

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Are They Half as Strong as They Used to Be? An Experiment Testing Whether Age-Related Social Comparisons Impair Older People's Hand Grip Strength and Persistence
<b>AUTHORS</b>	Hannah Swift, Ruth A. Lamont & Dominic Abrams

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Dr Rhiannon Turner Institute of Psychological Sciences University of Leeds
<b>REVIEW RETURNED</b>	02/03/2012

<b>GENERAL COMMENTS</b>	<p>This article examines whether social comparisons with younger people result in physical impairments on a commonly used diagnostic test of capability, hand grip strength. Older adults were randomly assigned to one of two conditions: either they were told they were taking part in a performance test or, in the social comparison condition, that their performance was being compared to younger people, who usually perform differently on such tests. Participants then gripped a manual hand dynamometer, and the strength and persistence of their grip was assessed. Participants who believed their performance was being compared with that of younger people performed significantly more poorly in terms of both strength and persistence, an effect that held even when controlling for age, gender, education level, and degree of arthritis.</p> <p>This is an excellent and fascinating piece of research, with powerful findings that have clear implications for wider society, particularly with respect to medical diagnosis, and perceptions of older adults, by themselves and by the rest of society. Crucially, the magnitude of the effect is huge: an age comparison created a stereotype threat effect which reduced older people's hand grip strength by up to 50% - the authors note that this is equivalent to the reduction in ability between middle age and old age. If we overestimate the extent of disability in older adults because of social comparison, this has concerning implications for their welfare, potentially fostering dependency and helplessness, but also perpetuating the negative stereotype that older adults are physically weak. By integrating two fields of research – applying social psychological theory to the health domain – the authors have produced findings which will be of real interest to both psychologists and physicians alike.</p> <p>As the authors note, the study involved a healthy sample of participants rather than older adults in ill health, but given that controls for arthritis in the current study did not diminish the stereotype threat effect, there is no reason to believe that the results would differ were the study to be carried out with less able older</p>
-------------------------	---

	adults.
<b>REVIEWER</b>	<p>Kevin McKee  Professor of Gerontology  School of Health and Social Studies  Dalarna University  Falun, Sweden</p> <p>There are no competing interests relating to this review.</p>
<b>REVIEW RETURNED</b>	14/03/2012
<b>THE STUDY</b>	<p>1. There are relatively few details about the participants. It is not clear, for example, why the recruitment locations were chosen from the options available, nor from what sampling frame they were selected. What were the inclusion/exclusion criteria for locations and for participants within locations? How many older people were approached, how did the approach occur? How many consented to participate in the study, and what information &amp; consent procedures were followed? When did the random assignment to condition occur, and when and where did the testing take place? Were participants alone (with the investigator) when tested, or in a public space? I appreciate that the manuscript is constrained in terms of length, but given that the authors' primary message is that the social context is very important for individuals' behaviour, so a clear description of the way in which the participants came to be in the study, and the manner of their testing, is also very important for interpreting the study's findings.</p> <p>2. Did the participants provide the study data (demographics etc.) through a structured interview or via questionnaire self-completion? From where were the items on residence, education, etc. obtained? In particular, from where was the item on arthritis obtained?</p> <p>3. The abstract could be a little better balanced: the conclusions are a little lengthy at the expense of more information about the study question, design and methods. I would delete the first line of the conclusions, and start the conclusions with something like: 'Due to the potential for age comparisons and negative stereotype activation during assessment of older people, such assessments may underestimate ....., etc. The 'key messages' I think could be a little better ordered and expressed. I would suggest the first message should be the general statement, i.e., 'Psychosocial factors strongly influence how the physical effects of ageing manifest themselves at the individual level' My second key message would be the current first key message, and my third key message would be a reworded version of the current third key message, i.e., 'health care professionals should be aware of the potential for age comparison and stereotype threat during assessment of older patients'. Finally, the sentence 'Age comparison....older people' in 'strengths and limitations' doesn't actually refer to a strength or limitation of the study – it's more a 'key message'! I would delete the line, and mention the small, convenience sample as a limitation, but a limitation countered by the use of population norms to control for this in the analyses.</p> <p>4. While I am happy with the analyses as a whole, I think there are three issues that perhaps the authors should reflect on. First, the recruitment from a number of different locations (we are not told how many) means that the study is effectively a clustered design. This means that the analyses might be under-powered (although there is no power calculation to justify the number of participants, so it might</p>

	<p>be that the investigators have over-sampled). Happily the authors obtained significant effects in their analyses anyway. Second, as the design is clustered, multi-level modelling should be used to partition out the effect of the clustering and ensure the estimates for individual-level effects are appropriately adjusted. Now, because of the small sample size I expect this would be problematic (i.e., to run MLM analyses where there are very few individuals per location). But I think this problem should be highlighted. It's a similar issue with my third point: because of multiple testing, the potential for Type I error is inflated. No adjustment for multiple testing is made, and while I don't necessarily think that is required, the authors should point out the inflated Type I error risk to readers. A final minor point is that I would chose Chi square as a test of the association between group membership and gender (Table 1). Presenting a 'mean gender' score, though perhaps technically leading to the same conclusion, seems rather odd.</p>
<b>RESULTS &amp; CONCLUSIONS</b>	<p>1. Very minor issues: Table 2 is poorly formatted, with the columns of rather random width. And generally in the tables, numbers should always be presented to two decimal places (for example, the SD given for arthritis in Table 2 is 1 – this should be presented 1.00)</p>
<b>GENERAL COMMENTS</b>	<p>This was an interesting and valuable piece of work, with important implications for clinical settings. Could I ask the authors to perhaps reflect in the manuscript on the following issues:</p> <p>1. The authors mention the potential importance of the study to health care professionals, but of course many assessments (some with physical aspects) are carried out by social care professionals, many in non-clinical settings. One could argue that the current study has as much or even greater significance for social care professionals and their assessment procedures than clinicians and clinical assessments, so perhaps a line or two on this would be useful?</p> <p>2. The statement on page 11 that 'one of the factors that is likely to make older people vulnerable to negative comparisons is age segregation' is slightly problematic. Does age segregation in itself make people vulnerable to negative comparisons? Well, if the segregation is absolute, no, as there would be no comparison group. Rather the issue is that age segregation creates in-groups and out-groups partly based on age, which in turn inflate the potential for group comparisons. Here, though, one could also ask why the comparisons should be negative? Under certain conditions, one could posit that age comparisons on dimensions such as experience, wisdom, kindness, etc., might be beneficial to older people. It would be the availability of stereotypes and the dimension on which comparison is made that would inform the comparison, and thereby the valence of the effect. Similarly with regard to the previous sentence, it would surely be the conditions under which older people maintain contact with younger people that influence their vulnerability to age comparisons? As an example, imagine a mixed-age social group in which membership is based on the capacity for physical activity – say a work rugby team! One could posit that older members of this group might be very vulnerable to age comparisons, because the activity of the group, and group performance, make age salient. Essentially I have no problem with the authors' argument (which is based on empirical work that they have published; they no doubt have a far better knowledge base than I on these particular, complex issues); rather the argument is presented rather starkly, whereas there are many inter-related factors at play. Perhaps there should be a few more lines provided to draw out the subtleties of the situation.</p>

<b>REVIEWER</b>	<p>Mette Aadahl Senior researcher, PT, MPH, PhD Research Centre for Prevention and Health Glostrup University Hospital Building 84/85, DK- 2600 Glostrup, Denmark</p> <p>I declare that I have no conflicts of Interests.</p>
<b>REVIEW RETURNED</b>	19/03/2012

<b>THE STUDY</b>	<p>This study presents an important research question that may have an impact on the way we conduct tests and interpret test results in older people. It is indeed an interesting field of research, however, I have some questions in relation to the methods in the study that I would like the authors to consider:</p> <p>1) The title says 'A field experiment'....., but the study appears to be conducted as a randomised trial with a control group. I suggest the authors include a mention of this in the title.</p> <p>2) The methods section does not include information on the randomisation procedure. TOP page 8 it says ' participants were randomly assigned to condition'. Please describe how.</p> <p>3) Recruitment of participants and inclusion/ exclusion criteria are not described.</p> <p>4) The method section should include a separate 'statistical analyses' section, where all the statistical methods are described, instead of reporting them in the 'results' section.</p>
<b>RESULTS &amp; CONCLUSIONS</b>	<p>As mentioned above, the results section should be 'clean' and not include description of the applied methods.</p> <p>Also, I recommend the following changes to the tables:</p> <p>Table 1: Please re-design table, substituting 'Gender' (1=female, 2=male) and 'Self-reported arthritis' with frequencies (i.e. number and %) instead of mean(SD). It is easier to read, makes better sense and hence, is more polite to the reader.</p> <p>Table 2: Please include units. (e.g. for grip strength= Kg, and persistence= time). Here again it makes no sense to have a mean gender or mean arthritis. I suggest you include correlations only in this table and report the total means, sd and frequencies elsewhere, e.g. in table 1 ? in a separate column for ?</p> <p>Table 3: I would prefer having the three statistical models for hand grip strength presented together and the models for persistence presented together.</p> <p>Please present first model as crude model and the following as adjusted for relevant co-variables.</p> <p>I am not sure whether the third model is adjusted for all variables or just for age and gender norms and education level?? Please clarify.</p> <p>Figure 1 &amp; 2: Please consider whether these figures contribute any information that is not already presented in the tables? If not the figures are not justified, and the 95% CI could be added in the table.</p>
<b>REPORTING &amp; ETHICS</b>	<p>3) There is no mention of ETHICS in relation to participation in the present study. Please include.</p>

### VERSION 1 – AUTHOR RESPONSE

1. Title now changed as requested by managing editor and R3. Strictly speaking this is not an RCT in that we do not administer a treatment. Rather it is an experiment testing the effect of different conditions on behaviour. The new title is 'Are They Half as Strong as They Used to

Be? An Experiment Testing Whether Age-Related Social Comparisons Impair Older People's Hand Grip Strength and Persistence'.

2. Abstract. We have redrafted the abstract and key messages as suggested by Reviewer 2.

3. R2, 3. Recruitment of participants. Participants were asked to volunteer so there was no question of refusal to participate because all of the testing sessions were filled. Demographics were measured rather than serving as a basis for recruitment. Fuller details of the participant recruitment and procedure, and a note that the study complied with appropriate ethical standards are now provided on page 8.

4. R2 – units of measurement for strength and time are indicated on page 9.

5. R2 demographics and arthritis measurement is explained as being by interview. We have also inserted more information about how these were coded.

6. R2. As noted by the reviewer there are insufficient numbers of participants or testing locations to conduct a multilevel analysis. However, we have now included evidence to show that random assignment to condition was successful for both gender and location, and we report the result of a multilevel analysis showing that once location is accounted for it makes no difference to the effect of condition. It is not statistically appropriate to run a full multilevel analysis with less than 30 groups so we feel it is best not to foreground that analysis. As R2 had helpfully noted, any effects of location would be likely to reduce rather than strengthen the differences between conditions, consistent with Type 1 error, suggesting that, if anything, the effects we obtained were an underestimate of the true effect. We have noted this in the Discussion.

7. R2 and R3 Gender distribution and type of residence are now reported using chi square in Table 1.

8. R2 and R3 Table 2 has been tidied and corrected for layout and the variables now follow the same sequence as Table 1. We have removed the means and standard deviations but inserted ns for gender and residence as they are binary variables.

9. As suggested by R3 we have placed a Statistical Analyses section at the end of the methods section and removed duplicate material from the Results section.

10. R3 suggested we should present two sets of hierarchically sequenced analyses in Table 3. This has now been done. The information on page 10 (in the original version) states that all variables are in the final ANCOVA.

11. We considered whether to remove the figures. However, because interactions are difficult to judge from numerical data, we feel that the findings are presented vividly using the figures and would prefer to retain them.

12. R2 We have expanded the section on the implications of segregation as suggested.

We hope that these changes meet with your approval.

Yours sincerely

Dominic Abrams, Hannah Swift and Ruth Lamont

**VERSION 2 – REVIEW**

<b>REVIEWER</b>	Mette Aadahl, Senior researcher, MPH PhD, REsearch Centre for Prevention and Health, Denmark  There are no competing interests.
<b>REVIEW RETURNED</b>	16/04/2012