PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Comparing responses to differently framed and formatted persuasive messages to encourage help-seeking for depression in Japanese adults: a cross-sectional study with 2-month follow-up
AUTHORS	Suka, Machi; Yamauchi, Takashi; Yanagisawa, Hiroyuki

VERSION 1 – REVIEW

REVIEWER	Amy Morgan University of Melbourne, Australia
REVIEW RETURNED	24-Jan-2018

GENERAL COMMENTS	This paper is a randomised trial of different help-seeking messages for depression. It tests the effectiveness of different health communication principles (framing and formatting) in an online study in Japan. I appreciated that the authors tested specific principles to include in help-seeking messages, rather than just the effectiveness of one message. Primary outcomes are intentions to seek help but actual help-seeking behaviour is also measured at 2-month follow-up, which is also a strength. Overall there is a lack of clarity around some of the methodology and I believe the interpretation of results are overstated. The paper would also be improved with attention to the following points: Introduction 1. I would have liked to see more context in the introduction to enhance the justification for the research. More specific information about the public health programs (p4 line 16) and how they relate to persuasive help-seeking messages would be useful. Furthermore, the authors should expand upon the second paragraph to explain which contexts loss-framed messages are more effective and whether these may apply to help-seeking for depression or mental health care. 2. It would be useful to include brief information about how the CDC criteria were developed. For instance were these based on expert consensus, or do they have supporting experimental evidence? Methods 3. Could the authors clarify the relationship between this study and that cited in reference 20. It's not clear whether the data are from the same study and this is a different analysis, or whether this is a new sample which builds on the prior study. It should be clear how the results in this paper are different. 4. Could the authors explain what the quotas were for K6 score in section 2.2. It appears that they have chosen a non-standard cut-off to indicate mental disorder (5 rather than 13). This should be justified, as this could have significant effects on

outcomes if this was chosen arbitrarily. For example, rates of help-seeking were quite low at follow-up in the 'depressed' sample, and this could be due to the use of a low cutoff for depressed status.

5.   More information about some of the measures is required. What validity data supports the perceived effectiveness scales? I couldn't understand how participants could rate 'keep for future reference' when the material was presented as part of an online survey. It's not clear what 'how best to do so' means (section 2.3.2). Further information about what is meant by tone and spacing would be helpful (section 2.3.4). The HLS-14 is mentioned on p11 but there is no information about this measure.

6.   The statistical methods describe the use of logistic regression with adjustment for confounders, but I could not find these results in the manuscript. These may be more useful for readers than the unadjusted analyses.

Results

7.   As this is a randomised trial with follow-up, I would like to see a flow chart of numbers in each group, and how many in each group were assessed at follow-up, to see if there was differential attrition.

8.   There is an error in Table 1 – the numbers for gender total 1967 but there are only 1957 participants.

9.   For ease of interpretation of results, I would have preferred different labels to indicate the different messages. Rather than 1N/1P etc, they could be labelled as neutral-visual and neutral-unformatted, for example.

10.   If there is scope to do so, I would like to see the interactions plotted, for a visual representation of these results, otherwise interactions can be difficult to interpret.

11.   Please provide the proportions for each message type on p12 line 33-37, rather than just the statement that they were not significantly different. This aids future research as the effect may have been too small to detect in this study.

Discussion

12.   My main concern with the discussion is that the authors have overstated the findings of their research. I don't believe these results "confirm the effectiveness of depression help-seeking messages in middle-aged Japanese people". Whilst self-reported help-seeking intentions may have increased in the short-term, this effect disappeared for half of the sample at follow-up, and there is little data to support effects on actual help-seeking behaviour. A control group who did not receive any depression-related message would also provide stronger evidence that the messages change help-seeking intentions.

13.   I also think there are limitations that the author has not addressed. These include the large number of analyses conducted and the potential for type I errors. This was not acknowledged or adjusted for in analyses. Another limitation is the low external validity of the study (the unnatural exposure to the messages in a study, rather than amongst a background of other activity during normal life).

Minor comments

14.   There are some minor typos and grammatical errors throughout the manuscript. Eg. 'any messages' in the abstract (should be 'all messages'), 'massage' on p5 line 11 (should be 'messages'), 'casually' on p5 line 39 (should be 'causally'), 'help-seeking intension' on p11 line 50.

15.   Some sentences should be re-written to enhance clarity, eg p5 line 13 "Applying the index items…"

REVIEWER	Hisateru Tachimori National center of neurology and psychiatry
REVIEW RETURNED	08-Feb-2018

GENERAL COMMENTS	This study provides important information on health communication of treatment for depression. Though this study has limitation in external validity, the study was executed well. I think this manuscript is suitable for publication.

REVIEWER	Juul Houwen Department of Primary and Community Care, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands
REVIEW RETURNED	19-Apr-2018

GENERAL COMMENTS	Dear editor and authors, Thank you for giving me the opportunity to review the manuscript about an interesting topic (A comparative study on persuasive health message design: effects of message framing and formatting on comprehensibility, persuasiveness, emotion, intention, and action). This is a cross sectional study which examines the effects of message framing and formatting on persuasive message effectiveness in the context of developing depression help-seeking messages. I carefully reviewed your manuscript and I have the following remarks. Major concerns Overall this is a well written paper which seems relevant to health communication research. However there are some issues regarding the key message and the structure of the manuscript. I would like to point out the following issues: Firstly, the authors are kindly invited to provide some explicit information regarding the key message of the manuscript. The authors describe there are no differences between loss-framed and gain-fraimed message in help seeking intention. On the basis of existing literature, we already know the influence of message framing and formatting on persuasive message effectiveness. I would encourage the authors to deepen their new findings and to describe these key messages more in detail. Table 3 shows that 2P and 3N change the help-seeking intention for both non-depressed and depressed groups. However, The biggest gap in changing help-seeking intention is found for 1 P (26.6 – 40.6%). What is the main conclusion of this manuscript? They should discuss these new key findings to 'the implications of further research' as this section is hardly described in the discussion. Secondly, the authors should improve the structure of the manuscript. For an example, some parts of the section 'methods' (number of patients who were included and excluded) should be written in the section 'results'. Further, the authors repeat in the section 'discussion' some findings which have been described already in the section 'results'. They also should discuss the 6 different framed and formatted message at the beginning of the section 'methods'. Minor concerns Introduction This paper aims to examine the effects of message framing and formatting on persuasive message effectiveness in the context of

developing depression help-seeking messages. However, they do not describe why they focused especially on depression. I advise the authors to extend and deepen the background and ratio concerning health communication and depression.

Methods
Page 7, line 14. The authors describe the third message (gain framed) and describe ''each message consisted of three part''. I think it should be parts instead of part.
The authors describe the difference between formatted and unformatted messages at the end of the subheading 2.1 ''messages''. I would encourage the authors to describe this difference between formatted and unformatted messages at the beginning of the section 'methods', as this would give a more clear overview of the methods (framing and formatting) and the 6 different messages.
I think the manuscript will improve when the authors describe some more information about the inclusion and exclusion criteria. Why did the authors choose to include just adults aged 35-45 years? And can they please give some information about the K6 score? According to me, the K6 score intended to yield a global measure of distress instead of a depression.

It seems to be a little confusing how many patients finally were included and excluded in the follow up survey. Maybe the authors can provide a clear overview of number of patients who were included and excluded and I will recommend the authors to describe the section 'number of patients' in the section 'results' and not in the section 'methods'.
Help seeking intention. The authors describe that the participants who gave affirmative answers (certainly yes and probably yes) were counted as having positive help seeking intention. Did the authors measure this on a likert-scale?

Results
The authors describe the characteristics of the study participants. They discuss in the first paragraph the age with university degrees. I should recommend the authors to mention this section in the 'discussion' and not in the section 'results'.
Further, the authors describe that no significant differences were found between the message groups in sociodemographic characteristics. Maybe they can add the p-values in table 1.
The authors describe that 143 people (57,4%) reported a positive help-seeking intention for depression again at follow up. What did the other people report?
The authors mention the term 'stable help-seeking intention'', can they please describe in more detail the meaning of this term?

Discussion
I think the authors can improve the structure of the discussion section. The authors should make four different paragraphs: summary of main findings, comparison with literature, strengths and limitations and finally implications for further research.
The authors mention that the results of this study successfully confirmed the effectiveness of depression help seeking messages in the middle aged Japanese people. What do they mean with 'successfully'. Can the authors give more information about this term? I think the authors may describe the main results more clear at the beginning of the discussion.

	The authors stated that no marked difference was found between the loss-framed and gain-framed message in help seeking intention. The authors are encourage to discuss these results. Can they explain why there is no significant difference between the loss-framed and gain-framed? Further, 2P and 3N seems to enhance help seeking intention in depressed and non depressed groups. Should the authors recommend these frames and formats in relation to help seeking intention? And what about 1 NP, it seems that this message would have the biggest change in help-seeking intervention.

VERSION 1 – AUTHOR RESPONSE

Dear Reviewer 1 - Dr. Morgan,

We greatly appreciate your favorable comments and helpful suggestions on our manuscript. We have incorporated your suggestions in the revised paper. Below we would like to offer our responses to your comments. We hope that our revision will meet with your approval. Please take note that those written in black are your comments while those in blue are our responses.

1.   Introduction
I would have liked to see more context in the introduction to enhance the justification for the research. More specific information about the public health programs (p4 line 16) and how they relate to persuasive help-seeking messages would be useful. Furthermore, the authors should expand upon the second paragraph to explain which contexts loss-framed messages are more effective and whether these may apply to help-seeking for depression or mental health care.

The background of this study has been explained in further detail. [p 5, 6]

2.   Introduction
It would be useful to include brief information about how the CDC criteria were developed. For instance were these based on expert consensus, or do they have supporting experimental evidence?

The Clear Communication Index is introduced as an evidence-based tool to plan and assess public communication materials. Previous studies have demonstrated that the materials revised using the Clear Communication Index are rated more favorably than the originals by possible audience members. The application of the Clear Communication Index makes it more likely that audience can correctly identify the intended main message and understand the words in the materials. [p 6]

3.   Methods
Could the authors clarify the relationship between this study and that cited in reference 20. It's not clear whether the data are from the same study and this is a different analysis, or whether this is a new sample which builds on the prior study. It should be clear how the results in this paper are different.

Our research project on depression help-seeking messages was launched to answer two research questions. The objective of this study was to examine the effects of message framing and formatting on the effectiveness of depression help-seeking messages. The objective of the study cited in reference 25 was to determine whether the effects of depression help-seeking messages are influenced by audience's depressive status. This study and the study cited in reference 25 used data collected from the same sample but analyzed them with respective objectives. The description of our research project has been modified to make it clearly understandable by the readers. [p 7]

4.   Methods
Could the authors explain what the quotas were for K6 score in section 2.2. It appears that they have chosen a non-standard cut-off to indicate mental disorder (5 rather than 13). This should be justified, as this could have significant effects on outcomes if this was chosen arbitrarily. For example, rates of help-seeking were quite low at follow-up in the 'depressed' sample, and this could be due to the use of a low cutoff for depressed status.

The Japanese version of the 6-item Kessler Psychological Distress Scale (K6) has been established as a screener for depression in Japan. A validation study revealed that a K6 score ≥5 is a reasonable cutoff to distinguish between depressed and non-depressed people. This definition has been commonly used in epidemiological studies in Japan. [p 9]

5.   Methods
More information about some of the measures is required. What validity data supports the perceived effectiveness scales? I couldn't understand how participants could rate 'keep for future reference' when the material was presented as part of an online survey. It's not clear what 'how best to do so' means (section 2.3.2). Further information about what is meant by tone and spacing would be helpful (section 2.3.4). The HLS-14 is mentioned on p11 but there is no information about this measure.

The comprehensibility scale (section 2.3.1) and the persuasiveness scale (section 2.3.2) were developed and validated to measure audience's perceptions of effectiveness of health messages in Japanese people (See reference 23). The six items for design quality (section 2.3.4) and the three items for intended future use (section 2.3.5) were derived from the Consumer Information Rating Form developed by Krass and colleagues (See reference 31). The use of these subjective measures is preferable when the survey aims to understand audience's perspective on health messages. The measures used in the survey have been explained in further detail. The sentence containing the HLS-14 has been deleted in the revised manuscript. [p 11]

6.   Methods
The statistical methods describe the use of logistic regression with adjustment for confounders, but I could not find these results in the manuscript. These may be more useful for readers than the unadjusted analyses.

This is our mistake. Multiple logistic regression analysis was conducted to compare the effects of 6 differently framed and formatted messages on help-seeking intention for depression with adjustment for potential confounders. Compared with the neutral-plain (1N) message as a reference group, the loss-visual (2P) message had a significantly greater effect, but the others did not: the adjusted odds ratios (95% confidence intervals) of the neutral-visual (1N), loss-plain (2N), loss-visual (2P), gain-plain (3N), and gain-visual (3P) messages were 1.31 (0.89-1.92), 1.29 (0.88-1.89), 1.57 (1.07-2.29), 1.39 (0.95-2.04) and 1.41 (0.97-2.06), respectively. This result indicates that the loss-visual (2P) message worked better than the other messages. [p 14]

7.   Results
As this is a randomised trial with follow-up, I would like to see a flow chart of numbers in each group, and how many in each group were assessed at follow-up, to see if there was differential attrition.

The flow of participants through the study has been described with a flow chart shown in Figure 1. [p 13]

8.   Results
There is an error in Table 1 ? the numbers for gender total 1967 but there are only 1957 participants.

The number of participants was 1,957 (980 men and 977 women). Table 1 has been corrected.

9.   Results

For ease of interpretation of results, I would have preferred different labels to indicate the different messages. Rather than 1N/1P etc, they could be labelled as neutral-visual and neutral-unformatted, for example.

As you suggested, the message labels have been changed from 1N, 1P, 2N, 2P, 3N, and 3P to neutral-plain, neutral-visual, loss-plain, loss-visual, gain-plain, and gain-visual, respectively.

10. Results
If there is scope to do so, I would like to see the interactions plotted, for a visual representation of these results, otherwise interactions can be difficult to interpret.

As shown in Table 2, significant frame×format interaction effects were found only on 'happiness' and 'anxiety'. This result was just as we had expected. We have no intension to bring focus to the frame×format interaction effects, so we would like to simply report the result in the text.

11. Results
Please provide the proportions for each message type on p12 line 33-37, rather than just the statement that they were not significantly different. This aids future research as the effect may have been too small to detect in this study.

The proportions of participants with help-seeking action by message group have been shown in Table 4. [p 14]

12. Discussion
My main concern with the discussion is that the authors have overstated the findings of their research. I don't believe these results "confirm the effectiveness of depression help-seeking messages in middle-aged Japanese people". Whilst self-reported help-seeking intentions may have increased in the short-term, this effect disappeared for half of the sample at follow-up, and there is little data to support effects on actual help-seeking behaviour. A control group who did not receive any depression-related message would also provide stronger evidence that the messages change help-seeking intentions.

As described in the methods (section 2.1), the aim of messaging was to increase people's help-seeking intentions for depression. All messages except the neutral-plain (1N) message produced significant increase in help-seeking intention after exposure to the messages. This result supports the effectiveness of communicating persuasive messages for increasing people's help-seeking intentions for depression. We think it the first step toward success in developing depression help-seeking messages. Further studies are needed to find a way to sustain the effect of messaging for a long time. As you pointed out, it is hard to say from these results that the depression help-seeking messages were effective. The conclusive discussion has been modified not to overstate the findings of this study. [p 15]

13. Discussion
I also think there are limitations that the author has not addressed. These include the large number of analyses conducted and the potential for type I errors. This was not acknowledged or adjusted for in analyses. Another limitation is the low external validity of the study (the unnatural exposure to the messages in a study, rather than amongst a background of other activity during normal life).

In order to reduce type I errors, we tried to keep the number of analyses to a minimum and to put a conservative interpretation on the results. The concern about generalizability of findings have been mentioned as limitations in the discussion. [p 17]

14. Discussion

There are some minor typos and grammatical errors throughout the manuscript. Eg. 'any messages' in the abstract (should be 'all messages'), 'massage' on p5 line 11 (should be 'messages'), 'casually' on p5 line 39 (should be 'causally'), 'help-seeking intension' on p11 line 50.

We have carefully checked the English throughout the paper and corrected the spelling and grammar mistakes.

15. Discussion
Some sentences should be re-written to enhance clarity, eg p5 line 13 "Applying the index items…"

The sentences have been modified to make them clearly understandable by the readers.

Dear Reviewer 2 - Dr. Tachimori,

We greatly appreciate your favorable comments on our manuscript. We have incorporated other reviewers' suggestions in the revised paper. We hope that our revision will meet with your approval.

Dear Reviewer 3 - Dr. Houwen,

We greatly appreciate the detailed thoughtful comments on our manuscript. We have incorporated your suggestions in the revised paper. Below we would like to offer our responses to your comments. We hope that our revision will meet with your approval. Please take note that those written in black are your comments while those in blue are our responses.

1. Major concerns
Firstly, the authors are kindly invited to provide some explicit information regarding the key message of the manuscript. The authors describe there are no differences between loss-framed and gain-framed message in help seeking intention. On the basis of existing literature, we already know the influence of message framing and formatting on persuasive message effectiveness. I would encourage the authors to deepen their new findings and to describe these key messages more in detail. Table 3 shows that 2P and 3N change the help-seeking intention for both non-depressed and depressed groups. However, the biggest gap in changing help-seeking intention is found for 1P (26.6-40.6%). What is the main conclusion of this manuscript? They should discuss these new key findings to 'the implications of further research' as this section is hardly described in the discussion.

First of all, for your information, according to the advice from Reviewer 1, the message labels have been changed from 1N, 1P, 2N, 2P, 3N, and 3P to neutral-plain, neutral-visual, loss-plain, loss-visual, gain-plain, and gain-visual, respectively in the revised manuscript.

As described in the methods (section 2.1), the aim of messaging was to increase people's help-seeking intentions for depression. All messages except the neutral-plain (1N) message produced significant increase in help-seeking intention after exposure to the messages. This result supports the effectiveness of communicating persuasive messages for increasing people's help-seeking intentions for depression. We think it the first step toward success in developing depression help-seeking messages. [p 15]

As for the percentage changes in help-seeking intention for depression by message group shown in Table 3, we think it difficult to compare the magnitude of the numbers. The loss-plain (2N) message

showed a 15.9% increase in the proportion of people who reported a positive help-seeking intention for depression after exposure to the message, and the gain-plain (3N) message showed a greater percentage increase (23.1%); however, the proportions of participants who reported a positive help-seeking intension after exposure these messages were equivalent (44.9% vs. 44.6%). Meanwhile, the loss-visual (2P) message showed a 29.1% increase in the proportion of people who reported a positive help-seeking intention for depression after exposure to the message, and the gain-visual (3P) message showed a smaller percentage increase (17.0%); however, the proportions of participants who reported a positive help-seeking intension after exposure these messages were equivalent (46.6% vs. 47.4%). From these results, we cannot say that the loss-visual (2P) message and the gain-plain (3N) message had greater effects than the other messages. Also, it is hard to say that the loss-framed messages were more effective than the gain-framed messages or vice versa. [p 16]

In order to compare the effects of 6 differently framed and formatted messages on help-seeking intention for depression, multiple logistic regression analysis was further conducted with adjustment for potential confounders. Compared with the neutral-plain (1N) message as a reference group, the loss-visual (2P) message had a significantly greater effect, but the others did not: the adjusted odds ratios (95% confidence intervals) of the neutral-visual (1N), loss-plain (2N), loss-visual (2P), gain-plain (3N), and gain-visual (3P) messages were 1.31 (0.89-1.92), 1.29 (0.88-1.89), 1.57 (1.07-2.29), 1.39 (0.95-2.04) and 1.41 (0.97-2.06), respectively. This result indicates that the loss-visual (2P) message worked better than the other messages. Loss-framing and formatting seemed to act synergistically to increase help-seeking intention for depression. [p 14, 16]

Previous studies have not provided a conclusive answer as to what kind of message will more satisfactorily motivate people to seek mental health care, but it is certain that message framing and formatting influence persuasive message effectiveness. Despite the potential limitations of this study, it would be recommendable to apply loss-framing and formatting to depression help-seeking messages. [p 15, 16]

2.   Major concerns
Secondly, the authors should improve the structure of the manuscript. For an example, some parts of the section 'methods' (number of patients who were included and excluded) should be written in the section 'results'. Further, the authors repeat in the section 'discussion' some findings which have been described already in the section 'results'. They also should discuss the 6 different framed and formatted messages at the beginning of the section 'methods'.

As you suggested, the structure of the manuscript has been revised. We believe that the revised manuscript has become better organized.

3.   Introduction
This paper aims to examine the effects of message framing and formatting on persuasive message effectiveness in the context of developing depression help-seeking messages. However, they do not describe why they focused especially on depression. I advise the authors to extend and deepen the background and ratio concerning health communication and depression.

The background of this study has been explained in further detail. [p 5, 6]

4.   Methods
Page 7, line 14. The authors describe the third message (gain framed) and describe ''each message consisted of three part''. I think it should be parts instead of part.

We have carefully checked the English throughout the paper and corrected the spelling and grammar mistakes.

The authors describe the difference between formatted and unformatted messages at the end of the subheading 2.1 ''messages''. I would encourage the authors to describe this difference between

formatted and unformatted messages at the beginning of the section 'methods', as this would give a more clear overview of the methods (framing and formatting) and the 6 different messages.

The description of message framing and formatting has been modified to make it clearly understandable by the readers. [p 8]

I think the manuscript will improve when the authors describe some more information about the inclusion and exclusion criteria. Why did the authors choose to include just adults aged 35-45 years? And can they please give some information about the K6 score? According to me, the K6 score intended to yield a global measure of distress instead of a depression.

The Comprehensive Survey of Living Conditions revealed that people who were feeling stressed or distressed were most frequently observed in the 40-49 age group (58.7% in men and 48.6% in women). In addition, the World Mental Health Japan Survey revealed that the 12-month prevalence of mental disorders was significantly higher in the younger age groups. Therefore, people aged 35-45 years seems to be a suitable target for persuasive messages encouraging help-seeking for depression. [p 9]

The Japanese version of the 6-item Kessler Psychological Distress Scale (K6) has been established as a screener for depression in Japan. A validation study revealed that a K6 score ≥5 is a reasonable cutoff to distinguish between depressed and non-depressed people. This definition has been commonly used in epidemiological studies in Japan. [p 9]

It seems to be a little confusing how many patients finally were included and excluded in the follow up survey. Maybe the authors can provide a clear overview of number of patients who were included and excluded and I will recommend the authors to describe the section 'number of patients' in the section 'results' and not in the section 'methods'.

The flow of participants through the study has been described with a flow chart shown in Figure 1. [p 13]

Help seeking intention. The authors describe that the participants who gave affirmative answers (certainly yes and probably yes) were counted as having positive help seeking intention. Did the authors measure this on a likert-scale?

The method of measuring help-seeking intention has been described in our previous papers (references 23-25), and so a brief outline has been given in this paper. Participants answered the question on a four-point scale (certainly yes/probably yes/probably not/certainly not). Those who gave affirmative answers (certainly yes and probably yes) were counted as having a positive help-seeking intention. [p 12]

5.   Results
The authors describe the characteristics of the study participants. They discuss in the first paragraph the age with university degrees. I should recommend the authors to mention this section in the 'discussion' and not in the section 'results'.

The possible selection bias has been mentioned as a limitation in the discussion. [p 17]

Further, the authors describe that no significant differences were found between the message groups in sociodemographic characteristics. Maybe they can add the p-values in table 1.

We tried showing the comparison of sociodemographic characteristics between message groups in Table 1 and found that the table contained too much information to read and grasp the point. We therefore would like to simply report the result in the text.

The authors describe that 143 people (57,4%) reported a positive help-seeking intention for depression again at follow up. What did the other people report? The authors mention the term 'stable help-seeking intention'', can they please describe in more detail the meaning of this term?

Participants in the follow-up survey were asked about help-seeking intention for depression using the same method as in the initial survey, as well as help-seeking action for their own mental health during the follow-up period. As described in the methods (section 2.3.6), those who chose 'certainly yes' and 'probably yes' for the question about help-seeking intention were counted as having a positive help-seeking intention. The other people chose 'certainly no' or 'probably no', meaning that they did not have a positive help-seeking intention. The term "stable help-seeking intention" meant that they had a positive help-seeking intention for depression in the initial survey and again in the follow-up survey. To be more understandable for the readers, the word "report (a positive help-seeking intention)" has been changed to "have (a positive help-seeking intention)", and the term "stable (help-seeking intention)" has been changed to "maintaining (help-seeking intention)". [p 12, 14]

6.   Discussion
I think the authors can improve the structure of the discussion section. The authors should make four different paragraphs: summary of main findings, comparison with literature, strengths and limitations and finally implications for further research.

As you suggested, the structure of the discussion has been revised. We believe that the revised manuscript has become better organized.

The authors mention that the results of this study successfully confirmed the effectiveness of depression help seeking messages in the middle aged Japanese people. What do they mean with 'successfully'. Can the authors give more information about this term? I think the authors may describe the main results more clear at the beginning of the discussion.

As described in the methods (section 2.1), the aim of messaging was to increase people's help-seeking intentions for depression. All messages except the neutral-plain (1N) message produced significant increase in help-seeking intention after exposure to the messages. This result supports the effectiveness of communicating persuasive messages for increasing people's help-seeking intentions for depression. We think it the first step toward success in developing depression help-seeking messages. [p 15]

The authors stated that no marked difference was found between the loss-framed and gain-framed message in help seeking intention. The authors are encouraged to discuss these results. Can they explain why there is no significant difference between the loss-framed and gain-framed? Further, 2P and 3N seems to enhance help seeking intention in depressed and non-depressed groups. Should the authors recommend these frames and formats in relation to help seeking intention? And what about 1 NP, it seems that this message would have the biggest change in help-seeking intervention.

As for the percentage changes in help-seeking intention for depression by message group shown in Table 3, we think it difficult to compare the magnitude of the numbers. The loss-plain (2N) message showed a 15.9% increase in the proportion of people who reported a positive help-seeking intention for depression after exposure to the message, and the gain-plain (3N) message showed a greater percentage increase (23.1%); however, the proportions of participants who reported a positive help-seeking intension after exposure these messages were equivalent (44.9% vs. 44.6%). Meanwhile, the loss-visual (2P) message showed a 29.1% increase in the proportion of people who reported a positive help-seeking intention for depression after exposure to the message, and the gain-visual (3P) message showed a smaller percentage increase (17.0%); however, the proportions of participants who reported a positive help-seeking intension after exposure these messages were equivalent (46.6% vs. 47.4%). From these results, we cannot say that the loss-visual (2P) message and the

gain-plain (3N) message had greater effects than the other messages. Also, it is hard to say that the loss-framed messages were more effective than the gain-framed messages or vice versa. [p 16]

In order to compare the effects of 6 differently framed and formatted messages on help-seeking intention for depression, multiple logistic regression analysis was further conducted with adjustment for potential confounders. Compared with the neutral-plain (1N) message as a reference group, the loss-visual (2P) message had a significantly greater effect, but the others did not: the adjusted odds ratios (95% confidence intervals) of the neutral-visual (1N), loss-plain (2N), loss-visual (2P), gain-plain (3N), and gain-visual (3P) messages were 1.31 (0.89-1.92), 1.29 (0.88-1.89), 1.57 (1.07-2.29), 1.39 (0.95-2.04) and 1.41 (0.97-2.06), respectively. This result indicates that the loss-visual (2P) message worked better than the other messages. Loss-framing and formatting seemed to act synergistically to increase help-seeking intention for depression. [p 14, 16]

VERSION 2 – REVIEW

REVIEWER	Amy Morgan University of Melbourne, Australia
REVIEW RETURNED	06-Jun-2018

GENERAL COMMENTS	The revised manuscript is much improved and you have addressed my concerns well. There are 2 minor issues that need correcting, (1) What does "the behaviour itself" refer to in the introduction? p5, line 18, (2) A few examples of misspelling 'intension' remain.

REVIEWER	Juul Houwen Primary and Community care, Nijmegen, Radboud university medical center, the Netherlands
REVIEW RETURNED	30-May-2018

GENERAL COMMENTS	Dear authors, The manuscript has significantly been improved. I do no have further suggestions.

VERSION 2 – AUTHOR RESPONSE

Dear Reviewer 1 - Dr. Morgan,

We greatly appreciate your favorable comments and helpful suggestions on our manuscript. We have incorporated your suggestions in the revised paper. Below we would like to offer our responses to your comments. We hope that our revision will meet with your approval. Please take note that those written in black are your comments while those in blue are our responses.

16. What does "the behavior itself" refer to in the introduction? p5, line 18
The phrase 'the behavior itself' has been specified as 'people's behaviors towards mental illness itself". [p 5]

17. A few examples of misspelling 'intension' remain
We have corrected the spelling throughout the paper.

Dear Reviewer 3 - Dr. Houwen,

We greatly appreciate your favorable comments on our manuscript. We have incorporated other reviewer's suggestions in the revised paper. We hope that our revision will meet with your approval.