.

2. As we think the label “treatment” is no specific (or protected) expression such as psychotherapy, group psychotherapy, group counselling or so on (which we already avoided), but (according to the oxford dictionary) instead is used in many contexts as a basic expression for health-oriented manipulation (medical treatment, physiotherapeutic treatment, internet-based treatment and so on…). Therefore. we would keep this expression in the manuscript. To avoid any confounding with above mentioned expressions, we contextualised the language convention of the expression “treatment” in many cases, such as treatment strategy, treatment duration, treatment process, or treatment intensification and fully avoided the sole use of the expression “treatment”.

3. We added a paragraph on the role of psychoeducation, resource-oriented psychotherapy principles and self-management for the CBT-based depression intervention. We think, this was an important missing link and we thank Ms.Mol for indicating this!

We hope we have been able to capture the essence of Ms. Mol´s concerns! We would kindly invite Ms.Mol to revise refinements of the applied nomenclature and hope, our suggestions are a practicable compromise!