Reviewer's report

Title: Development and initial validation of the Patient Safety Practices Questionnaire

Version: 1 Date: 15 March 2013

Reviewer: Justin Presseau

Reviewer's report:

The manuscript reports the development and preliminary validation of a questionnaire assessing theoretical factors associated with behaviours promoting patient safety. The theoretical factors identified are based on an operationalization of the Theoretical Domains Framework (TDF; Michie et al 2005). The manuscript makes the important implication that patient safety can be tackled through behavioural and behaviour change approaches (an under researched but highly promising area), and provides a description of a quantitative scale under development which may be used to potentially identify theoretical domains that could be targeted to promote change in clinician behaviours to improve patient safety. Such a tool, particular given its parsimony, would be quite useful in the field for diagnosing potential barriers. While the scale shows promise, I have a number of concerns, which I hope can be clarified, listed below:

Major Compulsory Revisions

- Paragraph 3 of the background suggests that no other quantitative questionnaire exists assessing the TDF for health professional behaviour. That is not the case. Beenstock et al 2012, Implementation Science, used a questionnaire to assess midwives smoking cessation support for pregnant women, and Amemori et al 2011 Implementation Science, 6:50, devised a TDF-based questionnaire assessing barriers to implementation of tobacco use prevention and smoking cessation counselling in dental providers. A stronger case should be made for the need for the present questionnaire given the existence of other similar questionnaires, which may be strengthened by a description of how they are similar and different from one another.

- Much is said about the need for establishing the key behaviour at the start of the methods, but the key behaviour does not seem to actually be specified anywhere, or if so, perhaps it could be made more explicit.

- Describing the 11 domains of the TDF as ‘domains of behaviour change’ is inaccurate. Most of the domains assemble constructs from theories of behaviour; only a subset focus specifically on constructs from theories of behaviour change. Unless the wording of the items in the questionnaire reflect behaviour change (which does not seem to be the case for this questionnaire) I would hesitate to refer to them as domains of behaviour change and suggest remove all such mention in the manuscript.
- It is not clear why ‘Action planning’ is selected as a domain instead of ‘behavioural regulation’, which is what the domain is named in the TDF. If the questionnaire is operationalizing the TDF domains, I recommend using the domain labels in order to maintain the consistency in the literature.

- It is also not clear why ‘Nature of the behaviour’ is not operationalised. Could this be clarified?

- In the ‘participants and procedures’ section, please further clarify the rationale for the sampling strategy, in particular the oversampling of junior doctors. Should the recruitment not be of those who actually do this behaviour rather than solely based on their professional job role? Was there any filter based on their actual involvement in this behaviour?

- More could be made of the use of CFA to validate the TDF domains, in the discussion. This appears to be among the first studies to use a confirmatory factor analytical approach applied to the TDF domains.

- The discussion notes that the ‘novel design allows for application to a range of patient safety behaviours’; however, this was not demonstrated in the data provided.

- The discussion suggests that the PSPQ could be used to develop ‘theoretically underpinned large-scale interventions’. It seems too early in the development of this scale to be able to claim this.

- The post-hoc language used in the third paragraph of the discussion (e.g., ‘marginally missed the boundaries’) is not particularly convincing. It would be more appropriate to be clear that the RMSEA did not meet agreed standards and that this has implications for the need for further development of the scale.

- The discussion and conclusion seems to waver between recommending this scale for use to develop interventions and recommending the need for further development of the scale itself. It would be important to present a consistent message to guide the reader about the status of this scale.

- Some more details are needed to justify the ‘total barriers’ score. What does such a score actually mean? The TDF domains are arguably inter-related, and many of the constructs from within different domains stem from complex mediating processes from the theories from which they are sourced. A sum score seems to equate all domains as equivalently proximal predictors of behaviour, which does not seem appropriate. This has implications for the total barrier score’s meaning.

- Results in Table 1 seem to imply that by and large, across domains barriers were relatively low for this behaviour. What implications does this have for intervention, and what implication does this have for the validity of the assessment if no clear domain appears to be problematic and yet the patient safety behaviour quality gaps persist?

- Some of the items are worded in such a way as to reflect a positive inclination (e.g., I am clear about what my role should be in the process’) while others are negatively worded (e.g., Training is not adequate). It should be clarified whether the items were reverse scored. Furthermore, perhaps it would be worth
discussing whether the meaning of a negatively-worded item is qualitatively different from a positively worded item and whether reverse scoring sufficiently addresses this.

- I have concerns about the operationalization of particular domains, detailed below. It may be worth some additional description about how the items were developed in the first place – where were the items sources from? Were existing validated scales used for any items? Or were all created de novo?:
  - Beliefs about capabilities: arguably, the concept of confidence is a key aspect of this domain. Removing it for statistical reasons seems to have resulted in a loss of construct validity.
  - Beliefs about capabilities cap3 (‘I have previously encountered problems when try to…’): It is not clear which construct within this domain this is assessing.
  - Motivation and goals: it seems inappropriate to remove the ‘I intend to (mg2)’ item, which arguably is the essence of the domain. It is not surprising that the other two items for this domain do not correlate well with intention as they both reflect goal conflict, which has been demonstrated to not correlate particularly well with intention and which accounts for variability in clinician behaviours and health behaviours over and above intention. Perhaps worth citing some of this literature to justify the decision to exclude the intention item, but still seems inappropriate that the motivation domain not include a measure of intention given its centrality in most theories of behaviour.
  - Memory, attention and decision-making: by omitting ‘I forgot to (cog1)’, the ‘memory’ aspect of this domain is no longer featured. This seems to lose an important diagnostic aspect of this domain as ‘forgetting’ has some elegant solutions for behaviour change techniques and may in some behaviours be an important factor.
  - Action Planning: action planning is a well developed theoretical construct with existing validated scales for assessing action planning (see e.g. Sniehotta et al 2005), it is not clear why these were not preferred.
  - Action Planning: the items ‘plans in my head often get muddled when trying to… (ap2)’ and ‘Things are too unpredictable to make plans to… (ap3)’ are arguably not to do with the concept of action planning. AP2 seems more a measure of decision processes than action planning. AP3 seems more a measure of environmental context than action planning. Overall, the two items selected to reflect this domain do not convincingly reflect it.
  - Why was ‘Action Planning’ selected as a domain? Is the domain not ‘Behavioural Regulation’ which involves a whole range of active techniques for promoting behaviour change? Why was only Action Planning selected from amongst that list?
  - Overall, many of the decisions to omit particular items from each domain seem to highlight that some domains are more heterogeneous than others and to assume that all constructs within the domain would reflect one underlying latent variable may not be appropriate. In addition, the decision to omit some items may be a feature of the behaviour and/or the populations studied. Would it be more
appropriate to recommend that future research use the full initial set of items to test whether the CFA results can be replicated?

Minor Essential Revisions
- unclear why ‘experts’ are in inverted commas (background, paragraph1, line6). Suggest removing.

- The line of argumentation in the first paragraph of the background could be tightened. For instance, the juxtaposition of ‘educating, persuading or reminding people’ against ‘adopting a more sophisticated and theoretical approach’ needs to be further expanded (background paragraph 1, line 7-9). Techniques associated with persuasion and reminders can be powerful strategies for promoting behaviour change and are themselves embedded within behaviour change theory.

- I can see (and I agree) that there is a desire to make a point about the need for theory for promoting a cumulative science permitting replication and assessment of generalizability. However a more convincing case needs to be made for why ‘a more sophisticated approach’ is needed. At the point, it comes across as the need for sophistication for sophistication’s sake, which I do not think embodies the idea behind the TDF or Occam’s Razor more broadly.

- There are too many ideas in the opening paragraph, and it may help to split into points about the advantages of theory over intuitive approaches followed by the challenges associated with theory selection.

- The term ‘determinants’ is used throughout. The term has causal implications which may not be appropriate for all constructs and certainly not for many designs that this questionnaire may be used for. I suggest avoiding the use of the term determinants.

- I welcome the comparison between methods of applying the TDF in paragraph 2 of the background but there is an underlying sceptic tone for interviews and qualitative methods which does not appear warranted. In particular, the purpose of qualitative research is often not for achieving generalizability nor is the purpose necessarily to achieve large sample sizes; this is not a limitation, but rather a reflection of the different aims of that methodological approach. I recommend rewording this section to highlight the differences in a more objective way. The authors may want to consider a pros and cons table of different methods of collecting TDF-based data.

- Also, the method of data collection (interviews) should not be confounded with qualitative methods; it is possible to conduct quantitative interviews.

- A clear statement of the aims in the background would be helpful.

- Paragraph 5 of the background seems more of a discussion point than a background point.

- It is helpful to have the Context section; however, it seems more of a methods point, perhaps worth relocating to that section?

- It would be helpful for the uninitiated to have a bit of background about the behaviour of interest in the present study. E.g., what is the recommended
process of care, what is the evidence for quality gaps, and what are the implications for patient safety.

- The last sentence in Footnote 2 can be omitted as the same sentence appears in the main text in paragraph 2 of the data analysis section.
- What was the method of missing value imputation used?

**Discretionary Revisions**

- The authors may wish to reconsider the name of the questionnaire (i.e., ‘The Patient Safety Practices Questionnaire’; PSPQ) for two reasons. First, use of ‘The’ feels a bit too grandiose for a questionnaire that is still in the fairly early stages of development. Rather than ‘The’, perhaps consider ‘A’? Secondly, the title does not really reflect the fact that the questionnaire is not about patient safety practices, but rather about beliefs, perceptions, and cognitions about patient safety practices. The title does not really seem fit for purpose or be sufficiently informative about the TDF-based aspects. For instance, some may assume that the questionnaire is assessing a range of skills associated with promoting patient safety. Perhaps reconsider the title of the scale?
- Fascinating and very important point about the challenges associated with how we select one behaviour from amongst the range of possible recommendations. I wonder whether more could be presented here, as this is a recurring issue in the application of behavioural theory to implementation-related behaviours: in a context involving multiple behaviours, what are our criteria for selecting one? And should we select only one if clinicians are performing multiple behaviours? Perhaps some elaboration on this issue would be helpful for readers.
- More could be made about patient safety behaviours. Behaviour change approaches to promoting patient safety are novel and this manuscript could present a larger case for considering patient safety behaviours (perhaps in the background?)
- Suggest omitting the ‘etc’ on the third line of the background
- Remove extra ‘the’ on line 10 of the first paragraph of the background

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests