PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Perceptions of cigarette pack inserts promoting cessation and dissuasive cigarettes among young adult smokers in the United Kingdom: A cross-sectional online survey
AUTHORS	Moodie, Crawford; Hiscock, Rosemary; Thrasher, Jim; Reid, Garth

VERSION 1 – REVIEW

REVIEWER	Judith McCool University of Auckland, New Zealand
REVIEW RETURNED	07-Oct-2017

GENERAL COMMENTS	A well-presented paper describing the added value of including smoking cessation advice via information inserts in packs and altered cigarettes. Some very interesting results emerged from this work that prompt discussion about how smokers (as in this sample) engage with altered products, smoking cessation information and novel means of attaching this information to a behavior or action (inserts inside a pack). This is useful evidence to contribute to developing interventions that nudge smokers towards quitting and deter uptake. All options for this end need to be considered and tested. The methods are clearly presented, but additional detail on the scales used would be useful. In particular whether the cessation, or quitting measures were adapted from other pre-validated scales or generated for this survey. Methods / measures What is a nominal incentive? Is it is not clear to me what that might entail, but would assume it to be a voucher? The sample method is too brief, need to add a little more information about the method and sample frame for recruitment. Were participants were show all four inserts (as pictured) or just one for each question relating to the inserts? Clarify. The inserts may have a different impact depending on where they included within the pack. For example, if they are expected to be folded and inserted on the inside front or back or fitted neatly inside and were somehow visible / presented when the pack was opened? It is not clear how they are presented inside the pack. Does the smoker need to extract the insert or it is presented at the top of the pack? Given other products with risk and use information sheets included inside, what have we learnt about the value they add to consumer behaviour. More detail on the methods required here.

	Discussion: The finding that the WC and GC were not as unappealing as expected is interesting, but perhaps only because they 'appear' more unhealthy etc. Although this study design is appropriate and effective for gaining a reasonable sample and some descriptive information about consumer responses to product alterations and their impact on perceptions, any subsequent work might look to return to a sample of smokers and examine the drivers behind these unexpected results. Perhaps the standard white cigarette has been tainted effectively via mainstream tobacco control measures that it extremely unlikely to viewed as benign whereas a WC or GC may have a more 'playful' or fake look about them. It is not clear from these results why they are not having the impact. This work is apparently underpinned by any theory of marketing, consumer behavior, risk perceptions but reference to these theories would be useful to support what the authors expected or the rationale for this approach. It has been based on previous studies where there was some indication that changing or tainting the products altered its appeal but that is missing here. I recommend including a statement or two in the background and again in the discussion to acknowledge where assumptions that this study was based on. Some discussion on tone of the insert content (focus being positively framed) and position in the pack (prominent and not able to be discarded until pack is completed?) or otherwise would also be useful. Tthe practical element of this intervention of providing inserts needs to be acknowledged - a removable insert, that would be discarded would may have other consequences, both positive and negative (positive being the information I may be diffused through inadvertent exposure among others to the information) and negative being litter. Finally, as the standard cigarette now occupies a space with e-cigarettes, pipes etc., the shift to having a green cigarette may have less impact or not be as disruptive. A comment acknowledging the changing or diversification of tobacco products may be warranted.

REVIEWER	Ingeborg Lund Norwegian Institute of Public Health Department for Drug Policy, Norway
REVIEW RETURNED	16-Oct-2017

GENERAL COMMENTS	This is a well-designed study based on replies from young adult smokers on questions regarding a selection of cigarette pack inserts as well as three differently designed cigarette sticks. Inserts have been implemented in Canada, and more knowledge about how they are perceived in other countries are particularly timely. My main concern with the current manuscript regards the methods section. In general I found it difficult to read, and difficult to understand. I would encourage the authors to rephrase parts of this section for better clarity. Specifically, this concerns ‘Design and sample’ and the section about the multilevel log.reg. Details Methods P6, Design and sample: The fact that the sample consists of 16-34 year old smokers is

	mentioned twice. One time would suffice. The bracket (n=193) would make more sense if it was moved to just after “completion time”. The size of the original sample is not mentioned explicitly here (although it is said in the abstract). It would be better to include it. A further explanation of the percentage 89.6 is warranted – it seems very high to be the response rate. P 10, first paragraph and first line of analysis: Double check that the references 21 and 22 are correct. It looks like they may have been switched around. P11, middle: Could the authors explain the meaning of the 0.5 in this sentence: “All eight variables loaded on a single factor>0.5.” P11, multilevel logistic regression: I found the explanations of this analysis difficult to understand. The authors should try to reformulate the text. It would be good if you could include some references in support of the more technical details. I also do not understand what is meant by a “cross classified variable”, or that the reference category of the cross classified variable is varied. P15, perceptions of cigarette desirability: I suggest the authors include a table with these results instead of just listing them in the text. P16, Figure 3: There is no figure 3 – should it say table 4? P15/16, table 4: The authors should explain the table better. If the vertical lines represent confidence intervals, I do not understand how one can see any significant differences between groups, as is maintained in the text. Results P14, bottom: The text reads participants were more likely to consider SC attractive, stylish and nicer to be seen with, but %-ages were higher for WC and GC. Is this correct? Also the sentence (p15, top) “the SC (17.8%) was viewed as not as appealing to people their age as the WC (51.5%) or GC (57.4)” is quite confusing, and I do not understand what it means. P15, top: Include % in the bracket (57.4)

REVIEWER	E. Paul Wileyto University of Pennsylvania, USA
REVIEW RETURNED	05-Feb-2018

GENERAL COMMENTS	This is a meat and potatoes kind of data paper. The research question is broad and open, and not hypothesis driven. The authors dichotomize all of the data, so they can analyze with logistic regression. The original data are ordinal in nature, and I am curious why they did not analyze the data as ordinal. That was within the capabilities of the software at hand. The authors used a piece of software called MLWin. Please capitalize the "L". I chased my tail for a while searching for info on MIWin. Looks like it does some things that SPSS cannot. The authors do not report unadjusted odds ratios. Model results should generally show both adjusted and unadjusted ORs side by side. Table 4 (figure 3?) shows an interaction. Authors should report the significance of the interaction somewhere in the caption or notes. Interactions in logistic regression can be parameterized and reported

	as the ratio of odds ratios (ORR), with confidence interval. Some of the language reporting results (odds ratios) in the text is a little clunky and monotonous. I found myself having to reread many of those because alarm bells were going off about sentence structure. They turn out to be fine, but I am guessing I will not be the only one with that response.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1
1) A well-presented paper describing the added value of including smoking cessation advice via information inserts in packs and altered cigarettes. Some very interesting results emerged from this work that prompt discussion about how smokers (as in this sample) engage with altered products, smoking cessation information and novel means of attaching this information to a behavior or action (inserts inside a pack). This is useful evidence to contribute to developing interventions that nudge smokers towards quitting and deter uptake. All options for this end need to be considered and tested.

Response: We thank the reviewer for these remarks.

2) The methods are clearly presented, but additional detail on the scales used would be useful. In particular whether the cessation, or quitting measures were adapted from other pre-validated scales or generated for this survey.

Response: Given the dearth of research on inserts and dissuasive cigarettes, most scales were created by the research team. There are some however that we have taken from past research, mostly for the cessation-related questions, and where this is the case we have referenced the source of these.

3) Methods / measures
What is a nominal incentive? Is it is not clear to me what that might entail, but would assume it to be a voucher?

Response: The incentive was 50 pence. This now reads: ‘Participants received a nominal incentive (50 pence)..’

4) The sample method is too brief, need to add a little more information about the method and sample frame for recruitment.

Response: We now explain in the ‘Design and sample’ that the Research Now UK panel has over 400,000 people. We have also provided more detail on sampling, response rate, and screening, at the start of the Procedure. This now reads: ‘An email invite was sent by Research Now to their online panel in the UK. Research Now is an established online market research company in the UK and elsewhere,21 with their panels recruited from a wide range of sources, such as internet sites, advertising and partnerships with other websites. Research Now, like other online panels, has details of their members’ demographics and other characteristics that are used to profile target samples. Response rate details are not available when using this sampling methodology however as recording contact, participation and refusal rates is not practical.22 For those that responded to the email invite, they answered screening questions about their age, smoking status and types of tobacco products used, with those that did not meet the inclusion criteria (factory-made cigarette smokers aged 16-34 years) excluded.’

5) Were participants were show all four inserts (as pictured) or just one for each question relating to the inserts? Clarify.

Response: We now explain that: ‘For each question about inserts, participants were shown the question and an image of one insert. Four different inserts were used in total, as shown in Figure 2, with these chosen from the eight used in Canada as they were considered most relevant to our sample.’

6) The inserts may have a different impact depending on where they included within the pack. For example, if they are expected to be folded and inserted on the inside front or back or fitted neatly inside and were somehow visible / presented when the pack was opened? It is not clear how they are presented inside the pack. Does the smoker need to extract the insert or it is presented at the top of the pack? Given other products with risk and use information sheets included inside, what have we learnt about the value they add to consumer behaviour. More detail on the methods required here.

Response: Inserts in cigarette packs, which have been used by tobacco companies across the globe for decades, are visible as soon as the pack is open, being positioned in the front of the pack between the outer packaging and inner foil. As noted within the Introduction, a qualitative study with smokers in Scotland found that ‘the general view was that they [inserts] would capture attention and be read due to their novelty and visibility when opening the pack.’ We have included an additional sentence explaining that ‘Inserts were also thought to have a long lasting impact as they would be removed from the pack and remain visible within the household or elsewhere, or as litter’. While smokers are most likely to extract the insert, as they present a barrier to reaching the cigarettes, it is possible for the inserts to be left within the pack – however in either case they are clearly visible. In comparison, the leaflets found in pharmaceutical packaging are less salient, typically being folded and less prominently displayed, and also less engaging, as they typically contain a lot of information.

We now explain in the ‘Procedure’ that participants were first shown an image of an insert in a pack, in order to demonstrate how these appear in packs, and then for each question they were shown the question and one insert, with four used in total. This reads: ‘Prior to the questions on inserts, participants were shown an image of a cigarette pack with an insert shown in the front of the pack – as they typically appear in packs – alongside the text ‘We have some questions on pack inserts, which can sometimes be found inside packs (see image for example)’. For each question about inserts, participants were shown the question and an image of one insert. Four different inserts were used in total, as shown in Figure 2, with these chosen from the eight used in Canada as they were considered most relevant to our sample.’

7) Discussion: The finding that the WC and GC were not as unappealing as expected is interesting, but perhaps only because they 'appear' more unhealthy etc. Although this study design is appropriate and effective for gaining a reasonable sample and some descriptive information about consumer responses to product alterations and their impact on perceptions, any subsequent work might look to return to a sample of smokers and examine the drivers behind these unexpected results. Perhaps the standard white cigarette has been tainted effectively via mainstream tobacco control measures that it extremely unlikely to viewed as benign whereas a WC or GC may have a more 'playful' or fake look about them. It is not clear from these results why they are not having the impact.

Response: We found, among our sample of smokers, that the warning cigarette and green cigarette were viewed as much more unappealing than the standard cigarette, and have explained this in the Discussion: ‘We found that the two dissuasive cigarettes were perceived as significantly more harmful and less appealing than the standard cigarette, and less likely to encourage trial. The harm, appeal and trial items loaded onto a single ‘undesirability’ factor, with the dissuasive cigarettes considered

much more undesirable than the standard cigarette.' The Results section was not as clear as it should have been however, and we have revised this, as explained in response to comment 10 from Reviewer 2.

8) This work is apparently underpinned by any theory of marketing, consumer behavior, risk perceptions but reference to these theories would be useful to support what the authors expected or the rationale for this approach. It has been based on previous studies where there was some indication that changing or tainting the products altered its appeal but that is missing here. I recommend including a statement or two in the background and again in the discussion to acknowledge where assumptions that this study was based on.

Some discussion on tone of the insert content (focus being positively framed) and position in the pack (prominent and not able to be discarded until pack is completed?) or otherwise would also be useful.

Response: Although theories such as the Extended Parallel Process Model have been cited in past research on pack inserts (Thrasher et al., 2016; Moodie, 2018; Thrasher et al., in press), and theories on marketing and consumer behaviour are relevant to dissuasive cigarettes, this work is instead underpinned by the lessons that public health can learn from how tobacco companies communicate with consumers. At the start of the inserts section and also the dissuasive cigarettes section, we explain that while tobacco companies have used promotional inserts since the 19th century, and cigarettes have long been used as a marketing device, only in Canada are inserts promoting cessation required and regulators have overlooked the possibility of using cigarette appearance to deter smoking. Within the first paragraph, we now mention that while public health has typically overlooked how the inside of the cigarette pack could be used to deter smoking, inserts and cigarettes have been a mainstay of tobacco companies' promotional efforts. This reads: 'Regulators and academics have typically focused on the exterior of the cigarette pack, with little consideration of how the pack interior, for instance pack inserts or cigarettes, which have long been used by tobacco companies to promote their brands, could potentially be used to encourage smokers to think about their smoking behaviour.' At the start of the Discussion we now state: 'Just as tobacco companies have used inserts and cigarette design to create interest in their products, our study suggests that greater attention to how these could be used to promote cessation appears warranted.' Within the Introduction we did reference the first studies to explore perceptions of pack inserts promoting cessation, cigarettes displaying health warnings and unattractively coloured cigarettes.

With respect to the tone of the messages on inserts, we now make it more explicit that these are positively framed: 'These were replaced with eight new inserts, with coloured graphics and positively framed messages about the benefits of quitting or tips on how to do so, in 2012.' We discuss the position of inserts in packs in response to comment 6.

9) The practical element of this intervention of providing inserts needs to be acknowledged - a removable insert, that would be discarded would may have other consequences, both positive and negative (positive being the information I may be diffused through inadvertent exposure among others to the information) and negative being litter.

Response: In response to comment 6, we have included a sentence in the Introduction explaining that inserts are typically removed from packs and are thought to have an enduring impact as they would remain visible within the household or elsewhere, or as litter. We have not mentioned litter as a potential negative because inserts have been in packs in Canada for over 15 years, just as leaflets have been used in pharmaceutical packaging for decades, and the fact that they may become litter does not appear to be an issue. In addition, if inserts do encourage cessation, as evidence from this study and from Canada suggests, then this would lead to a reduction in the number of cigarettes smoked, with non-biodegradable cigarette butts a far greater issue where litter is concerned.

10) Finally, as the standard cigarette now occupies a space with e-cigarettes, pipes etc., the shift to having a green cigarette may have less impact or not be as disruptive. A comment acknowledging the changing or diversification of tobacco products may be warranted.

Response: We have not made a change here because even though the study was conducted in a country where e-cigarettes are very popular, the green cigarette was viewed as undesirable. We do however agree that future research with vapers would be of interest, and mention in the Discussion that 'further research with dual users, or indeed those also using vaping devices (not assessed in this study), would be fruitful.'

Reviewer 2
1) This is a well-designed study based on replies from young adult smokers on questions regarding a selection of cigarette pack inserts as well as three differently designed cigarette sticks. Inserts have been implemented in Canada, and more knowledge about how they are perceived in other countries are particularly timely. My main concern with the current manuscript regards the methods section. In general I found it difficult to read, and difficult to understand. I would encourage the authors to rephrase parts of this section for better clarity. Specifically, this concerns 'Design and sample' and the section about the multilevel log.reg.

Details
Methods
P6, Design and sample:
The fact that the sample consists of 16-34 year old smokers is mentioned twice. One time would suffice.

Response: The second mention of age has been replaced with 'this age group'. This now reads: 'An online survey was conducted in January-February 2016 with smokers aged 16-34 years old in the UK; an online survey is a suitable approach given that 99% of this age group in the UK are recent internet users.'

2) The bracket (n=193) would make more sense if it was moved to just after "completion time".

Response: Change made.

3) The size of the original sample is not mentioned explicitly here (although it is said in the abstract). It would be better to include it. A further explanation of the percentage 89.6 is warranted – it seems very high to be the response rate.

Response: We thank the reviewer for identifying this. We have now included the size of the original sample at the start of the 'Design and sample' section. The 89.6% is a reference to the completion rate rather than the response rate, although we should have made this clearer. We have removed this from the Abstract and more explicitly explain this in the text, which reads: '…the final sample was 1766 (89.6% of completed surveys).' We now also explain that we are unable to determine the response rate for the survey, as explained in response to comment 4 from Reviewer 1.

4) P 10, first paragraph and first line of analysis: Double check that the references 21 and 22 are correct. It looks like they may have been switched around.

Response: Thank you for spotting this, change made.

5) P11, middle: Could the authors explain the meaning of the 0.5 in this sentence: “All eight variables loaded on a single factor>0.5.”

Response: We now explain that the 0.5 relates to factor loadings. This now reads: ‘All eight variables loaded on a single factor with factor loadings that were >0.5.’

6) P11, multilevel logistic regression: I found the explanations of this analysis difficult to understand. The authors should try to reformulate the text. It would be good if you could include some references in support of the more technical details. I also do not understand what is meant by a “cross classified variable”, or that the reference category of the cross classified variable is varied.

Response: We have included a reference for readers interested in knowing more about the approach: ‘Multilevel logistic regression modelling, with second order PQL estimation,29 was undertaken …’ We have also rewritten the last part of this paragraph, providing more information about the cross classified variable and supporting references. This now reads: ‘Only one interaction was found, between cigarette type and SES. The interacting variables (cigarette type and SES) were substituted by a cross classified variable which merged cigarette type and SES. This cross classified variable was split into six categories: low SES standard cigarette, low SES warning cigarette, low SES green cigarette, not low SES standard cigarette, not low SES warning cigarette, not low SES green cigarette. To understand the interaction several models were run with the reference category of the cross classified variable different each time.30,31'

7) P15, perceptions of cigarette desirability: I suggest the authors include a table with these results instead of just listing them in the text.

Response: We have now included a new table (Table 4).

8) P16, Figure 3: There is no figure 3 – should it say table 4?

Response: Yes, it should have said Table 4.

9) P15/16, table 4: The authors should explain the table better. If the vertical lines represent confidence intervals, I do not understand how one can see any significant differences between groups, as is maintained in the text.

Response: We have removed this table because it has created some confusion, we already provide the findings within the Results section, and the same reviewer has requested another table on cigarette desirability (comment 7).

10) Results
P14, bottom: The text reads participants were more likely to consider SC attractive, stylish and nicer to be seen with, but %-ages were higher for WC and GC. Is this correct? Also the sentence (p15, top) “the SC (17.8%) was viewed as not as appealing to people their age as the WC (51.5%) or GC (57.4)” is quite confusing, and I do not understand what it means.

Response: The SC was more attractive, stylish and nicer to be seen with, as the reviewer suggests, but we accept that this was not clear because for each item lower percentages indicated more positive responses. We have now removed the percentages with respect to appeal and harm as we have taken the reviewers advice (comment 7) and included a table showing these figures. We have also made this section a bit more concise so that it will be easier to follow. This now reads: ‘With respect to harm, participants were less likely to think that the standard cigarette (SC) looked harmful than the warning cigarette (WC) or green cigarette (GC) ($p<0.001$), and less likely to think that the SC

made them think more about the dangers of smoking than the WC or GC ($p<0.001$). Participants were also less likely to indicate that the GC would make them think of the dangers of smoking than the WC ($p=0.01$). In terms of appeal, participants were more likely to consider the SC attractive, and stylish, than the WC or GC (both $p<0.001$). The SC was also considered to be nicer to be seen with, and more appealing to people their age, than the WC or GC (both $p<0.001$).'

11) P15, top: Include % in the bracket (57.4)

Response: We thank the reviewer for identifying this error. However, as noted in response to the previous comment, we no longer mention the percentages for Appeal and Harm within this section as we have included a table with these figures.

Reviewer 3
1) This is a meat and potatoes kind of data paper. The research question is broad and open, and not hypothesis driven.

Response: We accept the point made but with so few studies having explored cessation promoting inserts and dissuasive cigarettes, two measures that could realistically be implemented across the globe, we believe the paper will be of broad interest and spur further research in this area.

2) The authors dichotomize all of the data, so they can analyze with logistic regression. The original data are ordinal in nature, and I am curious why they did not analyze the data as ordinal. That was within the capabilities of the software at hand.

Response: We considered using ordinal regression but discounted it because the output coefficients are less meaningful than output coefficients from logistic regression, from which odds ratios can easily be calculated. Odds ratios are commonly used in papers and easy for readers to understand.

3) The authors used a piece of software called MLWin. Please capitalize the "L". I chased my tail for a while searching for info on MIWin. Looks like it does some things that SPSS cannot.

Response: Change made. MLWin is software designed for multilevel analysis and thus is often better than SPSS for such modelling.

4) The authors do not report unadjusted odds ratios. Model results should generally show both adjusted and unadjusted ORs side by side.

Response: We did not include unadjusted ORs in the relevant tables because this would have made them cluttered and more difficult to digest. Instead, given that the inclusion of unadjusted ORs is usually to allow readers to assess multicollinearity, we now explain in the Analysis section that 'The models were assessed for multicollinearity by comparison of the standard errors27 and none was found'.

5) Table 4 (figure 3?) shows an interaction. Authors should report the significance of the interaction somewhere in the caption or notes.

Response: As noted in response to comment 9 from Reviewer 2, we have removed Table 4 as it has created some confusion, the findings are already presented within the Results section, and Reviewer 2 has requested an additional table on cigarette desirability. We have added that the interaction was significant at the .05 level to the results: 'Only one significant interaction, between cigarette type and SES, was found ($p<0.05$).'

6) Interactions in logistic regression can be parameterized and reported as the ratio of odds ratios (ORR), with confidence interval.

Response: There is some discussion/dispute about the meaning of odds ratios output for interactions so we are wary of using them. Instead we have chosen to use a cross classified variable to understand the interaction. As explained in response to Reviewer 2 (comment 6), we now provide more detail on the cross classified variable in the Analysis section and reference past research that has used this approach.

7) Some of the language reporting results (odds ratios) in the text is a little clunky and monotonous. I found myself having to reread many of those because alarm bells were going off about sentence structure. They turn out to be fine, but I am guessing I will not be the only one with that response.

Response: That the text in the Results could perhaps be considered somewhat monotonous is, we feel, a consequence of the lengthy section on 'Smoking-related differences'. This is due to the many differences identified between the products used (factory-made cigarettes only vs factory-made cigarettes and other tobacco products), previous quit attempts (previous attempts vs no previous attempts) and likelihood of a successful quit attempt (likely to be successful vs unlikely to be successful or unlikely to make an attempt). These are important findings however and, as such, we have not made a change here, particularly as it helps ensure consistency in the reporting of ORs throughout the Results.

VERSION 2 – REVIEW

REVIEWER	Judith McCool University of Auckland, New Zealand
REVIEW RETURNED	18-Apr-2018

GENERAL COMMENTS	The changes or improvements made to the manuscript are well conducted and leave little further need for change. The only questions that remain in my mind are the potential negative impacts of inserts - what do they do with them after reading - if at all? Has there been any interest or concern by the tobacco industry in relation to this intervention? Are they really concerned that information inserts, in particular, are likely to be dissuasive more so than the text and graphic warnings on the outside of packs? In terms of the broader smoke-free environments and policies that are in place or planned, how would inserts support or be consistent with these broader approaches? I am not suggesting here that you need to include these considerations, but they came to mind when reviewing this paper. Overall, the authors have addressed the requested/suggested changes effectively and the result is a much improved manuscript.

REVIEWER	Ingeborg Lund Senior Researcher Norwegian Institute of Public Health Department of Alcohol, Tobacco and Drugs, Norway
REVIEW RETURNED	18-Apr-2018

GENERAL COMMENTS	The paper gives a good and generally well-written analysis of the effects and perceptions of pack inserts and dissuasive cigarettes in

	a group of young smokers recruited from a web-panel. This is an interesting and relevant area of research, and I do not have many comments. However, I found the methods section confusing, and not always obviously related to the results section. Methods/measures: Although it is stated that a count procedure is used to calculate SES, the regression in table 3 only uses education as a proxy for SES, while the analysis where the SES-measure seems to have been used is not reported in table. Why is this and what does the proxy add? Please explain or at least refer your choices in the analysis-section Methods/analysis: P12/top: the explanation of how the cigarette-variables are measured and operationalised should be moved to the measurement section. Also, the expression "percentages calculated for those indicating one of the three points nearest the undesirable anchor" is confusing, and it remains unclear to me if you have calculated one overall percentage for all these three points, or three separate percentages. Also, this seems to be poorly matched with the results reported in table 4, as they seem to be related to the top three points (although, the table is insufficiently explained, so I might be wrong). Please check that the measurement explanation is coherent with results in table 4, and explain the measurement more clearly. There are rather long and elaborate explanations of procedures related to a FA and a PQL multilevel regression, and yet none of the results from this seem to be reported in tables. Re FA: I do not think that there is a normality requirement for FA. Re QPL: Please explain more clearly what are random and what fixed effects, and why a multilevel regression was chosen. Are individuals clustered/nested within cigarette type? Also, unless I've misunderstood, you've used cigarette type as both a random and a fixed effect. Please explain why this is done, and how it impacts the interpretation of results. Report the results from PQL in a table. Some details: P6, 3. line: looks like a word is missing (…with it less likely…) P16, perceptions of cigarette desirability: Please insert an explanation here and/or in table 4 for the abbreviations SC, WC and GC. Figures are numbered wrongly (1 should be 2) Table 3 a & b: Although there is a footnote explaining that these do not represent separate regressions ((1) demographic and 2) tobacco related), the a) and b) is confusing. The table should be reorganized to fit into one page, f.ex. by removing all the lines representing the reference categories. The reference category can easily be named with the variable-heading (only needed when there are more than two categories). Table 4: Explain the meaning of the percentages (what does it mean that the answer is within the three highest agreement categories).

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

1) The changes or improvements made to the manuscript are well conducted and leave little further need for change. The only questions that remain in my mind are the potential negative impacts of

inserts - what do they do with them after reading - if at all? Has there been any interest or concern by the tobacco industry in relation to this intervention? Are they really concerned that information inserts, in particular, are likely to be dissuasive more so than the text and graphic warnings on the outside of packs? In terms of the broader smoke-free environments and policies that are in place or planned, how would inserts support or be consistent with these broader approaches? I am not suggesting here that you need to include these considerations, but they came to mind when reviewing this paper. Overall, the authors have addressed the requested/suggested changes effectively and the result is a much improved manuscript.

Response: We thank the reviewer for their response to our changes. As the reviewer is not suggesting that we include these considerations, and given the length of the paper and the fact that we did not specifically test any of these considerations, we have not made any revisions in the Introduction or Discussion. There is clearly much more work to do on this topic, and as mentioned in response to one of the reviewers in our previous set of comments, we believe that the paper will stimulate further research and discussion.

Reviewer 2

1) The paper gives a good and generally well-written analysis of the effects and perceptions of pack inserts and dissuasive cigarettes in a group of young smokers recruited from a web-panel. This is an interesting and relevant area of research, and I do not have many comments. However, I found the methods section confusing, and not always obviously related to the results section.

Methods/measures:

Although it is stated that a count procedure is used to calculate SES, the regression in table 3 only uses education as a proxy for SES, while the analysis where the SES-measure seems to have been used is not reported in table. Why is this and what does the proxy add? Please explain or at least refer your choices in the analysis-section

Response: We have now clarified this within the ‘Measures’ section. This reads: ‘Preliminary analysis showed that education was associated with how pack inserts were perceived, whereas both education and economic status were associated with how cigarettes were perceived. As such, for the analysis of the cigarettes a count procedure was used to create a variable for low socioeconomic status…’ We used a composite variable for parsimony reasons and to reduce the risk of multicollinearity.

2) Methods/analysis:

P12/top: the explanation of how the cigarette-variables are measured and operationalised should be moved to the measurement section.

Response: The calculation of percentages was part of the analysis rather than variable derivation so we have not moved this. However, to make this clearer we have inserted a paragraph break after this section.

3) Also, the expression “percentages calculated for those indicating one of the three points nearest the undesirable anchor” is confusing, and it remains unclear to me if you have calculated one overall percentage for all these three points, or three separate percentages. Also, this seems to be poorly matched with the results reported in table 4, as they seem to be related to the top three points (although, the table is insufficiently explained, so I might be wrong). Please check that the measurement explanation is coherent with results in table 4, and explain the measurement more clearly.

Response: We have clarified this by explaining that a total of 24 percentages were calculated, one percentage for each of the eight items and each of the three cigarette types. This reads: ‘For each of the eight seven-point semantic scales, the percentage of participants choosing one of the three points nearest the undesirable anchor (e.g. unattractive, not nice to be seen with, looks harmful to health) was calculated for each of the three cigarette types (SC, WC, GC). Thus, 24 percentages were calculated.’

We have also changed the footnotes to Table 4 so that this matches the text in the main document: ‘1 Percentages shown indicate participants choosing one of the three points nearest the undesirable anchor on a seven-point semantic scale.’ Additionally, we have revised the direction of the trial scales in Table 4 to be consistent with the other scales.

4) There are rather long and elaborate explanations of procedures related to a FA and a PQL multilevel regression, and yet none of the results from this seem to be reported in tables.

Response: The factor analysis was to derive the outcome variable in the regression so is best considered part of the variable derivation. As such, we have moved this to the ‘Measures’ section. As it was part of variable derivation we do not feel it warrants a table, but we have summarised the multilevel, multivariable results in a new table (Table 5), as the reviewer subsequently recommends (comment 8).

5) Re FA: I do not think that there is a normality requirement for FA.

Response: The normality requirement arises because the factor was used as the outcome of a regression analysis. We have now made this clearer in the Measures by stating that ‘The factor was used as the outcome measure of cigarette desirability in the regression analysis.’

6) Re QPL: Please explain more clearly what are random and what fixed effects, and why a multilevel regression was chosen. Are individuals clustered/nested within cigarette type?

Response: The random effect was individual participants. The cases were opinions of cigarettes. As opinions of cigarettes were clustered within individual participants this required a multilevel model. The fixed effects were cigarette type and sociodemographic and smoking-related characteristics and we have clarified this in the text: 'Multilevel logistic regression modelling of cigarette desirability, with second order PQL estimation,29 was undertaken with cigarette evaluations (participants' response to each of the three cigarettes) clustered within individual participants. Therefore, cigarette evaluations were level one cases and participants were entered at level two as a random effect. All models included cigarette type as a fixed effect, where the standard cigarette was compared with the warning cigarette and green cigarette. Other fixed effects at the individual (participant) level were sociodemographic and smoking-related characteristics, which were significantly associated with the outcome in multivariable models.'

7) Also, unless I've misunderstood, you've used cigarette type as both a random and a fixed effect. Please explain why this is done, and how it impacts the interpretation of results.

Response: Each case within the analysis is a cigarette evaluation and each respondent gave three evaluations, thus the random effect is the participant. The classification of each cigarette evaluation as SC, WC or GC is a fixed effect. As explained in response to comment 6, we have clarified this in the text.

8) Report the results from PQL in a table.

Response: We now report the multilevel results in Table 5.

9) Some details:

P6, 3. line: looks like a word is missing (…with it less likely…)

Response: No, the sentence is as intended, so we have not made a change here.

10) P16, perceptions of cigarette desirability: Please insert an explanation here and/or in table 4 for the abbreviations SC, WC and GC.

Response: These acronyms are now introduced in the second paragraph of the Procedure, when describing the cigarette types. We have also added the acronyms to Table 4.

11) Figures are numbered wrongly (1 should be 2)

Response: Thank you for spotting this, the Figures are now correctly numbered.

12) Table 3 a & b: Although there is a footnote explaining that these do not represent separate regressions ((1) demographic and 2) tobacco related), the a) and b) is confusing. The table should be reorganized to fit into one page, f.ex. by removing all the lines representing the reference categories. The reference category can easily be named with the variable-heading (only needed when there are more than two categories).

Response: Thank you for this suggestion, we have merged Tables 3a and 3b into a single table (Table 3).

13) Table 4: Explain the meaning of the percentages (what does it mean that the answer is within the three highest agreement categories).

Response: We have made this clearer by using the same language in the footnote that is used in the text. As noted in response to comment 3, the footnote now reads ‘1 Percentages shown indicate participants choosing one of the three points nearest the undesirable anchor on a seven-point semantic scale.’

VERSION 3 – REVIEW

REVIEWER	Ingeborg Lund Norwegian Institute of Public Health
REVIEW RETURNED	06-Jul-2018

GENERAL COMMENTS	The paper has improved after the first review. I have no further comments