Author's response to reviews

Title: Using the theory of planned behaviour as a process evaluation tool in randomised trials of knowledge translation strategies: a case study from UK general practice

Authors:

Craig R Ramsay (c.r.ramsay@abdn.ac.uk)
Ruth E Thomas (r.e.thomas@abdn.ac.uk)
Bernard L Croal (bernie.croal@nhs.net)
Jeremy M Grimshaw (jgrimshaw@ohri.ca)
Martin P Eccles (Martin.Eccles@newcastle.ac.uk)

Version: 3 Date: 28 June 2010

Author's response to reviews: see over
<table>
<thead>
<tr>
<th>Reviewer 1 - Pierre Durieux</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is a superb article of outstanding interest. To my knowledge, it is the first time that a study demonstrates that a theory based process evaluation can provide useful information on mechanisms that aid the interpretation of the results of a trial. I have no request for revision, but just two points of clarification.</td>
<td>We thank the reviewer for his kind comments. We agree with the reviewer that this is likely the first time that a formal meditational analysis has been performed on a KT trial result.</td>
</tr>
<tr>
<td>1) The statistical analysis is sometimes difficult to follow for a non specialist, but I do not know how this problem could be improved. As an example, I have difficulties to understand the mediation analysis which is presented table 5 but this may be due to my poor knowledge of the statistical methods employed. However, other readers may encounter the same problem.</td>
<td>We have addressed this issue by supplying an additional figure 2 that illustrates diagrammatically how a trial effect is mediated (Pg 10 and pg 21).</td>
</tr>
<tr>
<td>2) Is it possible to present the full copy of the questionnaire on the internet site of the review?</td>
<td>We have added the questionnaire as additional file 3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reviewer 2 – Alison Brown</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strengths of this paper include its innovative approach to process evaluation, its use of theory, and its complex statistical analysis. I commend the authors for tackling the challenge of applying a theoretical construct to the design and interpretation of their evaluation and I believe there is much to be gained scientifically from the continued pursuit of this application in implementation science as a field. However, the paper has some components that I feel could be confusing to readers, and some components that overstate the implications of their design and findings. Overall the authors need to take a more cautious and critical approach to the presentation of the results and their interpretation.</td>
<td>We thank the reviewer for their positive comments and address each concern below.</td>
</tr>
<tr>
<td>1. The concept that process evaluation data (especially post-intervention data only) could be utilized to identify “causal mechanisms” needs to be more fully explicited in order to be convincing. In this study, as the authors note, behavioural data was not available for the individuals who completed the questionnaire, but rather, was analyzed at the level of each general practice. A design that includes pre- AND post-intervention questionnaires AND individual-level behavioral data would be able to better address the issue of causal mechanisms, but such a design would have to be triangulated with a number of other data sources and</td>
<td>We agree with the reviewer that to fully identify causal mechanisms a more thorough process evaluative model would have to be employed. Randomisation enables us to be more confident in making causal inferences, however, we agree that the causal models are still observational analyses and as such require cautious interpretation. We have edited paragraphs 6 and 8 of the discussion section (pages 15, 16 &amp; 17) accordingly.</td>
</tr>
<tr>
<td>Edits made :</td>
<td>Edits made :</td>
</tr>
<tr>
<td>• pg 15 addition of sentences “Ideally to operationalise … by practice”;</td>
<td>• pg 15 addition of sentences “Ideally to operationalise … by practice”;</td>
</tr>
<tr>
<td>• Pg 16 addition of “Our findings and</td>
<td>• Pg 16 addition of “Our findings and</td>
</tr>
</tbody>
</table>
would still be limited from an empirical point of view. The authors do note that it is difficult to analyze “individual psychological theories when the behavioural outcome is collected at the practice level,” but further critical analysis of this design issue needs to be provided.

<table>
<thead>
<tr>
<th>2. The second major concern is the use of one post-intervention measure, highly obvious in its use of TBP constructs (at least as presented in Box 1), as constitutive of a process evaluation. I am concerned that this paper may mislead readers who are not familiar with process evaluations in the sense that very little is said about what would or could have made this a more thorough, comprehensive process evaluation. As currently written, a reader naïve to process evaluation might conclude that utilizing one post-intervention measure makes a process evaluation. A discussion of how a theoretically-driven questionnaire could be PART of a thorough process evaluation would likely necessitate references to the ever-growing literature on process evaluations in implementation science.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The TPB model as used in this study is the most frequently used social cognitive theory in studying professional behaviour and professional behaviour change. As identified by the systematic review conducted by Godin et al (Implementation Science 2008, 3:36).</td>
</tr>
<tr>
<td>In addition to the responses made above, we have added a paragraph (pg 17) at the end of the discussion starting “the aims of process evaluations” to put this study in to context.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Related to the second concern, the lack of any literature cited in the discussion makes it difficult for the reader to consider the authors’ findings in relation to other studies of a similar nature. A discussion of where this study departs from and/or augments the work done by other scientists is particularly necessary in light of the complex findings that the authors present.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The use of theory in health professional behaviour change studies is limited (Godin et al) and the specific use within randomised trials is extremely rare. The additional paragraph in the discussion (pg17), as mentioned above, describes this and makes reference to the findings of the relevant reviews.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. The authors should provide an explanation of how they randomly selected the 50% of the sample of the “randomized groups of general practitioners,” and they should provide some speculation as to why their response rate of 57% was rather low, and what efforts (if any) were undertaken to increase the response rate. Such information about who was selected would allow for interpretation as to how representative the data are in terms of the enrolled general practitioners and of general practitioners outside this study. Information about the response rate would also be valuable from an implementation science perspective in that participation in process evaluation endeavors is often difficult to achieve to the extent that we seek in such studies.</th>
</tr>
</thead>
<tbody>
<tr>
<td>On page 7, first paragraph of study design and population section we have added the following: “The random sampling was performed by a statistician independent of the research team.”</td>
</tr>
<tr>
<td>A response rate of 57% is a reasonable for a survey in UK general practice. It is similar to that of other postal surveys of health professionals. We have added a sentence to demonstrate this to paragraph 5 in the discussion with the reference to the relevant work. (Cook et al. BMC Health Serv Res 2009)</td>
</tr>
<tr>
<td>Data were presented in the “Survey sample” paragraph of the results section on how sampled practices requesting levels differed between sampled and non-sampled practices. There were no significant differences between the groups.</td>
</tr>
</tbody>
</table>
As regards representativeness of general practitioners outside of the study, all GPs in Grampian were in the original study so the sampled versus non-sampled figures correspond to likely differences between study and non-study GPs.

<table>
<thead>
<tr>
<th>Reviewer 3 – Jeffrey Smith</th>
<th>No response necessary.</th>
</tr>
</thead>
<tbody>
<tr>
<td>General Comments:</td>
<td></td>
</tr>
<tr>
<td>The stated purpose of the manuscript is to “illustrate the applicability of causal methods in randomised trials by undertaking a theory-based process evaluation study to explore whether the cognitions of general practitioners predicted their test requesting behaviours and secondly, whether the trial results were mediated by the theoretical constructs.” Although I agree with the authors’ premise that theory-based process evaluation can offer useful information on causal mechanisms of intervention effects and can enhance interpretation of study data, the approach utilized in this case raises concerns about the description of this work as a ‘process evaluation’ (see below).</td>
<td></td>
</tr>
<tr>
<td>Process evaluation (PE) conducted alongside a randomized trial can offer useful information on fidelity of intervention implementation (ie, whether the intervention was implemented as planned), possible mechanisms/mediators of intervention effects, and (for future studies) potential refinements or modifications to consider in intervention design/implementation to enhance its impact on targeted clinical behaviors. PE is typically conducted either concurrent with intervention implementation (so information on the above can be assessed while the intervention is ongoing) or immediately following implementation… typically involving a combination of qualitative and quantitative methods. Concerns for this study are:</td>
<td></td>
</tr>
<tr>
<td>The quantitative survey for this study was administered 12 months after the interventions had been initiated, and the data collected in the surveys were based on ‘fixed’ scenarios that may or may not have been representative of ‘real’ patients seen by providers during the course of the trial. Data apparently were not collected while the interventions were ongoing,</td>
<td></td>
</tr>
<tr>
<td>Whilst it is true that the survey was administered 12 months after the interventions had been initiated, the interventions were ongoing over that 12 months and were not “turned off” at 12 months. Within the context of the randomised controlled trial conducting the survey at 12 months enabled a comparison of the constructs between intervention and control groups. ie at 12 months, the intervention group was receiving the intervention, but the control groups were still intact control groups. This enabled a comparison of the intervention effect without introducing possible contamination of the</td>
<td></td>
</tr>
</tbody>
</table>
and it is possible that the scenarios included in the survey may not adequately represent patients presenting to providers with potentially more complex clinical issues.

Original trial interventions with a questionnaire survey.

It is possible that the scenarios may not adequately represent patients presenting to providers with potentially more complex clinical issues but the point of the scenarios was to operationalise the clinical problem as described and targeted by the brief educational messages of the intervention. We have added to a sentence in the Methods section under Data Collection (pg 8) and to pg 16 of the discussion to make this clear.

No pre-intervention survey was administered, so it is not possible to assess with any confidence whether the intervention had an effect on TBP constructs (ie, we cannot calculate pre-/post-change in survey scores). Although there are differences in the post-intervention survey scores between trial groups, we have no way of knowing whether these differences may have existed prior to the intervention. Presumably, randomization of providers to trial groups should have limited the potential for such pre-intervention differences, but there are no pre-intervention survey data available to confirm (or refute) that presumption.

We do not agree with the reviewer that we cannot be confident that the interventions had an effect on the TPB construct. The randomised trial result clearly demonstrated that all randomised groups were balanced in test requesting behaviour at baseline and statistically significant decreases in test requesting was observed after the intervention.

Whilst it may seem theoretically possible that the intentions to request a test were imbalanced at baseline and continued to have the same level of imbalance post intervention, we cannot determine any obvious mechanism for such an effect to have occurred given the balance at baseline in requesting behaviour.

The use of practice rather than individual provider test requesting behavior as the dependent variable (DV) is a major study limitation (which the authors acknowledge), but particularly in this case since my understanding of the Theory of Planned Behavior is that it focuses on individual (rather than group) behavior. Data on individual provider test requesting behavior would have been more appropriate (the authors do note that such data were not available).

We accept these reservations and have acknowledged them in the manuscript, (see additions of sentences to the discussion described previously and below) but we do not think it is a major study limitation. We have added reference to the methodology and its limitations (Eccles 2009).

Further, the authors point out another problem with the dependent variable in that it may not match the context of the scenarios included in the survey (ie, DV does not distinguish between initial and repeat tests).

The use of the practice level DV will only result in more “blunt” correlational analyses ie correlations between constructs will be smaller on aggregated data. The fact that the correlations are significant in this study suggests that they are real. We also included sensitivity analyses (on minimum and maximum) and these did not change the direction of effects.

As above, this limitation has the effect of reducing the correlations.

Additions made
The above concerns lead me to conclude that ‘process evaluation’ is not an appropriate description of this work; or if it is appropriate, represents process evaluation in its most limited form.

See response to comment 2 from the 2nd reviewer above.

More information is needed on the rationale for selecting the 3 tests (ferritin, FSH, HPS) that were focused on in this study. Specifically, why were these 3 tests selected out of the 9 tests that were targeted in the randomized trial? What were the other 6 tests and why were they excluded?

We have added the word “choice of ” and “(ie they showed a range of effectiveness of the intervention)” to the study design section of the methods.

This is to clarify that these tests were selected because they demonstrated the breadth of the effectiveness of the interventions and were requested at similar levels prior to the intervention. In particular, following the intervention there were statistically significant reductions in the requesting of FSH, non statistically significant reductions in Ferritin requesting and HPS requesting was unaffected. It was not practical or necessary in our opinion to focus on all nine tests.

Further details of the other tests can be found in the original trial manuscript.

Although completed surveys were returned from 42 of the 43 practices participating in the trial, the authors should moderate assertions of robust study results given that only 57% of the 50% of providers receiving the survey actually completed it (equating to only 28.5% of providers participating in the trial).

Related to this issue, do the authors have any data on survey non-responders that could be compared to responders to assess whether they differ from one another on potentially important characteristics?

See response to comment 4 from the 2nd reviewer above.

Figure 1 was not included in the version of the manuscript I downloaded from the website for the review, though it apparently was only an illustration of the TPB, with which I am familiar.

Not sure why, we will liaise with the journal to solve this.

In the discussion section, it is too strong to state that the meditational analysis “suggested” that intentions to request an FSH or ferritin test was part of the causal pathway in the trial. Because the survey was not administered to providers before the trial began, we cannot be confident that the

We do not agree with the reviewer that we cannot be confident that the interventions had an effect on the TPB construct. The randomised trial result clearly demonstrated that all randomised groups were balanced in test requesting behaviour at baseline and statistically significant decreases in test
interventions had an effect on the TPB constructs (ie, no pre-/post-assessment is available).

requesting was observed after the intervention.

Whilst it may seem theoretically possible that the intentions to request a test were imbalanced at baseline and continued to have the same level of imbalance post intervention, we cannot determine any obvious mechanism for such an effect to have occurred given the balance at baseline in requesting behaviour. We therefore prefer to keep “suggested”.

The following sentence from the discussion section should be rewritten…”In this example, the changes in constructs between intervention and control practices were large and provided compelling evidence of differences." The survey was administered only once (post-intervention), so although the results reflect some differences between interview and control groups, they do not represent any change in scores from an earlier assessment.

See response above. The randomised trial gives us compelling evidence of differences with or without a “change score”.

However we have reworded this sentence to “… In this example, the difference in constructs scores between intervention and control practices were large and provided compelling evidence of changes in construct”

It would be helpful if the authors could discuss concordance/discordance of their study findings to results from the systematic review conducted by Godin et al (citation #9 in the Reference list).

See response to comment 3 from the 2nd reviewer above.