PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Social deprivation, gender and obesity: multiple stigma? Results of a population survey from Germany
AUTHORS	Makowski, Anna; Kim, Tae; Luck-Sikorski, Claudia; von dem Knesebeck, Olaf

VERSION 1 - REVIEW

REVIEWER	Sherry Kit Wa Chan Department of Psychiatry The University of Hong Kong
REVIEW RETURNED	12-May-2018

GENERAL COMMENTS	Social deprivation, gender, and obesity multiple stigma? Results of a population survey from Germany The research theme and questions of this article are interesting and important. Below are some suggestions for further improvement of the manuscript: 1. It would be good to state clearly the rationale for selecting a share of 30% mobile numbers given the use of mobile-only in the population is 13% 2. Please explain in detail what Kish-Selection-Grid method of selecting the person from this household. 3. It was not clear the methods that the study used to approach the subjects. Has a systematic way been adopted to reach the identified subjects given the low response rate? 4. It was unclear the reasons that only 49.4% of the total interviewed subjects were included in the final analysis. This may lead to bias. 5. No comparison of the sociodemographic information of the study sample and the population. Given the low response rate, it was unsure the representativeness of the study sample. Therefore one cannot claim that this is a representative cross-sectional survey study. 6. Since comparisons were conducted for individual items, correction for multiple comparisons should be conducted. 7. What was the sequence of the vignette presentation? Any randomization of the sequence? As only one vignette was presented to each respondent, the proportion of subjects answering each vignette should be reported. 8. Since no neutral vignette was used, that is all vignette were an obese person, therefore it is difficult to conclude on the addictive effect of the stigma of gender or SES on obesity. The results and discussion may adjust based this limitation.

REVIEWER	Lesley Gray University of Otago, Wellington, New Zealand
REVIEW RETURNED	25-Jun-2018

GENERAL COMMENTS	Overall: The authors are to be commended for their approach to this question and a well presented manuscript. The results will be of interest to readers. There are however a number of aspects of the manuscript requiring attention. I note my comments below for the authors' attention. Throughout the document: The authors refer to “public obesity stigma” this seems unnecessary, “obesity stigma” would suffice, although my preference would be “weight stigma” throughout the document. There are inconsistencies in use of the term “obesity stigma” and “weight stigma”, please address. Many sentences commence “Moreover” suggest adjust to reduce occurrences. When referring to obesity please refer to “people with obesity” or “female/women with obesity” or “male/men with obesity” (e.g. page 14 line 9 – 11, male obese, female obese). Line specific: Page 1 line 3 – delete comma after gender in title? Page 2 line 38-49. Please revise for clarity and can you remove the ‘assume’ reference? Page 3 – consider referring to “pre-recorded audio vignette” to be clear. Page 4 line 4 – delete ‘share’ prefer proportion or percentage or prevalence. Page 5 line 9 – Reword first sentence for clarity. Page 5 line 21-22 – Reword last sentence for clarity. Page 7 line 12 – suggest add ‘audio’ to all audio vignettes Page 7 line 18 – the term used in the vignette is “severely overweight” please correct comment Page 7 line 25 – add in ‘self-reported’ weight and height Results section – I have not conducted statistical tests, however the results as presented appear acceptable. Page 14 line 52 – suggest re word possibly “However, these results can shed light on…” Page 16 line 34 – the sentence “If affected by obesity…” needs to be clearer. References Generally very good, although I was somewhat surprised the authors did not include the following: World Health Organization. (2017). Weight bias and obesity stigma: considerations for the WHO European Region. Available at http://www.euro.who.int/en/health-topics/noncommunicable-diseases/obesity/publications/2017/weight-bias-and-obesity-stigma-considerations-for-the-who-european-region-2017 The authors may also be interested to read the following, although not directly related to the topic of this manuscript, the article makes linkages between SES, gender, severe and morbid obesity in relation to a different issue: Gray L. Social Determinants of Health, Disaster Vulnerability, Severe and Morbid Obesity in adults: Triple Jeopardy? International Journal of Environmental Research and Public Health, 2017. 14, 1452.

REVIEWER	Aapo Hiilamo Finnish Institute of Occupational Health, Helsinki, Finland
REVIEW RETURNED	18-Jul-2018

GENERAL COMMENTS	Thank you for the opportunity to review this manuscript. The paper examines the association between various stigma indicators and the obesity vignettes. Overall, the topic and the research questions are certainly very important ones, and the data is valuable. However, I have some concerns regarding the method description, results and discussion section. Methods: - Several terms are used inconsistently making the manuscript confusing to follow. Tthere is no name for the emotional reaction variables in the method section but in the results, these are described either as “negative/positive emotional reactions” (table3) or “scale prosocial / scale anger/disgust” (table 5). This is also the case with low occupational position and low SES (these are not synonyms). The manuscript would be much easier to follow if the authors would use terms consistently thorough the entire ms. - Description of statistical analyses is inadequate: it would be useful to describe why given method is used. For example, the statement that “To test significant mean differences between… mann-whitney-u tests were applied” does not describe this method correctly. Why was mann-whitney-u chosen? How wereas the indicators distributed? - What was the analytical sample when “prefer not to says” and “don’t know” answers were treated missing? Please include the model N to the tables.. - Also, it is not clear if was the vignettes presented were chosen by random? Were the covariates equally distributed among the vignettes presented? Results: - The results section does not include a single numeric reference to tables. It also and lacks analytical clarity. It would be important to report not only the associations but also the magnitude of the associations found. - To provide readers some more concrete results, it might be useful to conduct some analyses with a dichotomised fat phobia variable, i.e. the cut-point score FPS>2.5. (Ee.g. it would be very interesting to know what was the prevalence of “greater fat phobia” while using this cut-point, and how much higher was the prevalence for more males/low occupational position). - Please show also the coefficient of the covariates also in the results or appendix. Discussion: - The authors could be a more careful about making any conclusion regarding the interaction effects. The sample size was not necessary large enough to detect interaction effects. Minor/voluntary points - It might be useful to show the respondents’ characteristics by the vignette shown.

	- The abstract could be strengthened. For example, I am not sure if the statement that “there are studies focusing on obesity stigma in German public” provides any meaningful information regarding this study. - The ms would benefit from additional language edits. There are some confusing sentences and also incorrect ones. - It would be very interesting to see (perhaps in the appendix) stratified results by the respondents’ gender.

VERSION 1 – AUTHOR RESPONSE

REVIEWER #1

The research theme and questions of this article are interesting and important. Below are some suggestions for further improvement of the manuscript:

Thank you for your evaluation and your efforts.

1. It would be good to state clearly the rationale for selecting a share of 30% mobile numbers given the use of mobile-only in the population is 13%

Rationale for selecting a share of 30% is added on p. 6.

2. Please explain in detail what Kish-Selection-Grid method of selecting the person from this household.

Details are now given on p. 6.

3. It was not clear the methods that the study used to approach the subjects. Has a systematic way been adopted to reach the identified subjects given the low response rate?

The telephone survey was conducted by an experienced social research institute (USUMA, please see: http://www.usuma.com). As described in the manuscript, a random sample of telephone numbers was used. Sample consisted of registered private telephone numbers, additional computer-generated numbers, and randomly generated mobile phone numbers (Random Digit Dialing). Due to the method, there was a large number of neutral losses (e.g. wrong or non-working telephone number). Eight contacts on different days and at different day-times were made to find out whether a loss was neutral or systematic and to check eligibility. When a person or a household was reached, information about the study was given. For a random selection of participants in the households, the Kish-Selection Grid was applied (please see also point 2 of your review). Among mobile users, target person was the owner or main user of the mobile device. After having been informed that participation in the study is voluntary and that withdrawal from the study is possible at any time, 1,401 individuals participated. Compared to other national telephone surveys in Germany, a response rate of 49% can be regarded as satisfactory. Nevertheless, we consider a possible selection bias as a limitation (p. 16).

4. It was unclear the reasons that only 49.4% of the total interviewed subjects were included in the final analysis. This may lead to bias.

The reviewer is absolutely correct that we used 692 of 1,401 cases (i.e. 49.4% of the sample). The reason is that we used a vignette design. Vignettes were varied according to gender, SES and migration, resulting in eight different vignettes (23). The eight vignettes were randomly presented to about 175 respondents (p.7). In the present analyses, we wanted to focus on SES and gender differences. Thus, we used the four vignettes in which gender and SES were varied (p.6). As

assignment to the vignettes was made by random and subsamples are similar in terms of sociodemographic characteristics, we don't expect bias due to the use of the subsample.

5. No comparison of the sociodemographic information of the study sample and the population. Given the low response rate, it was unsure the representativeness of the study sample. Therefore one cannot claim that this is a representative cross-sectional survey study.

We now provide some information on comparisons of sociodemographic sample characteristics with official statistics on p. 9. Additionally, we provide information on comparisons with other (representative) studies on the prevalence of obesity (p. 9). In the discussion it is conceded that we cannot rule out selection bias due to non-response (p. 16).

6. Since comparisons were conducted for individual items, correction for multiple comparisons should be conducted.

We agree that the number of tests is relatively large. Therefore, level of statistical significance was defined at $p<.01$ instead of $p<.05$ (p.9). Respective changes were made throughout the manuscript and the tables.

7. What was the sequence of the vignette presentation? Any randomization of the sequence? As only one vignette was presented to each respondent, the proportion of subjects answering each vignette should be reported.

We now describe the vignette design in more detail on p. 6 and 7.

8. Since no neutral vignette was used, that is all vignette were an obese person, therefore it is difficult to conclude on the addictive effect of the stigma of gender or SES on obesity. The results and discussion may adjust based this limitation.

We agree. In the limitations we discuss that as follows (p. 17): "The lack of a neutral control condition impedes the interpretation of results. For example, it remains unclear whether respondents associate adjectives such as low self-esteem or insecurity with the fact that the individual in the vignette is obese or pursues the profession of a janitor when compared to a lawyer. This is a limitation that has to be considered when interpreting our findings as an indication of multiple or double stigma."

REVIEWER #2

The authors are to be commended for their approach to this question. The results will be of interest to readers. There are however a number of aspects of the manuscript requiring attention. I attach my comments for the authors' attention.

Thank you for your evaluation and your efforts.

1. Throughout the document: The authors refer to "public obesity stigma" this seems unnecessary, "obesity stigma" would suffice, although my preference would be "weight stigma" throughout the document.

We use the term "public obesity stigma" to make clear that we do not examine other stigma concepts like self–stigma. We agree that it is not necessary to do that throughout the manuscript and revised it accordingly.

2. There are inconsistencies in use of the term "obesity stigma" and "weight stigma", please address.

We now use the term "obesity stigma" consistently.

3. Many sentences commence “Moreover” suggest adjust to reduce occurrences.

We revised the manuscript accordingly.

4. When referring to obesity please refer to “people with obesity” or “female/women with obesity” or “male/men with obesity” (e.g. page 14 line 9 – 11, male obese, female obese).

We agree and revised the manuscript accordingly.

5. Page 1 line 3 – delete comma after gender in title?

Done.

6. Page 2 line 38-49. Please revise for clarity and can you remove the ‘assume’ reference?

Done.

7. Page 3 – consider referring to “pre-recorded audio vignette” to be clear.

Done.

8. Page 4 line 4 – delete ‘share’ prefer proportion or percentage or prevalence.

We use “proportion” now.

9. Page 5 line 9 – Reword first sentence for clarity.

We were not sure which sentence was meant, but we checked the whole manuscript for clarity.

10. Page 5 line 21-22 – Reword last sentence for clarity.

We were not sure which sentence was meant, but we checked the whole manuscript for clarity.

11. Page 7 line 12 – suggest add ‘audio’ to all audio vignettes

We added “audio” when the vignettes are introduced (p.7).

12. Page 7 line 18 – the term used in the vignette is “severely overweight” please correct comment

Corrected.

13. Page 7 line 25 – add in ‘self-reported’ weight and height

Done.

14. Results section – I have not conducted statistical tests, however the results as presented appear acceptable.

Thank you.

15. Page 14 line 52 – suggest re word possibly “However, these results can shed light on…”

Sentence was deleted.

16. Page 16 line 34 – the sentence “If affected by obesity…” needs to be clearer.

Sentence was deleted.

17. References: Generally very good, although I was somewhat surprised the authors did not include the following: World Health Organization. (2017). Weight bias and obesity stigma: considerations for the WHO European Region. Available at http://www.euro.who.int/en/health-topics/noncommunicablediseases/obesity/publications/2017/weight-bias-and-obesity-stigma-considerations-for-the-whoeuropean-region-2017

We thank you for suggesting this interesting work from the WHO that aims to give an overview of obesity stigma. We however realized that we already included most references that were mentioned in this WHO report.

18. The authors may also be interested to read the following, although not directly related to the topic of this manuscript, the article makes linkages between SES, gender, severe and morbid obesity in relation to a different issue: Gray L. Social Determinants of Health, Disaster Vulnerability, Severe and Morbid Obesity in adults: Triple Jeopardy? International Journal of Environmental Research and Public Health, 2017. 14, 1452.

Thank you for pointing to interesting additional references.

REVIEWER #3

The paper examines the association between various stigma indicators and the obesity vignettes. Overall, the topic and the research questions are certainly very important ones, and the data is valuable. However, I have some concerns regarding the method description, results and discussion section.

Thank you for your evaluation and your efforts.

Methods:

1. Several terms are used inconsistently making the manuscript confusing to follow. There is no name for the emotional reaction variables in the method section but in the results, these are described either as “negative/positive emotional reactions” (table3) or “scale prosocial / scale anger/disgust” (table 5). This is also the case with low occupational position and low SES (these are not synonyms). The manuscript would be much easier to follow if the authors would use terms consistently thorough the entire ms.

We agree that there were inconsistencies in the manuscript and revised it accordingly. We now consistently use the terms “negative emotional reactions” and “positive emotional reactions”. We also added a comment that the occupational position in the vignette (lawyer vs. cleaner) is used as an indicator of SES (p. 7). Therefore, we use the term SES throughout the manuscript.

2. Description of statistical analyses is inadequate: it would be useful to describe why given method is used. For example, the statement that “To test significant mean differences between… mann-whitney-u tests were applied” does not describe this method correctly. Why was mann-whitney-u chosen? How wereas the indicators distributed?

We referred to non-parametric tests since the investigated stigma responses for all components did not follow a normal distribution (p. 8). For your information, the distribution of stigma components can be viewed in the appendix (FIGURES_STIGMA DISTRIBUTION).

3. What was the analytical sample when “prefer not to says” and “don’t know” answers were treated missing? Please include the model N to the tables..

Information on sample sizes has been added to all tables throughout the manuscript.

4. Also, it is not clear if was the vignettes presented were chosen by random?

Vignettes were randomly assigned to respondents. (p.7)

Results:

5. The results section does not include a single numeric reference to tables. It also and lacks analytical clarity. It would be important to report not only the associations but also the magnitude of the associations found.

We carefully revised the result section and included numeric references (p.10-14).

6. To provide readers some more concrete results, it might be useful to conduct some analyses with a dichotomised fat phobia variable, i.e. the cut-point score FPS>2.5. (Ee.g. it would be very interesting to know what was the prevalence of “greater fat phobia” while using this cut-point, and how much higher was the prevalence for more males/low occupational position).

For the sake of clearness and comprehensibility we decided not to add more results. Moreover, Reviewer 1 mentioned that there may be a problem of multiple testing.

7. Please show also the coefficient of the covariates also in the results or appendix.

Full regression analyses with all included covariates are available in the appendix (TABLES_FULL REGRESSION ANALYSES) for reviewing purposes. We however decided not to add these full analyses to the manuscript since they are not fundamental for our research question and may limit the clearness and comprehensibility of our results.

Discussion:

8. The authors could be a more careful about making any conclusion regarding the interaction effects. The sample size was not necessary large enough to detect interaction effects.

We agree and added this point to the discussion (p.16).

Minor/voluntary points

9. It might be useful to show the respondents’ characteristics by the vignette shown.

For your information, please find a table showing the characteristics of the respondents by vignettes (gender/SES) in the appendix of this cover letter (TABLE_VIGNETTE CHARACTERISTICS).

10. The abstract could be strengthened. For example, I am not sure if the statement that “there are studies focusing on obesity stigma in German public” provides any meaningful information regarding this study.

We revised the Abstract accordingly.

11. The ms would benefit from additional language edits. There are some confusing sentences and also incorrect ones.

The manuscript has been checked by a native speaking colleague.

12. It would be very interesting to see (perhaps in the appendix) stratified results by the respondents’ gender.

Please see point 6.

VERSION 2 – REVIEW

REVIEWER	Sherry Kit Wa Chan The University of Hong Kong, Hong Kong
REVIEW RETURNED	15-Oct-2018

GENERAL COMMENTS	Thank you very much for the effort in doing the corrections.

REVIEWER	Lesley Gray University of Otago, Wellington, New Zealand
REVIEW RETURNED	24-Oct-2018

GENERAL COMMENTS	Overall, the results will be of interest to readers. Language and sentence layout still require further attention prior to publication. I attach marked up manuscript with tracked changes and comments. One specific question around the vignette's: what was the exact comment provided to respondents please? Strong or severe overweight? The vignettes as described for a person with an approximate BMI of 32 mean that the category of severe obesity was not reached (this would be BMI 35-39 according to usual BMI charts). If the term severe was used this may have overly influenced respondents responses and if so, this needs to be explained appropriately in the manuscript. The reviewer provided a marked copy with additional comments. Please contact the publisher for full details.

REVIEWER	Yevgeniya Gokun University of Arkansas for Medical Sciences
REVIEW RETURNED	13-Dec-2018

GENERAL COMMENTS	1. In the statistical analyses section, you indicated that Mann-Whitney U tests were applied since responses to stigma items didn't follow normal distribution. Please provide methods you used to test for normality. 2. In the same section, you indicated that determinants of stigmatizing attitudes were introduced into regression models. Please be more specific what type of regression models were used. 3. Table 2, 3 and 4: Mann-Whitney U Test (non-parametric test) is used. Mean and SD's are provided. It will help also to include medians as well as interquartile range. Also given you have decent sample size (almost 700 subjects), instead non-parametric testing, please perform two sample t-tests and see how its p-values differ from Mann-Whitney U tests. 4. Table 5: Linear regression was used for Fat Phobia Scale, Positive emotional and Negative emotional reactions as well as Social Distance. It would have been nice to include their distributions in the table to see if they satisfy normality assumption in order to proceed with linear regressions. You mentioned that BMI was adjusted as continuous variable into the regressions, would the results change if it was adjusted as categorical one?

VERSION 2 – AUTHOR RESPONSE

REVIEWER #2

Language and sentence layout still require further attention prior to publication. I attach marked up manuscript with tracked changes and comments.

Authors' response: Thank you very much for your helpful comments. We agree to your suggestions and revised the manuscript accordingly.

1. Reword "additive" to "additional"?

Authors' response: Wording was changed (p. 3, line 6).

2. "Live with obesity" instead of "who are obese"

Authors' response: Wording was changed (p. 4, line 1).

3. There are later publications – see Lancet 2016 for adult trends

Authors' response: We updated the first reference. Thank you (p. 4, line 2).

4. Delete "often"

Authors' response: The word "often" was deleted (p. 4, line 5).

5. Change "For Germany" to "One German study"

Authors' response: The sentence was revised accordingly (p. 4, line 30).

6. Add "In recent years"

Authors' response: "In recent years" was added (p. 5, line 11).

7. "This may be" instead of "a similar approach is"

Authors' response: The sentence was changed accordingly (p. 5, line 14-15).

8. Reword for clarity please

Authors' response: The sentence was rephrased for clarity (p. 5, line 15-18).

9. Reword for clarity please

Authors' response: The sentence was rephrased for clarity (p. 5, line 19-22).

10. Suggestion for an additional citation:

Authors' response: The suggested reference from Gray was added to the manuscript (p. 5, line 22-24).

11. Rephrase the sentence.

Authors' response: The sentence was rephrased (p. 5, line 24).

12. Delete the word "especially".

Authors' response: The word "especially" was deleted (p. 5, line 28).

13. Suggestion: Change “we incorporated” to “was incorporated”

Authors’ response: Wording was changed (p. 6, line 3).

14. Add “the”.

Authors’ response: Done (p. 6, line 5).

15. Delete “to this aim”

Authors’ response: “To this aim” was deleted (p. 6, line 7).

16. Change “the respondent” to “one person”

Authors’ response: Wording was changed (p. 6, line 10).

17. Change “in the beginning” to “at the start”

Authors’ response: Wording was changed (p. 6, line 10).

18. Delete “N” and insert “persons”.

Author’s response: Wording was changed (p. 6, line 12-14).

19. Suggestion: Shorten the sentence “It is comprised of 14 items and constitutes the first factor of the original 50-item scale”.

Authors’ response: As suggested, the sentence was shortened (p. 7, line 23-24).

20. Delete “the”.

Authors’ response: Done (p. 7, line 28).

21. Rephrase “German speaking area”

Authors’ response: “German speaking area” was changed to “applied in German by Luck-Sikorski and colleagues” (p. 7, line 29).

22. All or some items inverted?

Authors’ response: “Some” items was added to the sentence (p. 7, line 2).

23. Change “sex” to male:female.

Authors’ response: Wording was changed accordingly (p. 9, line 10).

24. Change “respondents are either overweight or obese” to “respondents reported overweight or obesity”.

Authors’ response: The sentence was revised accordingly (p. 9, line 15).

25. Change “The share of those who are overweight/obese to “The share of those with are overweight/obesity.

Authors’ response: Wording was revised (p. 9, line 16).

26. Be specific, provide percentage or similar.

Authors’ response: Percentages are now provided (p. 9, line 17).

27. Table 1: In data analysis – were you able to stratify according to respondents own weight status (and did this make a difference?) – there could be potential for individual's internalised bias to affect responses. (p. 10).

Authors' response: The question of internalized stigma is indeed very interesting and of general interest in the field of stigma. To document stratified analyses for different weight status groups (normal weight, overweight, obese), another 12 regressions and four tables would be needed. Given that our research explicitly aimed at the question of multiple stigma, we decided not to further discuss the role of one's own weight for stigma. Nonetheless, we are very grateful for this advice and will attempt to further elaborate the role of internalized stigma in future studies.

28. Reword "person who obese and has a low SES" to "person with obesity and a low SES"

Authors' response: Sentence was revised accordingly (p. 12, line 11).

29. In introduction/background you indicate this is the first study, here in discussion you indicate this is one of the first – please adjust for consistency throughout manuscript.

Authors' response: The sentence was changed to "The present study is the first to analyze" (p. 15, line 10).

30. This does not make sense – please clarify.

Authors' response: The sentence was deleted to avoid misunderstandings (p. 15, line 18-20).

31. Be clear – how many years? Or an indication of change over decades.

Authors' response: The information "over the past decade" was added (p. 15, line 22).

32. Change "Next to being obese" to "next to obesity".

Authors' response: The sentence was changed accordingly (p. 15, line 31).

33. See my earlier comment re analysis of respondent BMI and implications around internalised bias.

Authors' response: Please see our response to point 27.

34. Re-word for clarity.

Authors' response: "inacceptable" was changed to "limited reliability" (p. 16, line 13).

35. Change "that the individual in the vignette is obese" to "individual presented with obesity" (p. 16, line 22).

Authors' response: wording was changed accordingly

36. Maybe you meant "counteract"?

Authors' response: "To encounter" was changed to "counteract" (p. 17, line 7).

37. One specific question around the vignette's: what was the exact comment provided to respondents please? Strong or severe overweight? The vignettes as described for a person with an approximate BMI of 32 mean that the category of severe obesity was not reached (this would be BMI 35-39 according to usual BMI charts). If the term severe was used this may have overly influenced respondents responses and if so, this needs to be explained appropriately in the manuscript.

Authors' response: The comment provided to the respondents was: "With a height of 5'5 and a weight of 200 pounds, she/he is severely overweight." (please also see the vignettes in the appendix of the

manuscript). We decided to use the term “severely overweight” in our survey because the term “obesity” is not very common in the general population. This holds especially true for Germany, where the translated word is “Adipositas”, a term that is used predominantly in the medical context and is not well understood in the general population. Moreover, we do not expect the general public to be familiar with the exact definition of obesity and severe obesity. Furthermore we do not expect the respondents to calculate the BMI based on the information provided on height and weight. Thus, we decided to use the term “severely overweight” to describe an obese person without using the words ”obesity” or “obese”. In this regard, it was not our intent to depict a person with “severe obesity”. Nevertheless, we agree that the lack of medical accuracy of the vignettes is a further limitation of our study that we now mention in the discussion section (p. 16, line 22-24).

--

REVIEWER #4

1. In the statistical analyses section, you indicated that Mann-Whitney U tests were applied since responses to stigma items didn't follow normal distribution. Please provide methods you used to test for normality.

Authors’ response: The requested information (Kolmogorow-Smirnow-Tests) was added to the statistical analyses section (p. 8, last paragraph).

2. In the same section, you indicated that determinants of stigmatizing attitudes were introduced into regression models. Please be more specific what type of regression models were used.

Authors’ response: Information on the type of regression models were added to the statistical analyses section (p. 8, last paragraph).

3. Table 2, 3 and 4: Mann-Whitney U Test (non-parametric test) is used. Mean and SD's are provided. It will help also to include medians as well as interquartile range.

Authors’ response: After including medians and interquartile range to Table 2, 3 and 4, we realized that the values were quite similar across all stigma items due to the limited range of possible responses. As a consequence, these values provide little additional information to the aforementioned tables. Therefore, we would like to keep the original tables as these changes rather limit the readability of our tables. For your information, we document medians and interquartile ranges in the Tables X2 to X4 in the Appendix of this cover letter.

4. Also given you have decent sample size (almost 700 subjects), instead non-parametric testing, please perform two sample t-tests and see how its p-values differ from Mann-Whitney U tests.

Authors’ response: Two sample t-tests with our stigma items only revealed small aberrations from non-parametric testing (Mann-Whitney) and had no further implications for the interpretation of our study results and conclusions. For your information, two sample t-tests in all stigma items can be retrieved as an Appendix of this cover letter (Table Y2 to Y4).

5. Table 5: Linear regression was used for Fat Phobia Scale, Positive emotional and Negative emotional reactions as well as Social Distance. It would have been nice to include their distributions in the table to see if they satisfy normality assumption in order to proceed with linear regressions.

Authors’ response: As for the satisfaction of the normality assumption in linear regression analyses, we inspected residuals in P-P-Plots. Results indicate that residuals satisfy normality assumption (please see the Appendix of this cover letter: Figure Z1 to Z4).

6. You mentioned that BMI was adjusted as continuous variable into the regressions, would the results change if it was adjusted as categorical one?

Authors‘ response: When inserting the respondent’s weight status as a categorical variable (normal weight (as reference), overweight, obesity) into the linear regression analyses, only small differences compared to the former analyses appeared (please see the Appendix of this cover letter: Table Z1 to Z4).

VERSION 3 – REVIEW

REVIEWER	Lesley Gray University of Otago, Wellington, New Zealand
REVIEW RETURNED	28-Jan-2019

GENERAL COMMENTS	The authors have attended to the amendments required. This is an interesting paper. The authors will wish to review the following article (just published) for relevance/alignment with the manuscript as it pertains to a rise in bias in relation to body weight, and reductions in bias for sexual orientation and race: Charlesworth, T. E., & Banaji, M. R. (2019). Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. Psychological science, 0956797618813087.

REVIEWER	Yevgeniya Gokun University of Arkansas for Medical Sciences
REVIEW RETURNED	23-Jan-2019

GENERAL COMMENTS	1. You indicated that Kolmogorv-Smirnov Tests revealed that the responses to the stigma items didn't follow a normal distribution. Did you verify these findings with other methods such as Shapiro Wilk Test or Q-Q plot to make sure all the methods were giving you consistent results? 2. Given that your sample size is over 300 subjects, have you tried to do two sample t-tests for Table 2 through Table 4 to see whether it is giving you similar (either statistically significant or statistically insignificant) p-values? 3. Table 5--performance of multivariate linear regressions among these 4 outcomes (Fat Phobia Scale, Positive emotional reactions, Negative emotional reactions and Social Distance). I would like to see each one of these scales' distributions (mean, SD, median, IQR) in table 1. Please justify that each of the outcomes is normally distributed if you are choosing to use linear regression to model them (perform testing for normality for all four outcome scales using various methods).

VERSION 3 – AUTHOR RESPONSE

Reviewer #2

1. The authors have attended to the amendments required. This is an interesting paper. The authors will wish to review the following article (just published) for relevance/alignment with the manuscript as it pertains to a rise in bias in relation to body weight, and reductions in bias for sexual orientation and

race: Charlesworth, T. E., & Banaji, M. R. (2019). Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016.

Authors‘ response: Thank you for suggesting another interesting paper that provides new insights on the development of weight bias over time. Once again, thank you for reviewing our manuscript and the helpful remarks.

Reviewer #4

2. You indicated that Kolmogorv-Smirnov Tests revealed that the responses to the stigma items didn't follow a normal distribution. Did you verify these findings with other methods such as Shapiro Wilk Test or Q-Q plot to make sure all the methods were giving you consistent results?

Authors‘ response: As requested, we performed additional testing for normal distribution with the Shapiro-Wilk-Normality Test. However, the results from this test did not reveal any differences in single stigma items or sum scales, if compared with the Kolmogorow-Smirnow Test (in terms of statistical significance). The complete table which compares the Kolmogorow-Smirnow-Test and Shapiro-Wilk-Normality test can be viewed in the Appendix of this cover letter (Table X1; exact p-values are provided in scientific notation since numbers were too small to be conveniently expressed in decimal numbers).

3. Given that your sample size is over 300 subjects, have you tried to do two sample t-tests for Table 2 through Table 4 to see whether it is giving you similar (either statistically significant or statistically insignificant) p-values?

Authors‘ response: By performing two sample t-tests for all included stigma items, results only revealed small aberrations from non-parametric testing (Mann-Whitney) with no further implications for the interpretation of our study results and conclusions. For your information, the additional information on two sample t-tests in all stigma items can be retrieved as an Appendix of this cover letter (Table X2 to X4).

4. Table 5: performance of multivariate linear regressions among these 4 outcomes (Fat Phobia Scale, Positive emotional reactions, Negative emotional reactions and Social Distance). I would like to see each one of these scales' distributions (mean, SD, median, IQR) in table 1.

Authors‘ response: As requested, we have now integrated information on distribution of all stigma sum scales with mean, SD, median and ICQ in Table 1 of the manuscript.

5. Please justify that each of the outcomes is normally distributed if you are choosing to use linear regression to model them (perform testing for normality for all four outcome scales using various methods).

Authors‘ response: We assessed the distribution of residuals for each regression analyses (fat phobia, positive emotional reactions, negative emotional reactions, desire for social distance) using standard diagnostic procedures (1. P-P-Plots, 2. Q-Q-Plots, and 3. normality plot of residuals; please see Figure Z1-Z4 in the Appendix of this cover letter). Although there are some deviations, results indicate that, in overall, residuals satisfy normality assumption.