

PEER REVIEW HISTORY

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

ARTICLE DETAILS

TITLE (PROVISIONAL)	Understanding the relative valuation of research impact: A best-worst scaling experiment of the general public and biomedical and health researchers
AUTHORS	Pollitt, Alexandra; Potoglou, Dimitris; Patil, Sunil; Burge, Peter; Guthrie, Susan; King, Suzanne; wooding, Steven; Grant, Jonathan

VERSION 1 - REVIEW

REVIEWER	Kai Ruggeri University of Cambridge, UK
REVIEW RETURNED	09-Jan-2016

GENERAL COMMENTS	<p>Thank you very much for the opportunity to read this incredibly important paper. The general topic and your key findings are hugely important on a vast number of levels, and I enjoyed going through it very much.</p> <p>That said (you knew this was coming), there are a large number of changes that I would strongly recommend. This paper would probably benefit by just cutting a lot of the measure development and method (which would knock off quite a few of these), perhaps publishing those elsewhere, and really focus on the public/researcher differences. That is really the compelling argument, in my view. The rest reads a bit more like a thesis, where all information seems to simply be compiled, even when most may not really be that necessary to include. However, as this is a very important topic and your main findings are definitely of value, I have provided comprehensive feedback. I would strongly urge the authors to remove a large amount of the writing not central to the results, and perhaps consider publishing those separately as a methods paper. This is not because they are unimportant, but because the key outcomes do appear to be buried with all of it there.</p> <p>Using the review checklist, I explain my selection for 1-12 below. I would hope these would provide a clear checklist of revisions that will hopefully lead to a very strong final publication, whether in this journal or elsewhere.</p> <ol style="list-style-type: none"> 1. Though this is a discrete choice design, it is not really accurate to present this study as an experiment. No controls exist and no conditions are altered. I say this as a person who highly values and uses DCM on a regular basis. Easily fixed, though. 2. The abstract is a bit confusing and indirect. The opening objectives statement uses 'and' three times. Would suggest being more direct. The design is really a survey. It says the setting is the United Kingdom but (unless I have missed it) there is no clear aspect of the method or analysis that ensures this to be the case. I checked a few
-------------------------	---

	<p>times but do not find confirmation in either way, though this may be my mistake. The results are not results but generalized conclusions. It would help if this actually had specific results, e.g. 'BWS results showed that...', 'general public has a stronger preference for x,y,z; researchers value a,b,c'. The conclusion should then reflect on those implications – the points started are mostly generalisations about the study.</p> <p>3. This is a yes but could certainly be improved in the writing by addressing point 1.</p> <p>4. Yes. Perhaps too much detail and reliance on appendices, though. A methods paper might be more appropriate.</p> <p>5. I'm concerned about the ethics – it says they followed guidelines but nothing about actual review. Could this be clarified? I realise this is not a high-risk study so I wouldn't necessarily make any major recommendations, but following guidelines is different from seeking approval. I raise this largely because the statement made about ethics in the manuscript about not requiring review from some institutions, but those institutions being clearly listed in the survey introduction that participants would see, and that the link provided is for King's College, London, who did not review it. Again, this is not a risky study, so perhaps just some clarification.</p> <p>6. This is actually the biggest concern of the current version – while the importance of the topic and the overall methodological quality seem to be very clear, the writing on actual findings is very complex, indirect, and hard to extract the meaning. Also, it feels very unclear where the method ends and the results begin – I realise there are multiple studies combined into a single paper, but it should be much more obvious where things are as it leads to the previous point. Hence several points here might better belong in the question about design.</p> <p>a. First, for the focus groups, the 'key observations' are not explicit on what was a discussion point generally and what was the actual finding. These are interesting remarks to read, but it would be very helpful to distinguish between what was asked of the group, what was observed, and then what has been extracted.</p> <p>b. The same goes for the interviews – a more engaging structure for presenting the three elements mentioned for the FGs would be really beneficial to the article.</p> <p>c. Could the authors also state how they define 'cognitive interview'? Was this a validation check of the measure or something further? I think this also is not an appropriate place to refer to a supplement, but rather to provide the most relevant insights. However, this is one of several areas where the paper leans more toward metrics and methods rather than scientific outputs.</p> <p>d. For the pilot, it is hard to see how piloting a survey can lead to a conclusion that the questionnaires were generally well understood. Was there a feedback mechanism and/or how did this differ from cognitive interviews? I see it is mentioned later but it would help if it could be clarified (or point me to the explanation if I have missed it).</p> <p>e. Given that this paper is aiming to present some hugely relevant findings, it seems the information about how the method/measures were finalised should really be published as a paper in order to establish validity of this work. There are really far too many supplemental sections being relied on – hence they probably deserve equivalent attention or parts of this manuscript broken apart as separate papers.</p> <p>f. Were any validation checks done of the crowdsourced information? Having used these services myself, this seems highly critical. Is the population representative of the UK (ignoring the earlier point about location for now)?</p>
--	--

	<p>g. While I do not disagree with the approach to excluding surveys on quality checks, is there any way to justify these criteria? They seem fine in practice, but do appear a little bit arbitrary (apart from the last one – no explanation needed).</p> <p>h. Is there precedent for comparing two different model approaches? It seems odd that there was no difference in the model fit and this resulted in choosing one biased and one naïve sample. It would seem that using both excluded or both inclusive is the preferred option if any two were equivalent, and you would only want to use one of each if the data forced it. However, I am happy to be convinced otherwise.</p> <p>i. The entire modelling section could use a human touch. It is really unclear how someone’s response on a survey can be converted into coefficient related to life expectancy, and the examples given do not have clearly explained values or origins. Why are results not simply expressed as asked in the survey? As I had to regularly bounce back and forth between the supplement pages and the section, I may have missed it, but the best interpretation I could make is that certain answers come in with known weights based on known evidence about those items’ real-world impact. Could this please be addressed?</p> <p>j. At least in the version I could see for the review, Figure 2 does not have sufficient resolution to determine the difference between shaded and not shaded. The pattern is well-presented overall, though.</p> <p>k. Box 2 is a nice way to present findings, but I would strongly request that this be reduced to a small number, and then present the rest in text. Or come up with another way to present this. I do like consolidating key messages, but this essentially becomes almost a text within the text.</p> <p>l. Some interpretations of the findings as ‘more’ and ‘less’ seem a bit inaccurate. As this is choice modelling study, should it not be discussed as a matter of ‘greater proportion’ and ‘lower proportion’? In other words, are certain things valued more, or simply do more people value it from one group? Please clarify this if I am mistaken. I was trying to check if the statistical technique might clarify this distinction but I was not 100% sure in the end, so felt necessary to raise.</p> <p>m. The same point about equal value – should it not just be ‘similar proportions’?</p> <p>n. It seems that key observation 7 belongs in a much more prominent place as this is clearly a highly relevant finding.</p> <p>o. KO #8 is an enjoyable finding. I look forward to the media picking up on that...and then using it in my public lectures.</p> <p>p. KO #9 is highly relevant, please also highlight this.</p> <p>7. I would actually request a slightly more lay explanation of the analyses. It is passed over rather briefly in the method and a few general statements on the process would really help with grasping the work done. I think it’s a bit risky to rely on the supplement, unless there was a more clearly described link to a published method applied precisely here.</p> <p>8. No major concerns with referencing, though it might not hurt if a bit more scientific theory was introduced into the introduction and/or reflected on in the discussion.</p> <p>9. Skip – all in point 6.</p> <p>10. If above points are addressed, then this should be fine.</p> <p>11. The discussion is fine but personally I would strongly suggest reflecting more on the 9 key observations and critiques therein, and far less on the method. The findings are really important and at least three of them are worth a page of reflection, in my view.</p>
--	--

	12. I think sample-only limitations is not particularly self-critical but as I was not party to the study itself, cannot say what else might go in there.
--	---

REVIEWER	Anna Williamson The Sax Institute, Australia
REVIEW RETURNED	07-Feb-2016

GENERAL COMMENTS	<p>This paper addresses the important, and poorly defined, topic of research impact. It is clearly written, well conducted and represents a useful addition to the field.</p> <p>I have no major comments but suggest the introduction may benefit from a brief outline of some of the key ways in which research might make an impact.</p> <p>While it is briefly alluded to, I think it is also worth noting more pointedly that another major piece of work missing in this area is the development of robust methods for assessing whether or not an impact has occurred.</p>
-------------------------	---

VERSION 1 – AUTHOR RESPONSE

Reviewers' Comments to Author:

RESPONSE:

We thank both reviewers for their thoughtful comments and valuable suggestions for improving the clarity and usefulness of the paper – and we are delighted that they found it both interesting and enjoyable. We have addressed each comment in turn below.

We did note, however, that reviewer 2 appeared content with the methodological aspects of our paper. In light of this, while we have carefully considered all of the suggestions made by reviewer 1 (and in many cases these have been very helpful), we did not always feel it appropriate to make extensive changes to the manuscript. We were also cautious of adding in too much more on the methods, given concerns over the length of the paper and, as raised by the reviewer, the balance between methods and findings. Further detail is provided in the responses to specific comments below.

Reviewer: 1

Thank you very much for the opportunity to read this incredibly important paper. The general topic and your key findings are hugely important on a vast number of levels, and I enjoyed going through it very much.

That said (you knew this was coming), there are a large number of changes that I would strongly recommend. This paper would probably benefit by just cutting a lot of the measure development and method (which would knock off quite a few of these), perhaps publishing those elsewhere, and really focus on the public/researcher differences. That is really the compelling argument, in my view. The rest reads a bit more like a thesis, where all information seems to simply be compiled, even when most may not really be that necessary to include. However, as this is a very important topic and your main findings are definitely of value, I have provided comprehensive feedback. I would strongly urge the authors to remove a large amount of the writing not central to the results, and perhaps consider publishing those separately as a methods paper. This is not because they are unimportant, but because the key outcomes do appear to be buried with all of it there.

RESPONSE:

We appreciate the reviewer's concerns about the amount of methodological description and have made a number of changes to the text as a result (detailed below). However, we have decided not to split the paper into separate methods and findings pieces for the following reasons:

- As this was an exploratory study applying a novel approach to the area of research impact, developing the method was a core part of the study. Indeed, the study was funded through a methods panel at the MRC. We state in the introduction that there were two aims of the study: contributing methodologically to the assessment of research impact and developing our understanding of how different stakeholders value different research impacts. Our hope is that others take this as a starting point to further develop methods in research impact assessment, an area in which there is a strong demand for better tools, and that to do so they would benefit from having all of the relevant information in one paper.

- Given the novelty of the method, we feel it is important to document our approach in detail, allowing others to critique, replicate or improve it. For this reason we selected a journal which would allow these details to be published alongside the main text as supplementary information. We will also make our data freely available to coincide with publication of the paper.

That said, we appreciate that there are steps we could have taken to help readers navigate the large quantity of information and that we could have documented more clearly how the various parts of the methods contributed to the development of the survey. As such, we have made the following changes:

- The methods section has been more clearly structured to reflect the stages of developing and testing the survey. Supplementary files have been aligned with these sections, reducing the total number and ensuring that all of the information on, for example, defining and categorising impacts, is contained in a single file.

- We have clarified how each step of the method affected the final survey tool (i.e. the changes made as a result of it), as well as adding more specific detail on the implementation of the qualitative stages of survey development. This has been balanced with a consideration of the overall length of the paper, though.

- Some of the technical detail on the methods has been moved to the corresponding supplementary file, to improve the balance between methods and findings, and to keep the paper to a reasonable length.

These points and other specific changes are detailed below, in response to the individual comments from the reviewer.

R1: Using the review checklist, I explain my selection for 1-12 below. I would hope these would provide a clear checklist of revisions that will hopefully lead to a very strong final publication, whether in this journal or elsewhere.

1. Though this is a discrete choice design, it is not really accurate to present this study as an experiment. No controls exist and no conditions are altered. I say this as a person who highly values and uses DCM on a regular basis. Easily fixed, though.

RESPONSE:

The term 'discrete choice experiment' (also known as 'choice experiment') is used to refer to a domain

of stated preference elicitation methods used to elicit individual preferences in the absence of actual (revealed) preference data. The term has been widely used across many disciplines including Health and Environmental Economics, Marketing and Transport (Louviere et al., 2010; Hensher et al., 2005). The term 'experiment' reflects the process by which stated choice questions are allocated within respondents from a carefully selected experimental survey design.

- Louviere, J.J., Flynn, T.N. and Carson, R.T. (2010) 'Discrete choice experiments are not conjoint analysis', *Journal of Choice Modelling*, 3(3), 57-72.
- Hensher, D.A., Rose, J.M. and Greene, W.H. (2005) *Applied Choice Analysis - A Primer*, New York: Cambridge University Press.

2. The abstract is a bit confusing and indirect. The opening objectives statement uses 'and' three times. Would suggest being more direct. The design is really a survey. It says the setting is the United Kingdom but (unless I have missed it) there is no clear aspect of the method or analysis that ensures this to be the case. I checked a few times but do not find confirmation in either way, though this may be my mistake. The results are not results but generalized conclusions. It would help if this actually had specific results, e.g. 'BWS results showed that...', 'general public has a stronger preference for x,y,z; researchers value a,b,c'. The conclusion should then reflect on those implications – the points started are mostly generalisations about the study.

RESPONSE:

Addressing each of the reviewer's comments in turn:

- We have set out the two aims of the study more explicitly in the objectives statement (and removed one of the uses of 'and')
- We have clarified that the design is a "survey-based BWS experiment"
- The setting is restricted to the United Kingdom, as the two populations who participated were researchers who had received grants from the UK MRC (all based at UK institutions) and a sample of the general public in the UK.
- In the 'results' section we have added more specific findings on the preferences of researchers on the public, but are unable to go into more detail due to space limitations.
- The 'conclusion' section has been refined to focus more on the potential implications of the findings about differences in valuation of research impacts and less on the methodological conclusions.

3. This is a yes but could certainly be improved in the writing by addressing point 1.

4. Yes. Perhaps too much detail and reliance on appendices, though. A methods paper might be more appropriate.

RESPONSE:

Addressed above

5. I'm concerned about the ethics – it says they followed guidelines but nothing about actual review. Could this be clarified? I realise this is not a high-risk study so I wouldn't necessarily make any major recommendations, but following guidelines is different from seeking approval. I raise this largely because the statement made about ethics in the manuscript about not requiring review from some institutions, but those institutions being clearly listed in the survey introduction that participants would see, and that the link provided is for King's College, London, who did not review it. Again, this is not a risky study, so perhaps just some clarification.

RESPONSE:

RAND Europe's internal ethics panel confirmed that risk to participants was minimal and ethics approval was not required (although in all studies involving personal data RAND Europe complies

with the Data Protection Act 1998). The grant was transferred to King's College London partway through the study. King's confirmed the previous decision that formal ethical review was not required, given that the data was being collected by an external body.

We have amended the text to read:

"RAND Europe, as the initial grant holding institution, took responsibility for assuring ethical conduct throughout the research and its internal ethics panel confirmed that formal approval was not required. On transfer of the grant, King's College London confirmed that ethical approval was not required as data collection was being undertaken by an external body."

6. This is actually the biggest concern of the current version – while the importance of the topic and the overall methodological quality seem to be very clear, the writing on actual findings is very complex, indirect, and hard to extract the meaning. Also, it feels very unclear where the method ends and the results begin – I realise there are multiple studies combined into a single paper, but it should be much more obvious where things are as it leads to the previous point. Hence several points here might better belong in the question about design.

RESPONSE:

We hope that our restructuring of the methods section has helped clarify how the various stages complemented one another and fed into the design of the final survey instrument. We have also aimed to reduce the complexity of the findings section and explain the results in less-technical language. Specific points are addressed below.

a. First, for the focus groups, the 'key observations' are not explicit on what was a discussion point generally and what was the actual finding. These are interesting remarks to read, but it would be very helpful to distinguish between what was asked of the group, what was observed, and then what has been extracted.

RESPONSE:

We have added the following details in the main text on what participants were specifically asked to do in the focus groups:

"After an opening discussion about how to define biomedical and health research, the majority of the time in the groups was used to discuss types of impact, initially by asking participants to suggest ideas, then prompting discussion on any from our draft impact framework that had not been mentioned. For each item discussed, we aimed to determine if people considered it to be a possible impact of research, how it might be categorised (e.g. health, economic, scientific) and how it might be measured."

We have also made it clearer how each observation (in box 1) affected the development of the final survey instrument.

b. The same goes for the interviews – a more engaging structure for presenting the three elements mentioned for the FGs would be really beneficial to the article.

RESPONSE:

We are conscious of not adding too much detail when the article is already long and quite methods-heavy, but have added in the following text describing the areas interviewees were asked about. "Interviewees were asked about their understanding of research impact, the kinds of impact that research in their field might have, and how these impacts might be categorised. Any items from our provisional framework not mentioned unprompted by the interviewee were asked about specifically."

We have also highlighted that the interviews contributed to the survey development by revealing that

some items in the draft framework were not considered to be impacts (and so these were removed from the framework).

c. Could the authors also state how they define 'cognitive interview'? Was this a validation check of the measure or something further? I think this also is not an appropriate place to refer to a supplement, but rather to provide the most relevant insights. However, this is one of several areas where the paper learns more toward metrics and methods rather than scientific outputs.

RESPONSE:

The following text has been added to clarify what we mean by cognitive interviews in this context: "Cognitive interviews are a structured, systematic interview technique used to understand the cognitive processes respondents use when interpreting and responding to questions."

We have also added in the most important findings from the Supplementary File. The text now reads as follows:

"The cognitive interviews proved valuable in refining the structure and wording of the survey instrument, in particular confirming that statements should be short, use varied wording to highlight differences between levels and be ordered randomly in each task, and that the maximum number of statements that it was manageable to consider in one task was eight."

d. For the pilot, it is hard to see how piloting a survey can lead to a conclusion that the questionnaires were generally well understood. Was there a feedback mechanism and/or how did this differ from cognitive interviews? I see it is mentioned later but it would help if it could be clarified (or point me to the explanation if I have missed it).

RESPONSE:

The cognitive interviews did allow us to remove ambiguities and clarify terms which might be misunderstood, but the pilot also contained specific questions to test understanding in a larger sample. These questions were also included in the final survey (Q5-7 in the general population questionnaire and Q10-12 in the researcher questionnaire). A final free text box in the surveys allowed participants to make further comments, and the content of this was carefully reviewed to identify any issues with the questions, particularly during the pilot.

We have amended the main text of the paper to read:

"Specific questions asked about the respondent's understanding of the questionnaire and these confirmed that the questions were generally well understood by both groups and that respondents were engaging with the BWS experiment."

e. Given that this paper is aiming to present some hugely relevant findings, it seems the information about how the method/measures were finalised should really be published as a paper in order to establish validity of this work. There are really far too many supplemental sections being relied on – hence they probably deserve equivalent attention or parts of this manuscript broken apart as separate papers.

RESPONSE:

We have elected not to publish a separate methods paper for the reasons set out above. However, we have combined and amended some supplementary files so that they now better reflect the various steps of developing, testing and implementing the survey. For example, the three previous supplementary files on the literature review, researcher interviews and public focus groups have been combined into one file on identifying and categorising impacts. These files align with the methods sections presented in the main text to clarify the purpose and key findings of each stage.

f. Were any validation checks done of the crowdsourced information? Having used these services myself, this seems highly critical. Is the population representative of the UK (ignoring the earlier point about location for now)?

RESPONSE:

Yes, the validation checks (chi-square tests) are reported in Supplementary File 5 – please see the corresponding p-values. The population is representative of gender and location, and significantly different in terms of age and social grade, as reported in the text. That said, all survey recruitment methods can have potential biases. We used an internet based panel as a cost-effective and faster way of collecting data. It also allowed us to present the questionnaire dynamically.

g. While I do not disagree with the approach to excluding surveys on quality checks, is there any way to justify these criteria? They seem fine in practice, but do appear a little bit arbitrary (apart from the last one – no explanation needed).

RESPONSE:

We are excluding some respondents to reduce the inclusion of responses that appear to be insincere from the internet panel of the general public. The market research agency has its own checks (e.g. based on total response time, pattern of response, etc.) to filter out such respondents. Additionally, we exclude respondents who explicitly stated that they may not have read the BWS choice questions and those who said that they did not understand the choice questions or were unable to make the comparisons.

We have amended the text (which is now in Supplementary File 5) to read:

“The robustness of the output from choice models can be improved by excluding data that are not valid (e.g. Hess et al., 2010). We conducted additional quality checks on the BWS data using three exclusion criteria.”

We have added a reference to the following article, which discusses issues around data quality and exclusion of responses in choice modelling:

S. Hess, J.M. Rose, J. Polak Non-trading, lexicographic and inconsistent behaviour in stated choice data *Transportation Research Part D*, 15 (2010), pp. 405–417

h. Is there precedent for comparing two different model approaches? It seems odd that there was no difference in the model fit and this resulted in choosing one biased and one naïve sample. It would seem that using both excluded or both inclusive is the preferred option if any two were equivalent, and you would only want to use one of each if the data forced it. However, I am happy to be convinced otherwise.

RESPONSE:

We feel that reviewer has slightly confused the criteria for screening participants/observations for inclusion in modelling with the model approach. We are effectively referring to the estimation of a model with all the available observations and the identical model structure (and specification) using a reduced sample according the sample-selection/exclusion criteria specified (lexicographic choices, etc.). In effect the same model approach is applied twice; once with a full sample and another with a reduced sample. Reporting that there is no improvement (or simply change) in model fit means that there is effectively no bias between the two models. In other words, the complete- and reduced-sample models are not (statistically) significantly different from each other. The decision to use one sample over the other is purely empirical and depends on a case by case basis independently from (in this case) the general population sample. In this study there are two independently estimated models and decision to exclude/include observations are sample specific.

i. The entire modelling section could use a human touch. It is really unclear how someone's response on a survey can be converted into coefficient related to life expectancy, and the examples given do not have clearly explained values or origins. Why are results not simply expressed as asked in the survey? As I had to regularly bounce back and forth between the supplement pages and the section, I may have missed it, but the best interpretation I could make is that certain answers come in with known weights based on known evidence about those items' real-world impact. Could this please be addressed?

RESPONSE:

- Firstly, we would like to emphasise that there are no pre-determined weights based on known evidence about those items' real-world impact.
- In effect, respondents' choices for most important (best), least important (worst), second most important (second least important) elicited in each BWS task in the survey are assumed to be stochastic (non deterministic).
- Thus for each attribute level (research impact) the probability of being selected as best (most important), worst (least important), etc. is associated with the research impact presented to respondents.
- A model coefficient (that in this case comes from a multinomial logit model) explains the level of association that is the 'influence of the research impact' on the choice made (as most important, least important, etc.). This coefficient is estimated relative to a research impact level used as a reference. All model coefficients are estimated on a common scale and can be ordered to obtain the 'importance ranking' of all research-impact levels presented in the experiment.
- In a given sample (general population or researchers), the ratio of any research-impact coefficient over a quality of life coefficient (additional year of life expectancy, as determined from the BWS task) allows the former to be expressed as 'additional years of life expectancy'.

We have clarified this by amending the main text as follows:

- A sentence has been added to clarify that the conversion is based on each group's preference for impacts in the life expectancy domain of the survey:

“To allow comparison between the preferences of researchers and the general public, we converted the coefficients of both groups to a common scale using a tangible unit – ‘additional years of life expectancy’, based on each group’s preferences for the life expectancy domain in the BWS task.”

- The detail of the calculations has been moved to Supplementary File 5, given its technical nature and since the additional methodological details provided in that file should aid its interpretation.

j. At least in the version I could see for the review, Figure 2 does not have sufficient resolution to determine the difference between shaded and not shaded. The pattern is well-presented overall, though.

RESPONSE:

We think that the version submitted should be high enough resolution, but will confirm with the editor and would be very happy to send a new version if necessary.

k. Box 2 is a nice way to present findings, but I would strongly request that this be reduced to a small number, and then present the rest in text. Or come up with another way to present this. I do like consolidating key messages, but this essentially becomes almost a text within the text.

RESPONSE:

We understand the reviewer's concern that the findings box was lengthy, perhaps making it difficult for the reader to understand the main messages. We have addressed this in two ways: (i) findings are

now summarised in the text as areas of agreement and disagreement between researchers and the public; (ii) we have reordered the statements in the text box so that they are now grouped according to agreement and disagreement between the two groups, and so that those considered most important are at the start of the list (including those highlighted by the reviewer). We still include all of our main findings in the text box, however, as we feel that it is important to highlight both the similarities and differences in valuations, as well as cover the wide range of impacts that can arise from biomedical research.

l. Some interpretations of the findings as 'more' and 'less' seem a bit inaccurate. As this is choice modelling study, should it not be discussed as a matter of 'greater proportion' and 'lower proportion'? In other words, are certain things valued more, or simply do more people value it from one group? Please clarify this if I am mistaken. I was trying to check if the statistical technique might clarify this distinction but I was not 100% sure in the end, so felt necessary to raise.

RESPONSE:

The model estimation results effectively correspond to the 'influence' (statistical association) of a research-impact being selected as most important, least important, etc. As a result, when estimating the probability of a research impact being selected as 'most important' the corresponding coefficient of that research impact is expected to be positive. Similarly, for another research impact, its coefficient is expected to be positive when estimating its probability of being selected as 'most important'. Because the two coefficients are estimated (within the same sample) on a common scale it is possible to compare the two coefficients and thus determine that one research impact is valued more (or less) than the other; hence the terms 'more' and 'less' in the text correspond to individuals' (average) valuation of a research impact – i.e, 'certain things are valued more by people'. Having said that, it is also true that there may be groups within the sample that place a higher (or lower) value on a research impact than other groups or the rest of the sample.

Our clarification in the text of the conversion of the model coefficients to a common scale should also help interpretation of our use of 'more' and 'less' in the text.

m. The same point about equal value – should it not just be 'similar proportions'?

RESPONSE:

Addressed in previous point.

n. It seems that key observation 7 belongs in a much more prominent place as this is clearly a highly relevant finding.

RESPONSE:

We agree – and indeed this is an area which survey respondents also deemed to be among the most important. We have moved it to the start of the list.

o. KO #8 is an enjoyable finding. I look forward to the media picking up on that...and then using it in my public lectures.

p. KO #9 is highly relevant, please also highlight this.

RESPONSE:

This finding has also been moved further up and grouped with the other statements indicating agreement between researchers and the public.

7. I would actually request a slightly more lay explanation of the analyses. It is passed over rather

briefly in the method and a few general statements on the process would really help with grasping the work done. I think it's a bit risky to rely on the supplement, unless there was a more clearly described link to a published method applied precisely here.

RESPONSE:

We have added in the three references below, which set out the BWS method as used in our study. Since this is an established technique, we have not provided a detailed account of its implementation in this paper (and given the reviewer's concerns about the length of the methods section).

- Louviere, J.J., Marley, A.A.J., & Flynn, T. (2015) *Best-Worst Scaling: Theory, Methods and Applications*, Cambridge University Press.
- Marley, A.A.J., Flynn, T.N. and Louviere, J.J. (2008) 'Probabilistic models of set-dependent and attribute-level best–worst choice', *Journal of Mathematical Psychology*, 52(5), 281-296.
- Marley, A.A.J. and Pihlens, D. (2012) 'Models of best–worst choice and ranking among multiattribute options (profiles)', *Journal of Mathematical Psychology*, 56(1), 24-34.

8. No major concerns with referencing, though it might not hurt if a bit more scientific theory was introduced into the introduction and/or reflected on in the discussion.

RESPONSE:

We are not entirely clear if the reviewer is referring to scientific theory on research impact or on the use of choice modelling in this kind of exercise.

If the former, we have referenced a number of studies exploring the variety of impacts that can come from biomedical and health research, but do not feel that it is central to the current study to go into these in detail. In part, this is because the literature review conducted as part of the study covered the range of impacts that might occur and various frameworks for categorising them, but it is also because our focus is less on the impacts that might occur and their measurement, and more on the value that different stakeholder groups place on them. To the best of our knowledge, this has only been studied specifically by Miller et al. (2013), a study which we describe in the introduction and refer to again in the discussion.

If the reviewer is referring to the choice modelling literature, we have now added several references relating to the BWS methodology employed in our study. Again though, we do not feel that it is appropriate to include extensive detail on the development of choice modelling techniques, since we are not aiming this paper at a particularly technical audience and such accounts are available elsewhere in the methodological literature. Our focus in this paper is on the transferability of the BWS technique to a novel field.

9. Skip – all in point 6.

10. If above points are addressed, then this should be fine.

11. The discussion is fine but personally I would strongly suggest reflecting more on the 9 key observations and critiques therein, and far less on the method. The findings are really important and at least three of them are worth a page of reflection, in my view.

RESPONSE:

We have considered the reviewer's point and agree that the study raises intriguing questions about the preferences for different types of impact. However, given that this study is a first step in applying a new methodology in this field to a question which has received little attention to date, we want to avoid over-interpreting our results, particularly since (as we say in the paper) we cannot draw

conclusions on how preferences are formed or why they differ. Indeed, our intention with this paper is that it serves as a starting point for discussion of an important and under-researched issue, and stimulates further research in this area.

We have, however, expanded on how our findings might be taken forward, adding the following text: “What we are unable to assess from our data are the reasons behind the differences we observe between the groups. Encouraging discussions around why preferences differ in specific instances might usefully inform future research objectives, as well as encourage more nuanced communication between researchers and the public on the potential benefits of research.”

12. I think sample-only limitations is not particularly self-critical but as I was not party to the study itself, cannot say what else might go in there.

RESPONSE:

We have added the limitation mentioned above. The new text reads as follows:

“What we are unable to assess from our data are the reasons behind the differences we observe between the groups. Encouraging discussions around why preferences differ in specific instances might usefully inform future research objectives, as well as encourage more nuanced communication between researchers and the public on the potential benefits of research.”

Reviewer: 2

R2: This paper addresses the important, and poorly defined, topic of research impact. It is clearly written, well conducted and represents a useful addition to the field.

I have no major comments but suggest the introduction may benefit from a brief outline of some of the key ways in which research might make an impact.

RESPONSE:

We are wary of the length of the paper, but have added the following (in addition to the references, which specifically discuss the diverse range of impacts that were revealed in the REF case studies): “There is strong evidence that research makes a significant contribution to society [8-11], and that contribution manifests itself in different ways [12, 13]. These benefits might occur within the research system itself or more widely in areas such as healthcare, the environment, technology, the economy or culture.”

The literature review which informed the impacts and categorisation used in our survey also provides an overview of different areas in which research impacts can occur.

R2: While it is briefly alluded to, I think it is also worth noting more pointedly that another major piece of work missing in this area is the development of robust methods for assessing whether or not an impact has occurred.

RESPONSE:

We have added this in the discussion, which now reads:

“With further research, it may be possible to use such valuations to develop metrics for assessing research impact, although we stress that this is a longer term objective, to be considered alongside the need for better ways of identifying and measuring societal impact more generally, and should not be advocated based on the current study.”

VERSION 2 – REVIEW

REVIEWER	Kai Ruggeri University of Cambridge, UK
REVIEW RETURNED	29-Mar-2016

GENERAL COMMENTS	Thank you for your very thorough response. I would like to give a full review again but there is an issue that needs to be clarified generally. Is it possible to get simply the precise measurement tool for this study? It appears there is clear disagreement on whether or not this approach qualifies as a DCE and as it is very difficult to search through the appendix, I would appreciate if a single document with only the metric (and not extensive instructions/algorithms) was provided.
-------------------------	--

VERSION 2 – AUTHOR RESPONSE

Reviewer 1:

Thank you for your very thorough response. I would like to give a full review again but there is an issue that needs to be clarified generally. Is it possible to get simply the precise measurement tool for this study? It appears there is clear disagreement on whether or not this approach qualifies as a DCE and as it is very difficult to search through the appendix, I would appreciate if a single document with only the metric (and not extensive instructions/algorithms) was provided.

Response:

We would like to thank the reviewer for taking the time to look at the revised version of the manuscript. To clarify exactly what the reviewer needed in order to complete his review and to avoid unnecessary iterations, we obtained permission from the editor to contact the reviewer directly. To address his concern we:

- Provided the complete versions of the surveys as seen by respondents (set out in Supplementary File 4)
- Confirmed that the data used to model respondents' choices was obtained from the tasks presented in Part 2 of the survey questionnaire (between questions 4 and 5 in the general population survey and between questions 9 and 10 in the researcher survey).

The reviewer then confirmed that this information was sufficient for him to complete his review, which we look forward to receiving in due course. For this reason, the paper itself has not been revised in this submission.

VERSION 3 - REVIEW

REVIEWER	Kai Ruggeri University of Cambridge, UK
REVIEW RETURNED	17-May-2016

GENERAL COMMENTS	General All in all, I find this a very, very strong revision. The emphasis on the paper itself is much better and the appendices act merely as an augment, which is great. My comments are entirely because I believe this is a very important piece of research and I feel these points will make it engage the widest of audiences. Abstract Results – are those really results? Shouldn't this be what was
-------------------------	--

	<p>learned, not just a summary of what was done? Maybe just cut the first part, unless you are saying this is validation. Also, 'different things in different ways' is a running theme in the paper. Perhaps reduce some of the vague nuance and simply state a stronger view (if in it is one of many that could be listed)?</p> <p>Conclusion – suggest cutting the restatement about 'different things in different ways' and simply highlight the key finding in the second half. This is really the most interesting part! Also suggest dropping the future research part and stick to a concluding message from your work as there is a large amount of information provided.</p> <p>Strengths & Limitations</p> <p>These seem to generally restate what has been said rather than saying what is strong/weak. To me, strong/weak is clear: Strong: valuable insight about who values what, which guides both research funding decisions and public messages about research Weak: Some conclusions should not be overstated as mechanism for eliciting is somewhat complex (though sufficiently direct to the point).</p> <p>Introduction</p> <p>Too much use of the word 'different' without going into any detail. An example: There is strong evidence that research makes a significant contribution to society [8-11], and that contribution manifests itself in different ways [12, 13].</p> <p>An example of the evidence? An example of manifestations? Something a bit more concrete would be helpful because ultimately this is what the study should link back to. Just the notion of 'different and different' doesn't really inspire much tangibility.</p> <p>Introduction is generally very short and lacks much theoretical context. Not saying it needs to be massively extended, but presumably a more concrete statement of the theory being tested would be helpful.</p> <p>Intro jumps pretty quickly to over-viewing the study rather than introducing it.</p> <p>Summary: there is a lot reliance on the word 'different' – would suggest doing a quick scan of the places where it is used and incorporate something a bit more precise. This would really add a lot to the full paper, particularly the theoretical backing. (Again, not saying massive changes, but I think this really ups the overall value of the paper given the importance of the work)</p> <p>Method</p> <p>Much improved from the first submission. Definitely will engage a wider and relevant audience as it now reads.</p> <p>Suggest pulling the lit review section and moving relevant topics into the introduction. This would satisfy concerns from introduction section comments. If there are word count issues, I would strongly suggest moving some detail about anything other than the survey itself to appendix, or simply reducing to a few lines.</p> <p>Not mandatory, but could some sort of flow diagram be included to</p>
--	--

	<p>cover the different elements of the study that involved generation of new data? I found myself trying to check what role each step had in determining future steps vs concluding all steps were discrete. Nothing fancy but I think this helps for the future research suggestions, particularly by including the numbers of individuals involved at each round to get an overall concept/context. Not mandatory, though.</p> <p>Would strongly suggest more detail about the survey completion process. It seems less described than the other steps used and is perhaps most vital to final interpretations.</p> <p>While I am now satisfied with the feedback I received about the method being an appropriate DCE, there is no mention of analysis plans in the methods section. This would seem important and relevant to what comes next, particularly including key assumptions (or failed ones) related to the data generated/analysed. Some sections in the results about various decisions in the process of analysis would also likely be better placed here.</p> <p>Results</p> <p>The observation box is great. I suppose the only thing I still feel to be lacking is clarity on what is the actual key finding(s) in text/analysis. Something that directly states X was tested and Y was found, but beyond the reliance on 'different'.</p> <p>Discussion</p> <p>One general concern: validation of BWS is stated a few times, but is this actually established? It keeps saying it is established but against what criteria was this done? I am not implying it is not, I just do not see clearly where some threshold has been surpassed or other expectation satisfied.</p> <p>Similar to the results concern: though stated generally, the key finding should be explored a bit more. It's highly relevant and deserves some unpacking! Fine to generalise as the open (again a 'different and different'), but I think any reader will want the key difference findings stated/discussed/critiqued. Not asking for a massive extension, but definitely worthy of some discussion.</p> <p>Conclusion</p> <p>I think you have undersold the importance of this work. Perhaps after going through the above points about the key findings/discussion points, you might want to conclude on something a bit stronger? I feel this work is very important and something that any reader could take away would be a strong way to close ('future research' I would advise as a small subsection prior to a concluding statement.</p>
--	---

VERSION 3 – AUTHOR RESPONSE

REVIEWER 1: All in all, I find this a very, very strong revision. The emphasis on the paper itself is much better and the appendices act merely as an augment, which is great. My comments are entirely because I believe this is a very important piece of research and I feel these points will make it engage the widest of audiences.

RESPONSE: We would like to thank the reviewer for taking the time to review the paper and for providing very helpful and thoughtful comments.

REVIEWER 1: Abstract

Results – are those really results? Shouldn't this be what was learned, not just a summary of what was done? Maybe just cut the first part, unless you are saying this is validation. Also, 'different things in different ways' is a running theme in the paper. Perhaps reduce some of the vague nuance and simply state a stronger view (if in it is one of many that could be listed)?

Conclusion – suggest cutting the restatement about 'different things in different ways' and simply highlight the key finding in the second half. This is really the most interesting part! Also suggest dropping the future research part and stick to a concluding message from your work as there is a large amount of information provided.

RESPONSE: We have amended the abstract along the lines suggested, by rewording the results section, although we have left point (1) in so that this section mirrors the structure of the objectives section (which we think aids clarity). Point (2) in the results now reads:

“The general public and researchers provided similar valuations for research impacts such as improved life expectancy, job creation and reduced health costs, but there was less agreement between the groups on other impacts, including commercial capacity development, training and dissemination.”

While we understand the reviewer's preference that the conclusion focuses on the substantive findings of the study, we feel that it is important for context to flag that this is a research area in its infancy and that our findings are only a first step. For this reason we have left the conclusion section unchanged.

REVIEWER 1: Strengths & Limitations

These seem to generally restate what has been said rather than saying what is strong/weak. To me, strong/weak is clear:

Strong: valuable insight about who values what, which guides both research funding decisions and public messages about research

Weak: Some conclusions should not be overstated as mechanism for eliciting is somewhat complex (though sufficiently direct to the point).

RESPONSE: We have partially revised the the strengths and weaknesses to pick up this comment. A penultimate bullet point has been added which reads:

“The conclusions should not be over interpreted given the methodological nature of the research, including the complex mechanisms for eliciting valuations.”

While we are conscious that some of the remaining points restate the study's approach, we do consider one of the major strengths of the work to be the introduction of novel methods to the valuation of research impact, a field in which existing methods are somewhat limited. We would be happy to follow the Editor's advice, though, on whether this is appropriate to the style of the journal.

REVIEWER 1: Introduction

Too much use of the word 'different' without going into any detail. An example: There is strong evidence that research makes a significant contribution to society [8-11], and that contribution

manifests itself in different ways [12, 13].

An example of the evidence? An example of manifestations? Something a bit more concrete would be helpful because ultimately this is what the study should link back to. Just the notion of 'different and different' doesn't really inspire much tangibility.

RESPONSE: We have added examples of impacts and reworded the paragraph slightly (particularly, instances of "different") to address this point. The revised text now reads:

"There is strong evidence that research makes a significant contribution to society [8-11], and that contribution manifests itself in different ways [12, 13]. For example we know that the total economic return from biomedical and health research is between 24 and 28% [9, 11], and from analysis of the REF impact case studies that there are a wide variety of impact topics [13]. These benefits might occur within the research system itself or more widely in areas such as healthcare, the environment, technology, the economy or culture. However there remains a lack of methodological and empirical research on how the public values research impact [14, 15] and how valuations may vary between stakeholder groups [16]."

REVIEWER 1: Introduction is generally very short and lacks much theoretical context. Not saying it needs to be massively extended, but presumably a more concrete statement of the theory being tested would be helpful.

Intro jumps pretty quickly to over-viewing the study rather than introducing it.

Summary: there is a lot reliance on the word 'different' – would suggest doing a quick scan of the places where it is used and incorporate something a bit more precise. This would really add a lot to the full paper, particularly the theoretical backing. (Again, not saying massive changes, but I think this really ups the overall value of the paper given the importance of the work)

RESPONSE: We have moved the description of BWS from the methods section to the introduction to provide some theoretical context on the approach used by the study (see second paragraph of page 5).

REVIEWER 1: Method

Much improved from the first submission. Definitely will engage a wider and relevant audience as it now reads.

Suggest pulling the lit review section and moving relevant topics into the introduction. This would satisfy concerns from introduction section comments. If there are word count issues, I would strongly suggest moving some detail about anything other than the survey itself to appendix, or simply reducing to a few lines.

RESPONSE: We carefully considered this suggestion but have decided not to implement it. As noted above we moved the description of the BWS approach to the introduction, but we think that retaining a section titled 'literature review' in the methods is important, as this was one of the ways that we developed the attributes used in the BWS experiment.

REVIEWER 1: Not mandatory, but could some sort of flow diagram be included to cover the different elements of the study that involved generation of new data? I found myself trying to check what role each step had in determining future steps vs concluding all steps were discrete. Nothing fancy but I think this helps for the future research suggestions, particularly by including the numbers of

individuals involved at each round to get an overall concept/context. Not mandatory, though.

RESPONSE: Thank you for this suggestion. We have added a flow diagram (Figure 1) at the beginning of the methods section.

REVIEWER 1: Would strongly suggest more detail about the survey completion process. It seems less described than the other steps used and is perhaps most vital to final interpretations.

RESPONSE: We are not certain that we fully understand what the reviewer means by the completion process, but have added details as set out below to clarify how a respondent completed the questionnaire (i.e. order of the questions) and how long it took to complete. We hope that these address the reviewer's concerns.

Text changes:

We have added the following sentences in the Results > Data Summary section.

General population (first paragraph of Data Summary):

"Just over half (52.5%) of the respondents who completed the survey took more than 15 minutes to do so, while 12.5% completed it in 10 minutes or less."

Researchers (second paragraph of Data Summary):

"The majority (67%) of the respondents who completed the survey took more than 15 minutes, while 7.6% completed it in 10 minutes or less."

We have also clarified the steps taken by respondents in the section 'Constructing and testing the survey instrument'. This now reads as follows:

"For the general public the survey questionnaire included: screening questions, an introduction to the BWS experiment and its tasks, questions relating to attitudes to science, and socio-demographic questions. Screening questions asked respondents about their age, gender, region of residence, social grade and work status. For researchers the survey questionnaire included: questions on research background, job title, clinical experience (if any), an introduction to the BWS experiment and its tasks, and socio-demographic questions. Respondents completed the survey by answering the questions in the same order. The survey could be saved and completed in multiple sessions. The researcher questionnaire was shorter than that for the general public."

REVIEWER 1: While I am now satisfied with the feedback I received about the method being an appropriate DCE, there is no mention of analysis plans in the methods section. This would seem important and relevant to what comes next, particularly including key assumptions (or failed ones) related to the data generated/analysed. Some sections in the results about various decisions in the process of analysis would also likely be better placed here.

RESPONSE: To address this we have added a new subsection entitled 'Data analysis' at the end of the methods section. For clarity, we have also moved some description from the 'BWS method' description to this new subsection. The subsection reads as follows:

"The data analysis in this study comprised two stages: (1) descriptive analysis and (2) modelling of the BWS data.

The aim of the descriptive analysis was to summarise the profiles of the participants in the general population and biomedical and health researcher samples by socio-demographic and other characteristics. We also conducted quality checks on the BWS data using three exclusion criteria as

detailed in Supplementary File 5. The remaining data in both samples were then tested for representativeness against various socio-demographic characteristics (see Supplementary File 5 for further details). Finally, we constructed segments within the researcher and general population samples defined by research activity codes and their attitudes to science respectively (see Supplementary File 5).

In the second stage of the analysis, modelling of the BWS data was conducted at the respondent level using discrete choice analysis [26]. The aim of the modelling was to derive weights reflecting the relative importance of the research impact levels for different stakeholder groups. The probability of an individual respondent choosing a research impact level as the 'most important' ('best') among a set of research impacts (attribute levels) can be modelled within a multinomial logit (MNL) framework as described in Supplementary File 1. The estimated coefficients (weights) of each research impact can then be expressed on a common scale allowing one to infer how respondents or different groups of respondents value different types of research impact. In this stage, we also examined how individuals' preferences varied according to their attitudes to science (in the general population sample) and research activity codes (in the researcher sample)."

REVIEWER 1: Results

The observation box is great. I suppose the only thing I still feel to be lacking is clarity on what is the actual key finding(s) in text/analysis. Something that directly states X was tested and Y was found, but beyond the reliance on 'different'.

RESPONSE: We have added the following text in the results section to help the reader interpret Figure 3 (Fig 2 in the previous version of the manuscript).

"For example, the first horizontal bar in Figure 3 relates to the impact statement "Research replicates the work of others, helping to strengthen the evidence of how some things work". This statement is valued at providing the equivalent of 11.16 [95% C.I.: 9.77, 12.55] additional years of life expectancy (AYLE) by the general public and 10.36 [95% C.I.: 10.12, 10.60] AYLE by researchers. The transparent bars in the figure indicate that there is no statistically significant difference between the two groups with respect to this statement. By contrast, the fourth impact statement in Figure 3 – "Research contributes to better care being provided at the same cost" – is valued more by the general public, the solid bars indicating that this difference is statistically different."

REVIEWER 1: Discussion

One general concern: validation of BWS is stated a few times, but is this actually established? It keeps saying it is established but against what criteria was this done? I am not implying it is not, I just do not see clearly where some threshold has been surpassed or other expectation satisfied.

RESPONSE: We agree that it would be inappropriate to claim 'validation' as such, since we don't have a formal test for this. However, we don't think there is any such specific language in the manuscript as it stands. We have edited text in the discussion which suggested that we have 'demonstrated' the methodology, as part of wider revisions in this section to make it more appropriate, and hope that this addresses the reviewer's concerns.

REVIEWER 1: Similar to the results concern: though stated generally, the key finding should be explored a bit more. It's highly relevant and deserves some unpacking! Fine to generalise as the open (again a 'different and different'), but I think any reader will want the key difference findings stated/discussed/critiqued. Not asking for a massive extension, but definitely worthy of some discussion.

RESPONSE: Given the above amendments on this issue we have not made a change here – partly as we think it is appropriate for the discussion section to be ‘high level’, as in its current form. We have, however, made minor edits to wording and structure to more clearly set out the study’s contributions. We have also added a new first paragraph to summarise, which reads as follows:

"We can identify three key findings from this study. First, that it is possible for different types of impacts to be directly compared and rated, and that BWS offers a potentially effective way to make such comparisons. Second, that there are similarities in views between the researchers and the public about the relative importance of social impacts, but also notable differences of opinion between these groups regarding other research-related impacts. These differences are important, as researchers are increasingly asked to make judgements about the value, or potential value, of research in the award of public funding. Finally, we note that our findings differ from those of a previous study by Miller et al [16], suggesting that further research is required. We explore each of these in turn."

REVIEWER 1: Conclusion

I think you have undersold the importance of this work. Perhaps after going through the above points about the key findings/discussion points, you might want to conclude on something a bit stronger? I feel this work is very important and something that any reader could take away would be a strong way to close (‘future research’ I would advise as a small subsection prior to a concluding statement.

RESPONSE: Thank you for this suggestion. We have restructured and strengthened the narrative in the discussion section, breaking out the discussion of future research into a separate paragraph after the summary and discussion of the key findings. We have then added a strong concluding statement to finish the paper, drawing out our key findings and their importance. This new concluding paragraph reads as follows:

“To summarise, this work sets out a new approach to elicit opinions about the relative importance of different types of research impact and highlights evidence for some important differences in opinion between researchers and members of the public. This has implications for policy making, since researchers and funders commonly assess the potential and realised impact of research as part of funding decision-making processes. The methods set out here might offer one way to understand and begin to address this, with the potential, through further research, to even develop a way to assess and compare different types of impact based on empirical evidence of their relative importance to members of the public. Exploring this question and these methods further could help better align publicly funded research with the needs and priorities of the public, strengthening accountability and public engagement with science, and perhaps, ultimately, offering better value to society.”

VERSION 4 – REVIEW

REVIEWER	Kai Ruggeri University of Cambridge, UK
REVIEW RETURNED	08-Jul-2016

GENERAL COMMENTS	Overall, I think this final version does much better service to the insights generated. I am very pleased the authors have responded to the many comments and think this article will now have greater visibility and impact. I would anticipate some critique on the methodological description and analytical choices, but these would be for specialist academics and likely boost the chance of further work on the topic.
-------------------------	--

	<p>I would maybe offer one suggestion: the authors may want to ask a non-expert to read through this to check for general style/language. It's massively improved but still, given the content, I would imagine we want the largest possible audience to see this work.</p> <p>Thank you again for the opportunity to comment on this manuscript and I look forward to seeing it in print.</p>
--	--