PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Accounting for health literacy and intervention preferences when reducing unhealthy snacking: Protocol for an online randomized controlled trial
AUTHORS	Ayre, Julie; Cvejic, Erin; Bonner, Carissa; Turner, Robin; Walter, Stephen; McCaffery, Kirsten

VERSION 1 - REVIEW

REVIEWER	Jennifer Inauen, PhD, Assistant Professor of Health Psychology Department of Psychology, University of Bern, Switzerland
REVIEW RETURNED	23-Jan-2019

GENERAL COMMENTS	Summary: This trial aims to test whether a health-literacy sensitive action plan (vs. a standard action plan) can reduce unhealthy snacking among patients with diabetes type 2. Additional manipulations include mode of allocation of plan type, and assessing plan preference. This protocol is generally well written, clear and concise. However, the trial design is complex, and the methodology therefore needs more detail. I highlighted some specific issues that should be clarified. 1. Abstract: The term health literacy might not be clear to all readers. A brief explanation would be helpful. Also, important methodological information is missing from the abstract, e.g. how the primary outcome was measured. 2. Electronic diaries of snack consumption might have helped overcome some limitations of the one-time self-reported consumption (i.e. memory bias). This could be mentioned in the discussion if the trial has already started. 3. Has this trial already started? What are the dates of the (expected) trial run and each respective measurement occasion? 4. There are many terms in the introduction (p. 4) that need explaining and definition, e.g. health literacy. 5. p. 5, lines 43-50: I had to read this sentence multiple times. By action plans, do the authors mean a literacy-sensitive vs. a standard action plan? Are health literacy and action plan selection randomized (in addition to allocation method)? Please clarify this sentence.

	6. p. 5, line 45: What is the Rucker protocol? Can the authors provide more detail (here or on p. 10)? 7. p. 5, lines 53-55: Is it the *assessment* of the preference, or the *preference* that authors hypothesize to influence intervention effectiveness? Please clarify. In the former case, I would expect assessment of preference to be randomized (this becomes more clear later, but here it is not). 8. p. 6, lines 22-24: The second part of this hypothesis seems somewhat related to the literature presented above. However, a rationale for why *assessing* preference might negatively impact plan effectiveness was not presented. Is it because the discrepancy between plan and preference is made salient through the assessment? 9. p. 6, line 34: Is it really a parallel trial? From the hypothesis, I would have expected a 2 x 3 x 2 factorial trial: plan (literacy-sensitive vs. standard) x allocation method (random vs. literacy-tailored vs. preference based) x assessment of preference (assessed vs. not assessed). In Figure 1, this is suggested, although it becomes apparent that assessment is only manipulated in one of the method arms. Overall, the trial design is very complex, and needs to be described more carefully, and in detail in this section. 10. p. 7, lines 25-30: Informed consent procedures need clarification. The sentence can be misread such that consent will be implied by the fact that participants filled in the online survey. I imagine, the authors rather mean that participants indicate informed consent after reading the information sheet. 11. p. 10, line 24: Please provide a reference for the Newest Vital Signal Measure. Also, how was this cut-off defined? 12. p. 12: The literacy-sensitive action plan contains more elements than planning (e.g. thinking of reasons why one snacks, judging the ease of the plan). These might be alternative mechanisms of action. Can authors make this explicit by adding the respective behavior change techniques from the behavior change technique taxonomy (Michie et al., 2013)? 13. The standard action plan contains additional elements that might change behavior more effectively than the sensitive plan, i.e. providing information that planning is effective. Furthermore, this planning task seems very much harder to complete compared to the literacy-sensitive plan. This might cause unwanted differences in plan effectiveness as research has shown that people are not used to making plans and have difficulty to do so without guidance. 14. p. 13, line 39: Can authors provide the definition of unhealthy snacks? 15. p. 13, lines 42-58: Why are intentions, habits, and action control (of all social-cognitive predictors of snacking) assessed? A rationale and theory in the introduction is lacking.

	16. p. 13: The intervention also includes goal reminders, which could enhance plan effectiveness. Please add this as a behavior change technique. 17. p. 14: Why is need for cognition assessed? Lacks rationale. Most importantly, what is the primary outcome, and what are secondary outcomes? 18. p. 16, line 15: How was the a priori effect size determined? 19. p. 17, line 13: What was the basis of the decision to analyze completers only? Methodological research on missings suggests that they are usually not random, wherefore imputation of missing data is recommended to avoid bias (e.g. by full-information maximum likelihood, or multiple imputation). 20. p. 18, lines 24-27: I am unsure how the authors plan to do the subanalysis of persons with vs. without diabetes as diabetes type 2 was specified an inclusion criterion (p. 6, line 56). 21. p. 18, line 31: It is commendable that authors are coding the action plans as a measure of intervention fidelity. It would be a strength if this measure could be included in the effects regression models, as a sensitivity analysis.

REVIEWER	Heather D. Gibbs Assistant Professor University of Kansas Medical Center United States
REVIEW RETURNED	05-Feb-2019

GENERAL COMMENTS	The article presents the protocol for a three parallel-arm online RCT to examine the effect of health literacy, the approach used in delivering an action plan, and the method use to select an action plan upon self-reported unhealthy snacking behavior. The novelty of the study is in its design to challenge the notion that all individuals do best with a 'universal precautions' approach to health-related messaging and intervention approaches. While overall, the protocol appears unique and valuable to the body of health literacy literature, details and justification within the methods requires further attention. Most importantly, a definition of 'unhealthy snacking' must be presented within the introduction. It becomes more apparent as one reads further into the protocol, that the emphasis is placed upon behavior surrounding snacking and not upon the nutritional considerations of snacking. Authors seem to assume that snacking is inherently bad behavior, and this is not well justified. Related concerns include: 1. Please justify why participants will not be excluded on the basis of snacking behavior. How is the intervention relevant for those who do not snack? Isn't there potential to increase snacking in participants who do not snack? (p.7, line 19) 2. The example literacy-sensitive action plan (p. 11, lines 11-48) suggests that the target of the tool is to influence motivation for snacking/not snacking, correct? The example standard action plan (p.12, line 18) indicates focus upon "situations when you are not hungry but find yourself snacking." It does not appear the target is shifting to healthy snacks. Please make the focus of the intervention approaches clear.

	3. Given their known impact upon public health, why are sugar sweetened beverages and alcohol excluded? (p. 13, line 52-55) Why are beverages not viewed as snacks when they contribute significant empty calories for a large portion of the population worldwide? 4. The snack measures lack precision. As the primary aim, it should be clear how snacking behavior will be measured with precision. Will your score be derived from the CSIRO measure, perceived snacking behavior, some combination? a. The CSIRO measures diet quality over a one month period, but it does not provide an assessment of quality per eating occasion (i.e. snacks). It appears this approach is intended to capture intake of 'discretionary foods' only. How will you know if people are using unhealthy behaviors with 'healthy' foods (ie. foods not considered discretionary)? For example, overeating fruit can raise blood glucose in individuals with diabetes. Further, this measures intake over a 1 month period, and the intervention is only 4 weeks. Is this long enough to affect and detect change? Lack of follow-up at 6 months means maintenance of any change in behavior will be unknown. b. The previous day measure (26 items including 'healthy' and 'unhealthy' items) is not provided and little description of its content is provided in literature. It does not appear to have been validated. Other minor concerns include: 5. Please explain Hypothesis 3 further (p. 6, lines 22-24). Does this mean that the actual process of asking people to note their preference will decrease the intervention's effectiveness? Is this based on an assumption that people do not want to change snacking behavior? 6. Collecting height and weight by self-report is an important limitation. 7. How have the materials used in the 'Choice' arm been tested in individuals with low health literacy? Both descriptions of intervention tools should be correctly interpreted by individuals with low health literacy. (p. 10, lines 36-52). Will this be part of the piloting process? 8. Please provide a brief description of the phrase 'need for cognition.' (p. 13, lines 34-39)

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1
Summary: This trial aims to test whether a health-literacy sensitive action plan (vs. a standard action plan) can reduce unhealthy snacking among patients with diabetes type 2. Additional manipulations include mode of allocation of plan type, and assessing plan preference. This protocol is generally well written, clear and concise. However, the trial design is complex, and the methodology therefore needs more detail. I highlighted some specific issues that should be clarified.

1. Abstract: The term health literacy might not be clear to all readers. A brief explanation would be helpful. Also, important methodological information is missing from the abstract, e.g. how the primary outcome was measured.

Response: we have added the following additional text to the opening paragraph of the abstract: "Health literacy describes the cognitive and social skills that enable individuals to access, understand and act on health information." We have added a sentence to the abstract to indicate that the primary outcome is "self-reported serves of unhealthy snacks during the previous month".

2. Electronic diaries of snack consumption might have helped overcome some limitations of the one-time self-reported consumption (i.e. memory bias). This could be mentioned in the discussion if the trial has already started.

Response: The trial commenced in February after receiving reviewer comments. We were unable to incorporate electronic diaries of snack consumption as this is outside the technical and financial capabilities of the study. We were also concerned about the greater risk of missing or incomplete data due to low completion of snack diaries. We have mentioned this limitation in the methods section with the following text:

"Although electronic diaries might overcome some limitations of once-off measures of snacking behaviour such as that described above, electronic diaries were unavailable at the time of study development due to the budget and technical constraints. This approach chosen also reduces participant burden and minimises the risk of missing data."

We have also amended the strengths and limitations box to incorporate this limitation: "This study uses a subjective outcome measure (self-reported monthly unhealthy snacking collected at a single time-point) rather than an objective measure (e.g. unhealthy snacking observed throughout the month-long period) or more frequently reported self-report measure, which may limit the study findings."

We will also highlight this limitation in the discussion section of the planned results paper.

3. Has this trial already started? What are the dates of the (expected) trial run and each respective measurement occasion?

Response: This trial commenced in mid-February, after receiving peer review comments. The dates are also available from the ANZCTR registry entry. We have also added text to state that this trial commenced on February 14th, 2019 (p7)

4. There are many terms in the introduction (p. 4) that need explaining and definition, e.g. health literacy.

Response: Definitions of health literacy have been added to the introduction as follows: Health literacy: Health literacy describes the cognitive and social skills that assist individuals to access, understand and act on health information.

Engagement: This was revised to clarify: "For example, participants would be able to factor in other relevant aspects (such as motivation to engage with the planning process)"

5. p. 5, lines 43-50: I had to read this sentence multiple times. By action plans, do the authors mean a literacy-sensitive vs. a standard action plan? Are health literacy and action plan selection randomized (in addition to allocation method)? Please clarify this sentence.

Response: We have tried to make the meaning less ambiguous and easier to read with the following phrasing:

"The second aim is to evaluate how the method of allocation to either intervention (literacysensitive or standard action plan) affects the overall effectiveness of the intervention. Three methods of allocation will be evaluated: 1) random allocation to an action plan; 2) allocation to an action plan based on individual health literacy; or 3) allowing participants to choose which action plan they use."

6. p. 5, line 45: What is the Rucker protocol? Can the authors provide more detail (here or on p. 10)?

Response: The Rucker design involves two-stage randomization so that the effects of choice can be isolated from the effects of individual treatments (Rücker G. A two-stage trial design for testing treatment, self-selection and treatment preference effects. Statistics in Medicine. 1989;8(4):477-85). In that sense its definition is already provided in the introduction. However, we have separated the sentence to better distinguish between the Rucker design and Rucker analysis:

“This study will employ a two-stage randomization (Rucker) design (16). This design allows estimation of the effects of each action plan on outcomes (the treatment effect), independent of the effects of a person receiving their preferred treatment (preference effect) and the effects of self-selection (selection effect). This is the only preference trial design that allows this delineation of all these effects.”

We have also added an additional reference in the analysis section (under ‘Analysis of effects of allocation method’) to emphasise why the two-stage randomisation design is important for this study (i.e. that it allows us to conduct an analysis to isolate the impacts of treatment, preference and selection effects on the primary outcome). This reference is 32: McCaffery KJ, Turner R, Macaskill P, Walter SD, Chan SF, Irwig L. Determining the Impact of Informed Choice: Separating Treatment Effects from the Effects of Choice and Selection in Randomized Trials. Medical Decision Making. 2011;31(2):229-36.

7. p. 5, lines 53-55: Is it the *assessment* of the preference, or the *preference* that authors hypothesize to influence intervention effectiveness? Please clarify. In the former case, I would expect assessment of preference to be randomized (this becomes more clear later, but here it is not).

Both are correct. Aim 2/hypothesis 2 includes evaluation of whether preference influences intervention effectiveness. This is indicated by the sentence: ‘Both of these allocation methods [health literacy screening tool or participant preference] will be more effective than random allocation to an action plan.’ Using the Rucker design, we will be able to numerically isolate the effects of a participant having the ability to select their preferred action plan from the effects of the action plan itself.

This aim led to a third research question: does asking a participant about their preference for an action plan impact on the action plan’s effectiveness (if they are then randomised to an action plan and may therefore potentially do not received their preferred option)? As such, aim/hypothesis 3 refers to assessment of preference influencing intervention effectiveness. We have added text to aim 3 to describe the gap in the literature (and the rationale) for measuring this:
“This addresses an unanswered question in research on the effects of treatment preference i.e., does assessment of preference introduce a methodological artefact by increasing the salience of discrepancies between an individual’s preferred and received treatments, and in doing so, have a negative effect on intervention outcomes?”

8. p. 6, lines 22-24: The second part of this hypothesis seems somewhat related to the literature presented above. However, a rationale for why *assessing* preference might negatively impact plan effectiveness was not presented. Is it because the discrepancy between plan and preference is made salient through the assessment?
Response: Thank you for bringing this to our attention. This comment has been addressed through the following statement (as above):
“The third aim is to evaluate whether assessment of participant preference for an intervention prior to random allocation influences the effectiveness of the intervention. This addresses an unanswered question in research on the effects of treatment preference on study outcomes i.e., does assessment of preference introduce a methodological artefact by increasing the salience of discrepancies between an individual’s preferred and received treatments, and in doing so, have a negative effect on intervention outcomes?”

9. p. 6, line 34: Is it really a parallel trial? From the hypothesis, I would have expected a 2 x 3 x 2 factorial trial: plan (literacy-sensitive vs. standard) x allocation method (random vs. literacy-tailored vs. preference based) x assessment of preference (assessed vs. not assessed). In Figure 1, this is suggested, although it becomes apparent that assessment is only manipulated in one of the method arms. Overall, the trial design is very complex, and needs to be described more carefully, and in detail in this section.

Response: We agree that the trial design is complex, but feel this complexity is necessary to achieve our specified aims in the most efficient manner.

The design you describe (2 x 3 x 2) can be visualised as below (Table 1). With this hypothetical design, in the choice arm, if preference were assessed, there would be no data on the effects of discrepancies between an individual's preferred and received treatments, which is the focus of that research question. Similarly, if we assessed preference in the screened arm, this would answer a different research question (does a discrepancy between preferred and received plan impact of effectiveness of an action plan when that discrepancy is based on a health literacy score). By only assessing preference in the random arm, our proposed design allows for a smaller sample size that can answer our specific research questions. This trial could be better thought of as a 4 x 2 design, where 2 of the 4 arms are similar (random, preference assessed; and random, preference not assessed). Figure 1 has been revised to clarify where randomisation occurs (these are now indicated by ovals) and uses different colours to indicate the four main arms of the trial.

Table 1: Parallel trial (2 x 3 x 2)

	Allocation method		
	Random	Screened (literacytailored)	Choice (preferencebased)
Preference assessed	Literacy sensitive action plan OR standard action plan	Literacy sensitive action plan OR standard action plan	Literacy sensitive action plan OR standard action plan
Preference not assessed	Literacy sensitive action plan OR standard action plan	Literacy sensitive action plan OR standard action plan	Literacy sensitive action plan OR standard action plan

10. p. 7, lines 25-30: Informed consent procedures need clarification. The sentence can be misread such that consent will be implied by the fact that participants filled in the online survey. I imagine, the authors rather mean that participants indicate informed consent after reading the information sheet.

Response: In line with the Australian National Statement on Ethical Conduct in Human Research, 2007 (updated 2018), under statement 2.2.5, "Consent may be expressed orally, in writing, or by some other means (for example, return of a survey, or conduct implying consent)…". All participants are invited to read the participant information sheet, and indicate their consent by successful submission of a completed survey. On the landing page for the survey we also reiterate in summary form what participants will be asked to do if they decide to take part in the study.

11. p. 10, line 24: Please provide a reference for the Newest Vital Signal Measure. Also, how was this cut-off defined?

Response: The reference for the Newest Vital Sign is listed under 'Measures' > 'Health literacy.' As per the reference, these cut-offs correspond to performance on the Test of Functional Health Literacy in Adults (TOFHLA), which is the gold-standard performance-based measure of health literacy. This reference has also been added to where it is first mentioned under 'Participant allocation' > 'Allocation method' > Screened (Arm B)

12. p. 12: The literacy-sensitive action plan contains more elements than planning (e.g. thinking of reasons why one snacks, judging the ease of the plan). These might be alternative mechanisms of action. Can authors make this explicit by adding the respective behavior change techniques from the behavior change technique taxonomy (Michie et al., 2013)?

Response: We have included an additional table as Supplementary Material which lists the behaviour change techniques present in these interventions. This is referenced on page 12 of the manuscript:

Supplementary Table 1: Behaviour Change Techniques present in intervention

Intervention	Intervention feature	Behaviour change technique
Literacy-sensitive	The text 'forming plans has shown to improve snacking habits'	Credible source
	Identifying situations for unhealthy snacking	Problem solving
	Identifying an alternative behaviour to enact in snacking situation	Behaviour substitution
	Generation of plan (with images) to reduce unhealthy snacking	Action planning
	Instruction to imagine enacting the plan	Mental rehearsal of a successful performance
	Reminder emails	Prompts/cues
Standard	The text 'forming plans has shown to improve snacking habits'	Credible source
	Identifying situations for unhealthy snacking	Problem solving
	Generation of plan to reduce unhealthy snacking	Action planning
	Reminder emails	Prompts/cues

Note: Behaviour Change Techniques are based on the Behaviour Change Technique Taxonomy v1 (Michie et al., 2013)

Michie, S., Wood, C. E., Johnston, M., Abraham, C., Richardson, M., Francis, J., . . . Cane, J. (2013). The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. Annals of Behavioral Medicine, 46(1), 81-95. doi:10.1007/s12160-0139486-6.

13. The standard action plan contains additional elements that might change behavior more effectively than the sensitive plan, i.e. providing information that planning is effective. Furthermore, this planning task seems very much harder to complete compared to the literacysensitive plan. This might cause unwanted differences in plan effectiveness as research has shown that people are not used to making plans and have difficulty to do so without guidance.

Response: To maintain consistency, we have now added the statement about 'planning being effective' to the literacy-sensitive plan (and see it added as a features of the literacy-sensitive plan in Table 2 above). You can see this in the text described the intervention in the methods section.

Until recently, planning interventions have often been free-text and so are inherently more difficult (for example, SMART goals ask people to come up with their own plans, and implementation intention instructions also typically ask people to come develop text-based plans.) In our previous study (Ayre J, Bonner C, Cvejic E, McCaffery K. Randomized trial of planning tools to reduce unhealthy snacking: implications for health literacy. PloS one. 2019; 14(1), e0209863.), we were interested in developing an online planning tool that could be used without a coach or expert guidance, and that would be suitable for people with lower health literacy. Our results indicated that this approach (the literacy-sensitive plan) was more effective for people with lower health literacy. In contrast, the standard plan (i.e. the 'harder' of the two) was actually more effective for people with higher health literacy than the literacy-sensitive plan. As such, previous research has shown different groups of people may stand to benefit more or less from each of the plans (literacy-sensitive or standard).

14. p. 13, line 39: Can authors provide the definition of unhealthy snacks?

Response: The text has been changed to reflect the discretionary foods category described under 'snack scores (previous month)':

Perceived unhealthy/healthy snacking (previous week): Two items, each on a 7-point Likert scale will assess perceived extent of healthy and unhealthy snacking in the previous week, respectively. Healthy snacks are described to participants as those that are low in kilojoules, fat, salt and sugars. Unhealthy snacks are described as high in kilojoules, fat, salt and sugars. Examples of healthy and unhealthy snacks are provided.

Snacks consumed (previous day): Participants will be asked to select from among 26 items the snacks that they had consumed in the previous day. This list includes 'discretionary foods' and nondiscretionary foods as defined by (20), that can be consumed as snacks. This is based on previous work (21).

15. p. 13, lines 42-58: Why are intentions, habits, and action control (of all social-cognitive predictors of snacking) assessed? A rationale and theory in the introduction is lacking.

Response: The focus of this study is the moderating effect of health literacy on the effectiveness of action plans, and how this finding can be incorporated into interventions (e.g. through screening for health literacy and then tailoring the intervention that the participant receives). This emphasis is also reflected in our aims (which do not refer explicitly to social-cognitive predictors). In view of the specific angle we have taken, we have limited the scope of the introduction to this research question. We have now included a table in the Supplementary Material which maps the behaviour change techniques for those who are interested.

We have also deliberately included a measure of intention and habit strength as these are known to be important influences on the effectiveness of implementation intentions (see Adriaanse, M. A. and A. Verhoeven (2018). Breaking Habits Using Implementation Intentions. The Psychology of Habit: Theory, Mechanisms, Change, and Contexts. B. Verplanken. Cham, Springer International Publishing: 169-188.). As such, it will be important to see if baseline intention or habit strength differ across groups.

Action control was included as some evidence suggests that implementation intentions improve self-regulatory processes. For example, a previous study that employed a similar kind of guided implementation intention and a standard implementation intention reported that both increased awareness of standards, self-monitoring and self-regulatory effort (Armitage, C. J. (2015). "Field experiment of a very brief worksite intervention to improve nutrition among health care workers." Journal of Behavioral Medicine 38(4): 599-608.). Lastly, for consistency we have collected data on these variables at baseline and follow-up, and it would be interesting to explore changes in action control and habit strength as these would provide insight into possible mechanisms of action. We will conduct exploratory analyses of the effects of the action plans on action control and habit strength, and their potential mediation of the effect of the action plans on snacking outcomes. However, it should be noted that this study is not powered to test these effects.

Other social-cognitive predictors were not included as we felt it was important to minimise participant burden and survey fatigue.

16. p. 13: The intervention also includes goal reminders, which could enhance plan effectiveness. Please add this as a behavior change technique.

This is a helpful observation. As 'goal reminders' is not listed in the BCT Taxonomy (v1), we have listed this as 'prompts/cues' in the aforementioned Supplementary Table.

17. p. 14: Why is need for cognition assessed? Lacks rationale. Most importantly, what is the primary outcome, and what are secondary outcomes?

Response: We have added 'primary outcome' and 'secondary outcomes' headings to the 'Measures' section to make this clearer. We are interested in need for cognition as another potential moderator of action plan effectiveness (and as such, in future studies could be used as a screening tool for allocating people to an action plan). Therefore this was included for transparency, but is not the main focus of this study.

18. p. 16, line 15: How was the a priori effect size determined?

Response: This effect size was based on our previous study in which we investigated the impact of very similar versions of both interventions on snacking behaviour, using the same outcome measure, and same market research company for recruitment (though inclusion criteria were Australians aged over 30 years in that study). This reference is now included with the additional text:
"Estimates of effect size are based on the outcome of previous work exploring the effects of action plans on unhealthy snacking (11)."

19. p. 17, line 13: What was the basis of the decision to analyze completers only? Methodological research on missings suggests that they are usually not random, wherefore imputation of missing data is recommended to avoid bias (e.g. by full-information maximum likelihood, or multiple imputation).

Response: We felt that imputation of missing data would not be appropriate in this instance given that we would be missing data for the outcome measure. At baseline, each question must be answered before progressing to the next one, so there is no missing data at baseline. Instead, we propose an intention-to-treat analysis in which outcome measures for those who have been lost to follow-up (and therefore do not provide follow-up data) will be estimated in a worst-case (no change in snacking score) and best-case scenario (average change in snacking score). Given that each question at baseline and follow-up had to be completed before continuing onto the next question, there should be no missing data. As such we anticipate that the range given by best- and worst-case scenario estimates will not be so wide as to be unhelpful. This approach will be incorporated into the methods as a sensitivity analysis. We have added text to describe this approach under the analysis heading:

"Follow-up outcome measures for participants who do not complete the follow-up survey will be estimated for worst-case and best-case scenarios (calculated as no-change and average-change in snack scores, respectively). These estimates will be incorporated as an intention-to-treat analysis. Baseline characteristics of completers and non-completers will be compared to assess bias and generalisability."

20. p. 18, lines 24-27: I am unsure how the authors plan to do the subanalysis of persons with vs. without diabetes as diabetes type 2 was specified an inclusion criterion (p. 6, line 56).

Response: the inclusion criteria specify that participants must "self-report that they have type 2 diabetes or self-report a height and weight that correspond to overweight or obese BMI." As such, not all participants are expected to have type 2 diabetes.

21. p. 18, line 31: It is commendable that authors are coding the action plans as a measure of intervention fidelity. It would be a strength if this measure could be included in the effects regression models, as a sensitivity analysis.

Response: In the previous study (Ayre et al., PlOS One, 2019) we analysed plans using content analysis. We intend to use the same methods in this study (described under 'additional analysis') on p19. As suggested, we will aim to incorporate this into a secondary analysis of the data. The following text has been added to the paragraph on p19:
"Results from this content analysis will also inform a secondary analysis of the impact of allocation method and action plan on snacking scores."

Reviewer: 2
Summary: The article presents the protocol for a three parallel-arm online RCT to examine the effect of health literacy, the approach used in delivering an action plan, and the method use to select an action plan upon self-reported unhealthy snacking behavior. The novelty of the study is in its design to challenge the notion that all individuals do best with a 'universal precautions' approach to health-related messaging and intervention approaches. While overall, the protocol appears unique and valuable to the body of health literacy literature, details and justification within the methods requires further attention.

Most importantly, a definition of 'unhealthy snacking' must be presented within the introduction. It becomes more apparent as one reads further into the protocol, that the emphasis is placed upon behavior surrounding snacking and not upon the nutritional considerations of snacking. Authors seem to assume that snacking is inherently bad behavior, and this is not well justified. Related concerns include:

1. Please justify why participants will not be excluded on the basis of snacking behavior. How is the intervention relevant for those who do not snack? Isn't there potential to increase snacking in participants who do not snack? (p.7, line 19)

Response: Unhealthy snacking is reasonably common, and in our previous study (Ayre et al, PloS One 2019) we found that people were motivated to improve their snacking behaviour, scoring on average 5.1 on a Likert scale for intention to change unhealthy snacking behaviour (1 = low intention, 7= high intention) (76% of participants reported positive intentions to change their unhealthy snacking behaviour).

2. The example literacy-sensitive action plan (p. 11, lines 11-48) suggests that the target of the tool is to influence motivation for snacking/not snacking, correct? The example standard action plan (p.12, line 18) indicates focus upon "situations when you are not hungry but find yourself snacking." It does not appear the target is shifting to healthy snacks. Please make the focus of the intervention approaches clear.

Response: The target of both interventions is to reduce unhealthy snacking. Both interventions invite the participant to create implementation intentions which focus on situations that act as cues for unhealthy snacking. By focusing on these cues, implementation intentions mostly target the action phase of behaviour change rather than the motivation phase. Plans from either version of the intervention could include strategies such as distraction (e.g. going for a walk), reduction (e.g. eating less of an unhealthy snack) or substitution (e.g. eating fruit or drinking a glass of water). In Figure 2 (the screenshot of the literacy-sensitive plan) you can see that there are options that are distraction techniques as well as substitution.

3. Given their known impact upon public health, why are sugar sweetened beverages and alcohol excluded? (p. 13, line 52-55) Why are beverages not viewed as snacks when they contribute significant empty calories for a large portion of the population worldwide?

Response: Upon reflection we have now included sugar sweetened beverages in the CSIRO diet measure. Whilst we also accept that alcohol can contribute towards snacking we feel that the factors that influence alcohol intake are different to those that influence snacking on unhealthy foods. We believe that an action plan intervention for reduction in alcohol intake could be better explored in a separate study.

4. The snack measures lack precision. As the primary aim, it should be clear how snacking behavior will be measured with precision. Will your score be derived from the CSIRO measure, perceived snacking behavior, some combination?

Response: We have clarified this by separating measure using the headings 'primary outcome' and 'secondary outcome measures.' The primary outcome is the CSIRO measure (to be reported as average serves of unhealthy snacks per week).

   a. The CSIRO measures diet quality over a one month period, but it does not provide an assessment of quality per eating occasion (i.e. snacks). It appears this approach is intended to capture intake of 'discretionary foods' only. How will you know if people are using unhealthy behaviors with 'healthy' foods (ie. foods not considered discretionary)? For example, overeating fruit can raise blood glucose in individuals with diabetes. Further, this measures intake over a 1 month period, and the intervention is only 4 weeks. Is this long enough to affect and detect change? Lack of follow-up at 6 months means maintenance of any change in behavior will be unknown.

Response: You are correct in identifying that the CSIRO measure does not capture the quality per eating occasion. If people select a plan that mentions eating a piece of fruit, they receive additional information about how much and which fruits to eat as a snack. This information was drawn from the Diabetes Australia / National Diabetes Services Scheme fact sheet on healthy snacks. It is true that it is still possible that participants may eat too much fruit and this is a limitation of the study. However, we would argue that this is still preferable to participants eating the unhealthy snack ('discretionary foods').

With regards to the comment about 1 month change and a 4 week intervention, in our previous study we were able to affect and detect change using this measure(Ayre et al, PloS One 2019.). Nevertheless, this is a useful comment and we will mention the lack of long-term follow-up in the limitations of the results paper. We would be very interested in following-up participants at 6 months if we could, but this option is not currently available to us within the time, budget and logistical constraints of the study.

b. The previous day measure (26 items including 'healthy' and 'unhealthy' items) is not provided and little description of its content is provided in literature. It does not appear to have been validated.

Response: This was a measure adapted from previous studies; we have decided to remove this measure.

Other minor concerns include:

5. Please explain Hypothesis 3 further (p. 6, lines 22-24). Does this mean that the actual process of asking people to note their preference will decrease the intervention's effectiveness? Is this based on an assumption that people do not want to change snacking behavior?

Response: Please see our response to comments 7 and 8 from Reviewer 1. We anticipate reduced effectiveness of the intervention if people answer a question about preference but then do not receive this preference (because they are randomised to an intervention). This has implications for the manner in which preference research is conducted.

We have added extra text to the third hypothesis:
The third aim is to evaluate whether assessment of participant preference for an intervention prior to random allocation influences the effectiveness of the intervention. This addresses an unanswered question in research on the effects of treatment preference i.e., does assessment of preference introduce a methodological artefact by increasing the salience of discrepancies between an individual's preferred and received treatments?

6. Collecting height and weight by self-report is an important limitation.

This will be noted as a limitation in the discussion of the paper presenting results.

7. How have the materials used in the 'Choice' arm been tested in individuals with low health literacy? Both descriptions of intervention tools should be correctly interpreted by individuals with low health literacy. (p. 10, lines 36-52). Will this be part of the piloting process?

Descriptions for the literacy-sensitive intervention uses simple language appropriate for an audience with lower health literacy. Descriptions for the standard intervention uses more complex words (e.g. 'situation'). We felt that the difference in the complexity of intervention descriptions would assist participants to select a tool that was most appropriate for their health literacy level by giving an indirect indication of the relative level of mental effort required to engage with the intervention, in addition to the explicit descriptions which describe the relatively greater mental effort required in the standard plan.

8. Please provide a brief description of the phrase 'need for cognition.' (p. 13, lines 34-39)
This has been added where it is listed under 'measures':

Need for cognition describes the extent that an individual engages in and enjoys cognitively effortful activities (i.e. activities that require a lot of thinking) (Cacioppo & Petty, 1982).

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. Journal of Personality and Social Psychology, 42(1), 116-131. doi:10.1037/0022-3514.42.1.116

VERSION 2 – REVIEW

REVIEWER	Jennifer Inauen University of Bern, Switzerland
REVIEW RETURNED	19-Mar-2019

GENERAL COMMENTS	The authors have fully addressed all my comments.

REVIEWER	Heather D. Gibbs University of Kansas Medical Center United States of America
REVIEW RETURNED	20-Mar-2019

GENERAL COMMENTS	Thank you for addressing many previously noted concerns including the addition of sugar sweetened beverages, justifying exclusion of alcohol, and removing the snacking measure that has not been validated. Although the authors have provided rationale in their response to reviewers, some of these explanations should be incorporated into the text to clarify rationale for the reader. Suggested changes include: 1. Unhealthy snacking remains undefined in the introduction. Changing this behavior is the primary aim, so it should be clear to the reader how this term is defined early on. Currently, it is not defined until page 14. If the intention is to change both the types of food snacked upon as well as the behavioral aspects of snacking, then both of these components should be apparent to the reader in the description of the intervention. Clearly, the primary aim is measuring difference in discretionary food intake, but it is not clear how the participant is informed of foods considered discretionary. It is unclear how the study will verify that participants understand what constitutes a healthy vs. unhealthy snack. Considering this is a health literacy study, attention should be paid to checking for understanding of terms and application of terms. Currently, the focus is nearly entirely on behaviors around snacking and not the quality of foods consumed when snacking. 2. Please justify in the text why 4 weeks is believed adequate for changing snacking behavior. 3. Please describe and justify in the text, your approach to describing the intervention options for the choice arm (i.e. your response to question 7).

VERSION 2 – AUTHOR RESPONSE

Reviewer 2:

Thank you for addressing many previously noted concerns including the addition of sugar sweetened beverages, justifying exclusion of alcohol, and removing the snacking measure that has not been validated. Although the authors have provided rationale in their response to reviewers, some of these explanations should be incorporated into the text to clarify rationale for the reader.

Suggested changes include:

1. Unhealthy snacking remains undefined in the introduction. Changing this behavior is the primary aim, so it should be clear to the reader how this term is defined early on.

Currently, it is not defined until page 14. If the intention is to change both the types of food snacked upon as well as the behavioral aspects of snacking, then both of these components should be apparent to the reader in the description of the intervention. Clearly, the primary aim is measuring difference in discretionary food intake, but it is not clear how the participant is informed of foods considered discretionary. It is unclear how the study will verify that participants understand what constitutes a healthy vs. unhealthy snack. Considering this is a health literacy study, attention should be paid to checking for understanding of terms and application of terms. Currently, the focus is nearly entirely on behaviors around snacking and not the quality of foods consumed when snacking.

We have addressed this comment with 2 changes to the manuscript:

1. Defining unhealthy snacking for the reader (in the introduction rather than methods): We have added a line in the paragraph that introduces the aims of the current study: Unhealthy snacking in this study includes consumption of discretionary foods as described in the Australian Guidelines to Healthy Eating (20). For the purposes of this study, this included sugar-sweetened beverages and excluded alcoholic beverages.

2. Defining unhealthy snacking for the participant: We have revised the text under ‘Baseline and follow-up surveys’ so that it now reads: [Participants] will then receive information about general reasons for reducing unhealthy snacking (such as avoiding weight gain), reasons for ‘smart snacking’ (i.e. choosing healthy snacks), definitions of healthy and unhealthy snacks (low and high in: kilojoules, fat, salt and sugars, respectively), and examples from each category.

2. Please justify in the text why 4 weeks is believed adequate for changing snacking behavior.

We have incorporated comments from the previous response to reviewers into the manuscript, under ‘Baseline and follow-up surveys’:

Consistent with our previous study, in which a change in snacking scores was detected as a result of the intervention after 4 weeks (15), participants in this study will also complete a follow-up survey after 4 weeks. In addition, the 4-week period was selected as it is a tangible time period over which participants can recall their behaviour, and the instrument for the primary outcome has been validated for recall over the previous month (20).

Again we would also like to emphasise that we will mention the lack of long-term followup as a limitation in the results paper. This limitation is now also mentioned in the ‘strengths and limitations’ box:

- This study uses a subjective outcome measure (self-reported monthly unhealthy snacking collected at a single time-point) rather than an objective measure (e.g.

unhealthy snacking observed throughout the month-long period) or more frequently reported subjective measure, and has a relatively short follow-up period; these aspects may limit the study findings.

3. Please describe and justify in the text, your approach to describing the intervention options for the choice arm (i.e. your response to question 7).

The description and justification from question 7 in the previous response has been added under point 3 in subheading ‘allocation method’:

The literacy-sensitive intervention is described in simple language whereas the standard intervention is described using more complex words (e.g. ‘tailored plan that suits your specific situation’ vs ‘simple plan using common ways to eat less snacks’). It is anticipated that the difference in language complexity will help participants select the most appropriate tool for their health literacy level by giving an indirect indication of the relative level of cognitive effort required, in addition to the explicit descriptions which describe the relatively greater cognitive effort required in the standard plan.