PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Assessing societal and offender perspectives on the value of offender health care: A stated preference research protocol
AUTHORS	Settumba, Stella; Shanahan, Marian Shanahan; Chambers, Georgina; Schofield, PW; Butler, Tony

VERSION 1 - REVIEW

REVIEWER	Dr Tim Kirkpatrick Plymouth University UK
REVIEW RETURNED	06-Sep-2018

GENERAL COMMENTS	This is a very interesting and worthwhile topic, especially given cuts to criminal justice budgets over recent years. The protocol describes a rigorous methodology and provides sufficient detail to enable replication. General comments: The protocol is well written, but in many parts is highly technical. Whilst I appreciate this is, in part, due to the methodologies employed, I wonder whether giving some examples (e.g. of anticipated attributes and possible levels) would make the technical language more digestible. Given poor literacy levels in offending populations, I wonder if any consideration was given to address this. Might be useful to include a sentence to say if this was considered, or if not, why not. Specific comments: There are no dates of the study included in the manuscript. Under the aims of the study, the abbreviation WTP is used, but it is not previously stated what this means. In the section on the Delphi method it states that the attributes and levels were further deliberated on by a team of experts. Later on, it states that an advantage of the Delphi method is that is it avoids domination of the consensus process by experts. Whilst is agree with this point, it seems irrelevant to the study as you are using a 'team of experts'. A minor change to the wording of the text could get round this matter. In the section on Experimental design and pilot (second paragraph) – should the end of the paragraph read ' treatment 1 and treatment 2' In the section on Contingent Valuation method – sometimes you write CV and sometimes CVM – I don't mind which you use, but you should be consistent. In the section on Data Anlaysis – first line – the word 'To' shouldn't have a capital letter.

REVIEWER	Ekant Veer
	University of Canterbury
REVIEW RETURNED	04-Nov-2018

GENERAL COMMENTS	Thank you for your protocol on assessing public and offender perspectives of offender healthcare. The manuscript is well written and described and covers most of the major areas necessary when completing a protocol. The authors propose to tackle quite a complex concept with some serious competing values, which could yield some fascinating results.
	My concerns are minor and I feel they would be simply addressed in a revision. The first relates to the methodology employed and whether the authors can provide greater information about the sampling procedure. How will the final sample be effectively representative of the wider population and how can the researchers avoid experimentation biases that are inherent with panel data. One could argue that the use of pre-determined panels are inherently biased and non-representative as many people in the population do not choose to take part in such surveys. How are the participants compensated for their time and how does this inducement impact their responses?
	Secondly, with regards to the analysis method, a little more information would be appreciated about the various measured variables in the study. More detail regarding the various measurements and their validity (both internal and external) would help to determine whether the data gathered is valid.
	Finally, I would like some further theoretical rationale for the rationale of comparing public sentiment with that of offenders and their families. One can assume that there would be greater sympathy for healthcare costs from offenders and their families towards offenders but not as much from the general public. What else are the researchers hoping the learn? What other variables would be included as covariates of the study to ensure we are getting a full picture of the phenomena in question. Are the public's political preferences being measured? If so, how is this incorporated into the model. This may be cleared with further examination of the variables as requested earlier.
	Overall, an interesting study and one I look forward to seeing the results of.

VERSION 1 – AUTHOR RESPONSE

RESPONSE TO REVIEWERS' COMMENTS

Reviewer: 1

Reviewer Name: Dr Tim Kirkpatrick

Institution and Country: Plymouth University, UK

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below.

This is a very interesting and worthwhile topic, especially given cuts to criminal justice budgets over recent years. The protocol describes a rigorous methodology and provides sufficient detail to enable replication.

General comments:

The protocol is well written, but in many parts is highly technical. Whilst I appreciate this is, in part, due to the methodologies employed, I wonder whether giving some examples (e.g. of anticipated attributes and possible levels) would make the technical language more digestible.

Response: In the manuscript we provided the following definition for attributes and levels in the section 'developing attributes and levels':

The attributes characterise the factors considered by offenders and society to be most valuable for acceptance, support and uptake of treatment programs by impulsive violent offenders and the levels are the ranges over which the attributes vary

We have added the following sentence to give an example of an attribute and attribute levels.

"An example of a common attribute used in DCE studies is 'cost of the treatment program'. The attribute levels would be the various dollar amounts the treatment might cost."

In figure 3, we provided an example of a DCE choice scenario with examples of attributes i.e. characteristics of treatment, and examples of attribute levels under treatment 1 and treatment 2.

Given poor literacy levels in offending populations, I wonder if any consideration was given to address this. Might be useful to include a sentence to say if this was considered, or if not, why not.

Response: For the FGDs, we had separate groups for offenders and the general public and the questions were tailor made for the different groups. The facilitator was an experienced qualitative researcher with vast experience working with offender populations. We have added this text in the manuscript.

While it is true that offender populations have poorer literacy levels compared to the general population, we also know that the general public has distorted perceptions of crime (Weatherburn and Indermaur, 2004). For all groups, we initially explored their views and perceptions and then provided them with correct information.

'During the FGDs, after exploring participants' knowledge and views on violence, impulsivity, incarceration, recidivism and the role of treatment programs, they were provided with precise definitions of terms, examples of available interventions and contemporary statistics on violent crime, incarceration and recidivism rates.'

For the quantitative component, recruitment for the online survey will be from the general public and people will self-identify as offenders. The questionnaire will be pilot tested to make sure it is understood by the participants.

Specific comments:

There are no dates of the study included in the manuscript.

Response: Dates have now been included and the following sentence added.

"This study will be conducted between July 2017 and May 2019"

Under the aims of the study, the abbreviation WTP is used, but it is not previously stated what this means.

Response: This has been included under the aims of the study.

In the section on the Delphi method it states that the attributes and levels were further deliberated on by a team of experts. Later on, it states that an advantage of the Delphi method is that is it avoids domination of the consensus process by experts. Whilst is agree with this point, it seems irrelevant to the study as you are using a 'team of experts'. A minor change to the wording of the text could get round this matter.

Response: Thank you. The wording has been changed to the following:

"It is popular because, in addition to providing an opportunity for everyone's views to be taken into consideration by the group, it allows anonymous voting and avoids the domination of the consensus process by a few individuals."

In the section on Experimental design and pilot (second paragraph) – should the end of the paragraph read ' treatment 1 and treatment 2'

Response: Thank you for pointing this out. Yes it should read 'treatment 1 and treatment 2'. The change has been made.

In the section on Contingent Valuation method – sometimes you write CV and sometimes CVM – I don't mind which you use, but you should be consistent.

Response: This has been changed and CV is consistently used throughout the manuscript.

In the section on Data Analysis – first line – the word 'To' shouldn't have a capital letter.

Response: This has been corrected.

Reviewer: 2

Reviewer Name: Ekant Veer

Institution and Country: University of Canterbury

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

Thank you for your protocol on assessing public and offender perspectives of offender healthcare. The manuscript is well written and described and covers most of the major areas necessary when completing a protocol. The authors propose to tackle quite a complex concept with some serious competing values, which could yield some fascinating results.

My concerns are minor and I feel they would be simply addressed in a revision. The first relates to the methodology employed and whether the authors can provide greater information about the sampling procedure. How will the final sample be effectively representative of the wider population and how can the researchers avoid experimentation biases that are inherent with panel data. One could argue that the use of pre-determined panels are inherently biased and non-representative as many people in the population do not choose to take part in such surveys. How are the participants compensated for their time and how does this inducement impact their responses?

Response: Members of the public for the DCE survey will be recruited from an outsourced online panel provider, Survey Sampling International (SSI) where respondent duplication and fraudulent completion of surveys is monitored. Participants are recruited via verified, certified sources and methods to create a vast pool of potential research respondents for clients.

SSI has actively managed proprietary panels across the globe, including in New South Wales, and these are at the core of their online sample. SSI improves the quality and representative nature of its panels by incorporating participants from partnership sources which SSI also actively manages. Participants are recruited via partnerships, invited via banners, invitations and messaging, and then go through rigorous quality controls before being included in panels. The survey population will be selected to represent the distribution of the general Australian population in terms of age, sex and geographical location (metropolitan and rural NSW).

Respondents are incentivised by the survey company in different ways, depending on the amount of effort required, the population, and appropriate regional customs resulting in higher panel respondent satisfaction.

Recognizing that motivations may be different across different demographic groups, SSI methodology allows participants to be rewarded in the way that makes most sense for them. Participants choose between the following pre-defined set of incentives: win a prize e.g. a movie ticket, earn cash, and donate to a charity. It is not possible to provide an absolute value that a participant might be paid to undertake the survey (but is generally the approximate value of 2 movie tickets). It is important to note that many participants choose not to be paid or rewarded, but participate altruistically.

Secondly, with regards to the analysis method, a little more information would be appreciated about the various measured variables in the study. More detail regarding the various measurements and their validity (both internal and external) would help to determine whether the data gathered is valid.

Response: In this protocol, we cannot provide more information about the variables as we do not know the final variables used in the DCE. These will be decided after the experimental design and pilot studies. We will report more details when we write up the methods and results for the DCE and CV.

As explained in the manuscript in the second paragraph of the section 'experimental design and pilot', attributes in this study will be described by a continuous, discrete or categorical scale and therefore the experimental design will be optimised for a multinomial logit model and this will then be evaluated in NGENE using a panel mixed logit model, which accounts for the parameter distribution, and a latent class model which accounts for non-uniformity of respondents. The final model used in our DCE analysis will depend on which model best describes the data.

External validity will be tested using convergent validity and the following text has been added to the second last paragraph in the manuscript.

'To test external validity, we will use convergent validity to compare the results from the DCE to those of the CV method.'

For internal validity, we demonstrate the rigor with which the study was conducted. We employ a mixed methods design. In developing attributes and their levels we used different group types: offenders (both with and without treatment), family members of offenders, the general public, and experts in the justice and health fields to make sure to capture all different societal attributes. Care is being undertaken to correctly apply the DCE and CV health economics methodology to this field.

Finally, I would like some further theoretical rationale for the rationale of comparing public sentiment with that of offenders and their families. One can assume that there would be greater sympathy for healthcare costs from offenders and their families towards offenders but not as much from the general

public. What else are the researchers hoping the learn? What other variables would be included as covariates of the study to ensure we are getting a full picture of the phenomena in question. Are the public's political preferences being measured? If so, how is this incorporated into the model. This may be cleared with further examination of the variables as requested earlier.

Response: We hypothesise that the attributes or attribute levels for treatment programs preferred will be different for the various groups. Since we are looking at a societal perspective, it is important that we capture the various group type differences. Offenders benefit directly from the treatment programs especially as they target their health, while families and the general public benefit through safer communities as a result of reduced crime. While offenders may be particularly interested in the actual delivery of treatment, the general public might be interested in treatment outcomes such as reduced societal crime. On the other hand, the general public tax-payers are the funders of treatment programs and it would be interesting to capture their perspectives on the programs they would like to see funded. While we cannot provide any results at this moment to show the differences by group, in our manuscripts reporting results from both the qualitative work and the DCE and CV studies the different covariates will be reported.

Overall, an interesting study and one I look forward to seeing the results of.

Response: Thank you.

VERSION 2 – REVIEW

REVIEWER	Tim Kirkpatrick
	Plymouth University UK
REVIEW RETURNED	04-Dec-2018

GENERAL COMMENTS	Thank you for submitting the revised manuscript which address all points I raised in the original review.
	I would have recommended to accept the manuscript, but there are 4 very minor points that probably need addressing (three of which I missed when I first reviewed the paper - apologies):
	Page 7 (line 19) - it states ".....bringing the total number of FDGs to 8 (4 offender, 2 general and 1 family members group)" - I presume there is a typo with regard to the numbers as 2+4+1=7 Page 7 (line 46 - it states ".......the Delphi method and the results from the voting and ranging exercises uses to provide a qualitative indication of the strength...." I would have thought the results of the voting and ranking exercises would provide a quantitative indication rather than a qualitative indication, but I could be wrong.
	Page 9 (line 58) - It states "Participants are recruited via verified, certified sources and methods and create a large pool of potential research respondents for our clients" - I don't understand why the words 'for our clients' are used.
	Page 9 (line 60) - the words 'The panel' have been added - I presume this is an error.

REVIEWER	Ekant Veer
	University of Canterbury, New Zealand
REVIEW RETURNED	16-Dec-2018

GENERAL COMMENTS	Thank you for your revision and commentary on the changes made. The only concern I have is that the authors have avoided providing a theoretical justification for the potential measures used in the final study. I appreciate that the final model will depend on the data collected and based on how the data interacts but the measurement of this data and the key variables being tested should still have a theoretical basis for analysis - if there is no theoretical basis then an explanation as to why this is the most appropriate means of measuring the concepts under scrutiny is being employed. I'll leave this decision to the editors as they will be more well versed with the current practices in the journal, but in my reading a protocol should be published as a means of guiding future researchers and as it stands I do not feel there is enough detail in the measures being employed to effectively guide future researchers.

VERSION 2 – AUTHOR RESPONSE

Response to reviewer's comments

Reviewer 1:

Page 7 (line 19) - it states ".....bringing the total number of FDGs to 8 (4 offender, 2 general and 1 family members group)" - I presume there is a typo with regard to the numbers as 2+4+1=7

Response:

Thank you for spotting this. The total number of FGDs was 8. A correction has been made on the number of general public FGDs. These were 3 and not 2. The sentence now reads:

Recruitment and FGDs were conducted until saturation was reached, i.e. when no new data was generated with additional groups, bringing the total number of FGDs to 8 (4 offender, 3 general public and 1 family members group).

Page 7 (line 46 - it states ".......the Delphi method and the results from the voting and ranging exercises uses to provide a qualitative indication of the strength...." I would have thought the results of the voting and ranking exercises would provide a quantitative indication rather than a qualitative indication, but I could be wrong.

Response:

We can see how this might be confusing. While voting and ranking may be quantitative in nature, are the methods still quantitative when used within a focus group discussion? To avoid the confusion we have made changes to illustrate that the indication of preference strength from FGDs will be compared with that from the DCE.

However, in this study, all attributes obtained from the FGDs were further assessed through the Delphi method and the results from the voting and ranking exercises used to provide an indication from the FGDs of the strength of the different attributes that will then be compared with results from the DCE.

Page 9 (line 58) - It states "Participants are recruited via verified, certified sources and methods and create a large pool of potential research respondents for our clients" - I don't understand why the words 'for our clients' are used.

Response:

'For our clients' has been deleted.

Page 9 (line 60) - the words 'The panel' have been added - I presume this is an error.

Response:

Thank you. It was an error and has been deleted.

Reviewer: 2

Reviewer Name: Ekant Veer

Institution and Country: University of Canterbury, New Zealand

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

Thank you for your revision and commentary on the changes made. The only concern I have is that the authors have avoided providing a theoretical justification for the potential measures used in the final study. I appreciate that the final model will depend on the data collected and based on how the data interacts but the measurement of this data and the key variables being tested should still have a theoretical basis for analysis - if there is no theoretical basis then an explanation as to why this is the most appropriate means of measuring the concepts under scrutiny is being employed. I'll leave this decision to the editors as they will be more well versed with the current practices in the journal, but in my reading a protocol should be published as a means of guiding future researchers and as it stands I do not feel there is enough detail in the measures being employed to effectively guide future researchers.

Response:

We probably misunderstand the reviewer in his earlier comment regarding this. We have included a theoretical basis for the analysis as follows:

The theoretic underpinning of the DCE analysis is based on Lancaster's theory of choice [42] and the random utility maximisation framework [43]. As shown in equation 1, the utility (U) that an individual n derives from the treatment alternative j in the choice set c is explained by an observed component V_{ncj} and an unobserved component ε_{ncj}.

$$U_{ncj} = V_{ncj} + \varepsilon_{ncj} \text{ (Equation 1)}$$

The observed component of the utility associated with alternative j, V_{ncj}, is a function of a vector of k attributes that describe treatment alternative t, x_{ncjk}, with associated preference weights, β, to be estimated. Such that:

$V\_ncj= \sum\_{(k=1)}^{k} \beta\_k \, x\_ncjk$ (Equation 2)

When faced with a choice task with treatment alternatives an individual will choose i over j if the utility obtained from i is greater than that from j. Such that:

$\llbracket(V)\rrbracket\_nci+\varepsilon\_nci) > \llbracket(V)\rrbracket\_ncj+\varepsilon\_ncj)$ (Equation 3)

VERSION 3 – REVIEW

REVIEWER	Tim Kirkpatrick Plymouth University, UK
REVIEW RETURNED	14-Feb-2019

GENERAL COMMENTS	All my comments have now been addressed. Thank you.

REVIEWER	Ekant Veer University of Canterbury
REVIEW RETURNED	13-Feb-2019

GENERAL COMMENTS	Thank you for attempting to address my previous concerns - the inclusion of a choice model is useful but does not go any further to determine what factors/variables are of particular interest. Again, it is up to the editor to decide whether this is acceptable but a protocol based on data-mining and hoping to find something that may interact is not a great approach, in my field. We must have a theoretical rationale for specific variables otherwise we are just hoping to find interactions and effects that will be retrospectively explained without a theoretical rationale. Best of luck with the research.