Author’s response to reviews

Title: Goal conflict, goal facilitation and health professionals’ provision of physical activity advice in primary care: An exploratory prospective study

Authors:

Justin Presseau (justin.presseau@ncl.ac.uk)
Jill J Francis (j.francis@abdn.ac.uk)
Neil C Campbell (n.campbell@abdn.ac.uk)
Falko F Sniehotta (falko.sniehotta@ncl.ac.uk)

Version: 2 Date: 2 June 2011

Author's response to reviews: see over
RE: MS 1073940868536086 (Health professionals' provision of evidence-based physical activity advice: A theory-based multiple behaviours approach)

Dear Editors,

Thank you for your positive comments and for the helpful critique from the reviewers. We have addressed your comments and each of the reviewers’ comments in our revised manuscript. Furthermore, we have itemised how we have addressed each of the comments below. We have included the Editor/Reviewer comments in bold, our response in bold italic and where applicable direct quotes from the revised manuscript in standard font. Also, we have altered the title of our paper following your helpful suggestion to: “Goal conflict, goal facilitation and health professionals’ provision of physical activity advice in primary care: An exploratory prospective study”.

On behalf of the co-authors, thank you for considering this manuscript for publication in *Implementation Science*.

Kind regards,

Justin Presseau
Response to Editor Comments

1) Justifying the novelty of this approach. You may already have done in the text but please review.

_We have added the following text to the first paragraph in the discussion to further emphasise the novelty of the approach:_

“This social cognition models and other theories applied in implementation science tend to focus on a single behaviour in isolation from the other behaviours performed in a given context. The present study presents a novel theoretical approach to understanding health professional behaviour. The novelty of this approach lies in the explicit consideration for the alternative behaviours that health professionals engage in and how these are perceived to facilitate and conflict with a particular clinical behaviour. The present study, while exploratory, shows that other clinical behaviours are perceived to help or hinder a given health professional behaviour and that such perceptions predict the reported provision of PA advice beyond PA advice-specific intention and PBC.

The potential of this approach is supported by testing the predictive utility of novel constructs against evidenced theory-based predictors of behaviour (such as those in the TPB). Given the preponderance of theories and theoretical constructs in the literature, the utility of novel constructs for predicting behaviour should be tested against existing theory [24]. Such tests promote theory development and move towards identifying a parsimonious set of constructs which each contribute independently to the prediction of behaviour.” (p.13-14)

“It is also among the few which prospectively measures health professional behaviour and thus heeds current calls from the literature for such longitudinal designs [10].” (p.14)

2) Given the issues highlighted by Reviewer 1, you might wish to consider adding the term 'exploratory' to the study title and elsewhere in the manuscript.

_We have added the term ‘exploratory’ in the manuscript and have revised the study title:_

“Goal conflict, goal facilitation and health professionals’ provision of physical activity advice in primary care: An exploratory prospective study”
3) There is no sample size justification. Can you address this?

_We have addressed this in the method section as follows:_

“To our knowledge, the present study was the first to test goal conflict and goal facilitation as predictors of health professional behaviour in primary care. There was little existing evidence upon which to estimate the effect sizes for a formal power calculation and thus this study was considered to be exploratory. We sent questionnaires to a random sample of health professionals from all 84 GP practices in NHS Grampian and all 69 practices in NHS Tayside, Scotland at baseline, targeting a final sample size of at least 157 health professionals. We estimated a 40% response at baseline and a 65% response at follow-up. Baseline questionnaires were sent to 606 health professionals (453 GPs and 153 nurses).” (p.8)

4) I also note that the effective response rate was 44/606, or 7.2%.

_We have addressed this as follows:_

“The cumulative response rate for the study was 7.3%.” (p.11)

5) Obviously, you did not set out to achieve a low response rate and it might beneficial to set this response rate in the context of similar work and highlight the methodological challenges in work of this type. Given this, it might be worth reinforcing the rationale for using single items in the questionnaire - whilst also acknowledging limitations.

_Thank you for this suggestion. We shortened the limitations and future research section of the discussion and relocated some of its content to a new section which elaborates on the recruitment challenges and use of single items (p.17 to 19). This section also addresses many of the points raised by Reviewer 2 (a direct quote from this new discussion sub-section is included in response to Reviewer 2’s point)._
Response to Reviewer 1 Comments

Overall, we agree with the methodological limitations raised by the reviewer. Nevertheless, we feel that the novelty of the questions being explored and the potential applicability of constructs reflecting the multiple behaviours that health professionals perform in a clinical consultation would be of interest to the Implementation Science readership. The present study is indeed exploratory, but we feel it raises issues concerning the use of social cognition models such as the TPB and theory more generally which are of relevance both in developing theory and also understanding health professional behaviour. In particular, this study recognises that additional theory-based factors need to be explored which may supplement the predictive validity of predictors in the TPB as suggested in the literature (Godin, Bélanger-Gravel, Eccles, Grimshaw, 2008, Implementation Science, 3:36). We nevertheless appreciate and take the reviewer’s points, and address them each in turn below.

1- Regarding sample size: It is well known that tests of the TPB (especially if based on multiple regression) would require a much larger sample size. Moreover, if the aim is to test the theory, structural equation modelling is now the acknowledged approach. A sample size of 44 respondents cannot be accepted for this type of theory testing.

We thank the reviewer for highlighting this point. We fully acknowledge and agree that a sample of 44 is small. However, sample size is independent of effect size. The present study was exploratory in the sense that it is the first to test goal conflict and goal facilitation as predictors of health professional behaviour over and above the TPB. There was little previous evidence to inform a formal power calculation. We have expanded on our recruitment strategy as follows:

“To our knowledge, the present study was the first to test goal conflict and goal facilitation as predictors of health professional behaviour in primary care. There was little existing evidence upon which to estimate the effect sizes for a formal power calculation and thus this study was considered to be exploratory. We sent questionnaires to a random sample of health professionals from all 84 GP practices in NHS Grampian and all 69 practices in NHS Tayside, Scotland at baseline, targeting a final sample size of at least 157 health professionals. We estimated a 40% response at baseline and a 65% response at follow-up.
Baseline questionnaires were sent to 606 health professionals (453 GPs and 153 nurses).” (p.8)

**We have also addressed the issue of sample size in a section in the discussion dedicated to recruitment challenges:**

“Despite a relatively small sample size, this study detected statistically significant effects because their magnitude was large. Although their confidence intervals may be wide, the effect sizes we found can help to inform sample size calculations for future research. Ideally, evidence-based recommendations for increasing responses rates should be used at all phases of data collection when feasible. We utilised many evidence-based methods of promoting questionnaire completion at follow-up, such as printing questionnaires in colour, sending questionnaires using recorded delivery, using shorter questionnaires including stamped return envelopes [30, 31]. However, besides using short questionnaires, pragmatic constraints limited our ability to use additional techniques at baseline.

Small sample sizes are not uncharacteristic of theory-based studies with health professionals. Of the 16 prospective studies testing the TPB in health professionals identified by Godin et al.’s [10] systematic review, seven had sample sizes of 50 participants or less. Furthermore, many of the reviewed studies using postal questionnaires to collect data reported response rates of less than 25%. This underscores the recruitment challenges involved in conducting theory-based research with busy health professionals. These challenges are not new. Indeed, we expected a degree of attrition and this was among the primary justifications for measuring constructs using single items. We kept the questionnaire short to promote a higher response and to reduce participant burden.” (p.17-18)

**We would disagree with the claim that structural equation modelling (SEM) is now the acknowledged or even the most appropriate approach for theory testing. Goodness of fit measures used in SEM take into account a range of hypotheses, including those about measurement models and covariances between remote predictive variables, which go beyond the hypotheses that can be derived from current theory. We would argue that theory should be tested using a multi-methodological approach and testing the additional predictive utility of new predictors over and above established predictors is relevant to a predictive theoretical approach.**

2- **Regarding the reliability aspect: The authors have used “one” item each to assess each theoretical variable. Again, test of a theory should be based on “known”**
psychometric qualities of the variables. One cannot accept as granted that the construct were correctly measured without this information.

Thank you for highlighting this issue. We consider the issue of ‘correct’ measurement to be an issue of validity rather than of reliability. The present study demonstrates support for our hypotheses and results that are consistent with the extant TPB literature in terms of direction and size of effect for predictors of both intention and behaviour. This evidence supports the predictive validity of the measures. We agree that the issue of reliability is important. However, we feel that there is sometimes an over-reliance on multi-item measures of TPB constructs in the literature and of the use of Cronbach’s alpha as a measure of reliability. The approach that we used, which involves deriving measures following standard procedures by Ajzen (2002; “Constructing a TPB Questionnaire: Conceptual and Methodological Considerations”) and Francis, Eccles, Johnston, Walker, Grimshaw, Foy, Kaner, Smith, Bonetti (2004; “Constructing Questionnaires based on the Theory of Planned Behaviour: A Guide for Health Services Researchers”), has been shown to produce reliable measures in a multitude of studies. Given the well-recognised recruitment challenges in this field, we would argue that requiring the use of multi-item measures for each study would unduly limit the advancement of behavioural science and implementation science. Instead, there is a trade-off between the ability to assess the reliability of measures, the explanatory value of any study (by including more variables) and, obtaining higher response rates. We have added a dedicated discussion point which directly addresses the reviewer’s concerns:

“By convention, TPB studies typically assess constructs using multiple items and report an index of internal consistency, but such operationalisations do not address issues of validity. Multi-item measures used to assess intention tend to vary a single word in each item, often using words with similar but not identical meaning, assuming that they tap the same construct. While this may promote a high Cronbach’s alpha, some wording reflects measures of demonstrably distinct constructs and may be theoretically unjustifiable. For instance, a prototypical intention item is worded: ‘I intend to do X behaviour in Y context at Z time’. However, additional items using similar wording such as ‘I want’, ‘I expect’, and ‘I plan to’ are commonly recommended intention items, despite being arguably related to separate constructs: ‘I want’ measures desire [32], ‘I expect’ measures behavioural expectation [33],
and ‘I plan to’ can be viewed as a facet of a post-intentional planning measure [34]. The single item measures used in the present study allowed us to circumvent this issue.

The quality or presence of psychometric properties of predictors of health professional behaviour does not appear to be an effect modifier of the magnitude of the relationship between predictors and behaviour [10]. Scores on single items measures may be associated with behaviour to a similar magnitude as scores from multi-item measures, as we observed.

The present study was exploratory and serves to demonstrate that, given the consistently observed good psychometric properties of standard items across numerous studies, single items might be considered as an alternative to multi-item measures. Observed means and standard deviations from single items were consistent with other studies using composite scores based on multiple items [27, 35]. In addition, the amount of variance in intention and behaviour accounted for (46% and 48%) was in line with mean frequency-weighted R2 observed across TPB studies reported by Godin and colleagues’ review [10] (59% and 31%, respectively). Despite the limitations of single item measures, this study shows that scores based on such items can be effectively used in multiple regression-based analyses, and means, standard deviations and effect sizes are similar to those garnered from composite scores based on multi-item scales.” (p.18-19)

3- Regarding the validity of the dependent variable: Given that it is a self-reported measure, we should at least be provided with its reliability value.

*We agree with the reviewer that self-reported measures have their limitations. We nevertheless feel that our measure of self-reported behaviour is defensible because it controls for the number of patients seen for whom physical activity advice could be provided and it reduces recall bias as it is limited to the previous two weeks. This is the best available measure of the dependent variable, as objective measure of this behaviour attributable to an individual clinician is currently not recorded in the UK. Obtaining this information from patients would require a complex consent procedure potentially biasing the target behaviour of the GP. Computing a reliability value, such as a Cronbach’s alpha would have required asking several questions essentially asking for the same content. We did not think that would have added to the study. We hope this addresses the reviewer’s point, but would be happy to consider this further if the reviewer feels strongly about it.*
4- I have noted the arguments listed by the authors in the discussion to support their methodological approach. However, these have not convinced me of their relevance in the present context. For instance, when only one item per construct is used, this should at least be controlled by using a large sample size and providing test-retest reliability. Also, it is not because others have published studies based on small sample sizes that it is justified to continue this practice nor to accept this practice.

We understand the reviewers concerns and hope that our responses herein and in our revised manuscript provide a more convincing case for our methodological approach in this context. We also wish to emphasise that the study was exploratory. We have made that clearer both in the title and throughout the manuscript:

“Goal conflict, goal facilitation and health professionals’ provision of physical activity advice in primary care: An exploratory prospective study”

We also agree that test-retest reliability indices and a large sample size would be preferred. Clearly more research is needed which addresses the limitations in the present study. However, as an exploratory study addressing an under-researched theoretical and applied issue, we feel that the present study contributes to the literature and would be of interest to researchers interested in developing behavioural theory and applied implementation issues. We have added a discussion point to suggest how future research could use different strategies to promote a higher response and use test-retest reliability (as mentioned in response to the reviewer’s second point).

We also agree with the reviewer that small sample sizes observed in previous studies does not justify aiming for small sample sizes. We regret that this was the impression that our previous wording provided. We wanted to highlight the recruitment challenges of conducting theory based studies with health professionals have also been shown in other studies. We also clarified in the text that we aimed for a higher response rate in the first instance:

“To our knowledge, the present study was the first to test goal conflict and goal facilitation as predictors of health professional behaviour in primary care. There was little existing evidence upon which to estimate the effect sizes for a formal power calculation and thus this study was
considered to be exploratory. We sent questionnaires to a random sample of health professionals from all 84 GP practices in NHS Grampian and all 69 practices in NHS Tayside, Scotland at baseline, targeting a final sample size of at least 157 health professionals. We estimated a 40% response at baseline and a 65% response at follow-up. Baseline questionnaires were sent to 606 health professionals (453 GPs and 153 nurses).” (p.8)
Response to Reviewer 2 Comments

Major Compulsory Revisions:
1) the length of time between baseline (when variables of primary concern in this manuscript are assessed) and follow-up assessment of primary outcome behaviour is reported, is 6 months. No rationale or explanation is provided as to why this time lag was so long, or how transient/stable measures of goal conflict or goal facilitation are. There is literature suggesting the appropriate length of time one can "safely" measure intentions and PA behaviour - this would be important for the reader to know as well so that the present study findings could be placed into context. I would encourage the authors to explain their rationale for the chosen measurement timepoints and the potential implications of trait vs. state goal conflicts and or facilitative GDB (e.g., is it plausible that goals and associated behaviours would have changed within 6 months?).

We thank the reviewer for these helpful comments. Regarding the rationale of the length of time between baseline and follow-up, we added the following to the manuscript:

“This length of follow-up is consistent with previous research testing goal conflict and goal facilitation in other settings [18, 19] and tests of the TPB in this population [22, 23].” (p.9).

Regarding the appropriate length of time one can ‘safely’ measure intention and PA behaviour, we are not aware that this has been evidenced in health professional behaviour. We are aware of literatures in other populations which show that the size of the relationship between intention and behaviour tends to decrease as a function of time between measurement occasions (McEachan, Conner, Taylor, Lawton, in press, Health Psychology Review). We are also aware of the literature concerning the stability of intentions as a moderator of intention-behaviour relationships (Conner & Godin, 2007, Psychology & Health, 22, 875-897). In the context of health professional behaviour, we would assume that intentions to engage in behaviour are relatively stable over time, though research is needed to test this hypothesis. The stability of goal conflict and facilitation is a very interesting point. Riediger and Freund (2004; Personality and Social Psychology Bulletin, 30, 1511-1523) showed that goal facilitation predicted behaviour 4 and 5 months after baseline, but not at months 1, 2 and 3. However, their study was conducted with members of the general population rather than the potentially more stable contexts defined
by clinical practice. Unfortunately the data in the present study do not allow us to say very much about transience vs. stability of goal conflict and facilitation. This could and should be addressed in future research, and we have added the following text to the discussion to highlight this point:

“It is not clear to what extent goal conflict and facilitation vary over time in health professionals. Future research could consider whether the stability of goal conflict and facilitation might moderate the relationship between these factors and behaviour.” (p.21)

2) While reading the manuscript, I was trying to envision what types of goals would conflict or assist with the “focal GDB” for a health care provider. Could you please provide some examples of what types of goals/scenarios you might think health care providers might be experiencing when proving PA advice? I think this simple change will assist the reader with understanding what you are referring to.

We again thank the reviewer for suggesting this revision. We previously conducted a qualitative study which highlighted the sorts of GDBs that primary care health professionals perceive as conflicting and facilitating with giving physical activity advice (Presseau, Sniehotta, Francis, Campbell, 2009, Implementation Science, 4:77). We added examples from this study to guide the reader as suggested, as follows:

“For instance, GPs have reported that addressing the patient’s agenda, treating acute illnesses, and prescribing to reduce cholesterol are among the GDBs perceived to conflict with giving PA advice [13]. Furthermore, taking a patient’s history, addressing alcohol consumption and smoking, checking BMI, and addressing well-being and stress are perceived by GPs to facilitate giving PA advice [13]. Thus, not only do health professionals engage in numerous behaviours but many of these are also perceived as facilitating or conflicting.” (p.6)

Minor Essential Revisions:

3) Upon first read, it was a little unclear to me as to whom was perceiving PA recommendations to be in conflict or facilitating with their other personal goals– the health care practitioners or the patients. In other words, I didn’t find a section where it was clearly outlined that health care practitioners had a personal goal to encourage PA in their patients with uncomplicated hypertension, and had other personal goals to that
would specifically influence that PA-Rx goal of theirs. If the authors are suggesting that the health care practitioners’ do not need to have a personal goal but rather a more “generic” practice-based goal (i.e., I strive to make all my patients with hypertension more physically healthy, and as such it is implied that I would want to tell them about PA), perhaps this could be specified?

Based on goal literature, I would think that knowing whether goals were self-selected, assumed/implied or imposed is an important piece of information for readers to know.

We thank the reviewer for this important point. The provision of PA advice was selected as a goal on the basis of it being recommended in clinical practice guidelines yet underperformed in clinical practice. Thus it is not so much a ‘personal’ goal necessarily (though it may be), but rather a professional goal that is suggested by guidelines. But we take the point that there is a need to be clearer that in this instance we are solely interested in goal conflict and facilitation at the health professional level. We added the following sentence to the introduction to further emphasise this point:

“As such, the present study was interested in conflict and facilitation between a health professional’s GDBs.” (p.7)

4) I felt that there was a slight discordance between the concluding paragraph and the actual measures used in this study. I agree with the authors’ summarization of how this data may be used in the applied setting, but wonder if the beginning of this paragraph could be tweaked to more align with the measures used.

Thank you for this suggestion. We reworded the first sentence of the concluding paragraph to more accurately reflect the measures used, as follows:

“The present study demonstrated that the strength with which primary care health professionals perceive their other GDBs to facilitate and conflict with them giving PA advice predicts how often they report providing such advice, over and above the TPB.” (p.21)
Discretionary Revisions:

5) With respect to whether the title and abstract adequately convey what has been found… I’m not sure if "evidence-based" is needed in the title?

*Thank you for the suggestion. We have altered the title to the following:*

“Goal conflict, goal facilitation and health professionals’ provision of physical activity advice in primary care: An exploratory prospective study”

6) I find some wording of a few sentences in the abstract and manuscript a tad confusing. Perhaps it is just me, but I wonder if the following statements could be clarified:

-ABSTRACT BACKGROUND "Goal conflict and goal facilitation may influence health professionals' performance of guideline-recommended behaviours". It comes across as if the professional is doing the performance of the behaviour, not prescribing it.

We appreciate this feedback as it is crucial that we convey whose behaviour it is that we are interested in. In this instance, we suspect the issue is the use of the word ‘performance’ to convey performing/engaging in a behaviour. This is slightly different from the literature on performance in the physical activity domain (i.e., doing physical activity). We tend to use ‘performance’ to convey engaging in any behaviour, including health professional behaviours. But we take the point that this may be confusing. We have altered the wording in the sentence highlighted by the reviewer to:

“Goal conflict and goal facilitation may influence whether health professionals engage in guideline-recommended behaviours and may supplement the predictive power of the theory of planned behaviour.” (p.2)

-pg 4, line 4 onwards

We clarified the sentence on line 4 of page 4 as follows:

“When viewed as such, theories of human behaviour can be employed to identify factors that predict the behaviours involved in translating research evidence into practice [2]” (p.4)
We reworded as follows:

“A core assumption in the TPB is that constructs in the theory sufficiently account for all effects on behaviour [25]. While this assumption argues for the necessity of including factors such as intention and perceived control when predicting behaviour, the present study suggests that these factors may be necessary but not sufficient.” (now p.14-15).

7) Use of the word “moderately” is used frequently when discussing relationships between study variables. Could you define and/or reference what you are terming “moderate”?

Thank you. When referring to relationships between variables, we used moderate to mean a medium effect size in Cohen (1992; Psych. Bulletin, 112, 155-159) terms. However, the reviewer’s point is well taken and we have altered the wording in these instances to the more conventional terminology and cited Cohen in the first instance of its mention (p.7).

8) The authors are experts in the topic of goal conflict and goal facilitation – but the readers may not be. I wonder if a brief explanation could be inserted in the introduction as to why it may or may not be appropriate to have a bidirectional measures of conflict – facilitation? Just a suggestion for the reader’s sake.

This is a helpful suggestion; thank you. We added the following sentence to the methods section:

“Factor analytic and predictive evidence has shown that goal conflict and goal facilitation are best considered as independent constructs, and were therefore measured separately [18, 19].” (p.10)

9) It is reported that gender was assessed… do you mean sex?

Yes. We changed the terminology to ‘sex’.

10) I am curious as to whether the measure of goal facilitation “During these consultations, other things I do helpfully lead me to five lifestyle advice for increasing PA” was clearly comprehended by participants. “Other things” may or may not refer to
goals, or goal-directed behaviours, and “helpfully” may be interpreted a number of different ways within the context of this item.

While ‘other things’ on their own may not refer to goals or goal directed behaviours, asking about ‘other things you do’ is in our view a defensible way to operationalize ‘goal-directed behaviours’ in a way that does not require too much additional explanation. We agree that ‘helpfully’ conveys a generalised level of facilitation, and this was explicitly designed to do so. The wording is consistent with wording used in previous research in another other population (Presseau, Sniehotta, Francis and Gebhardt, 2010, British Journal of Health Psychology, 15, 905-919) as noted on p.9 of the manuscript:

“Measures for goal facilitation and goal conflict were adapted from existing scales [18, 19] into two single items (to maximise response rates) and assessed at baseline.” (p.9)