PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Protocol for the Process Evaluation of a Complex Intervention Delivered in Schools to Prevent Adolescent Depression: The Future Proofing Study
AUTHORS	Beames, Joanne; Lingam, Raghu; Boydell, Katherine; Calear, Alison L.; Torok, Michelle; Maston, Kate; Zbukvic, Isabel; Huckvale, Kit; Batterham, Philip; Christensen, Helen; Werner-Seidler, Aliza

VERSION 1 – REVIEW

REVIEWER	Judi Kidger University of Bristol UK
REVIEW RETURNED	07-Sep-2020

GENERAL COMMENTS	Thanks for the opportunity to review this paper. It discusses an exciting intervention and it is great to see a thorough process evaluation; this is so important in school-based studies to understand why something worked or did not and the role of context. It is good to see a logic model and that this was tested against pilot data. However, there were some points of clarity/missing information that I would like to see in a published version: The authors reference the MRC framework but there is no mention of mechanisms of impact in the logic model, nor in the aims. I would expect to see one aim to be identifying the mechanisms of change of the intervention. The aims section could be clearer, for example the authors mention a secondary aim is to evaluate individual level factors likely to impact outcomes but isn't this included under aim 2 as it includes 'individuals' in the bracket? Fidelity is sometimes considered an implementation outcome (e.g. line 11 page 13) and sometimes is separated from the implementation outcomes, for example in the aims. I don't think it is an outcome but a feature of implementation which may explain the outcomes. But it would be good for the authors to clarify either way. It is not clear how many schools/participants will take part in the process evaluation? Is it all of them for the quantitative data? But how many for the qualitative interviews? It is not clear when the interviews will take place. The process evaluation is very quantitative, and I am curious as to why there is no qualitative work planned with the students to

	triangulate their data, particularly to understand acceptability of the intervention. There may be good reason for this but it would be good to include the reasoning behind why staff interviews and not students. I suggest the authors double check that they are following Braun and Clarke as some of the terminology suggests they are planning as 'codebook' TA which is different to Braun and Clarke's 'reflexive' TA: https://www.tandfonline.com/doi/full/10.1080/14780887.2020.1769238 Observations are mentioned in methods but not in the analysis section. While I accept the authors' claim that this is the first process evaluation for digitally delivered MH school-based interventions, there are plenty of other process evaluations of school-based MH interventions. This should be acknowledged, otherwise suggesting this study will be a model or road map for other school-based MH studies implies this is the first process evaluation of any school-based MH intervention.

REVIEWER	Cristine Glazebrook University of Nottingham UK
REVIEW RETURNED	11-Sep-2020

GENERAL COMMENTS	The protocol is well-written and reflects significant expertise in what is an important and rapidly emerging area of research. I feel that some refinements in the aims would allow for more focussed analyses, but I accept that guidelines for process evaluations are broad and open to interpretation. My comments are really reflections on the complexity of the issues. There seems to be some confusion between implementation analysis at the school level (service implementation?) and a process analysis which should seek to explain intervention implementation and intervention impact ie as stated in the MRC guidelines a process evaluation needs to understand how an intervention works, for whom and in what context. The authors say that they are following MRC guidelines for the process evaluation of a complex intervention. However, the overall study aim is to understand how SPARX is implemented and delivered in schools, and to identify systematic differences and variation in delivery between clusters (individual schools). This feels more like a service implementation aim. Whilst it's important to understand the implementation in terms of knowing what has been delivered, a process evaluation in the context of an RCT also needs to try to explain the impact of the intervention(ie whether or not it was effective in reducing depression). In this way a process evaluation can inform future service implementation outside of an RCT. This protocol seems to over emphasise the implementation at the school level. In-fact what is described as the implementation strategy in the logic model (support from school) feels like part it should be conceptualised as part of the intervention package. This specific study aims are to 1) To evaluate the reach (including completion), uptake, and acceptability of the intervention (school- and student-level).

	Comment: The reach considers how well the intervention was accessed by the targets of the intervention ie what proportion of children received the intervention, completion is fidelity of delivery. The strategies to deliver the intervention eg reminders and staff support are arguably part of the complex intervention. The school is part of the context in which this happens. School uptake presumably reflects issues around participating in a randomised trial which are unrelated to the reach of the intervention itself. 2. To understand the contribution of contextual factors (e.g., characteristics of the outer/inner setting, intervention, individuals) on school-level fidelity to the implementation strategy and on Implementation outcomes (intervention reach, uptake, acceptability), as assessed from the perspectives of school staff, teachers and students. Comment: The intervention isn't a contextual factor. I think here the aim is to look at contextual factors influencing level of uptake/ child engagement with the intervention (number of modules completed?) – would be good to specify more clearly how this is measured? I think that the school support for the intervention is a contextual factor but some aspects of what is called the implementation strategy I think are part of the intervention. 3. To examine the impact of school-level variation (in implementation fidelity and outcomes) on clinical effectiveness outcomes at the school- and student-level. Comment: How is school level clinical effectiveness defined? Here the authors need to look at contextual factors influencing the effectiveness of intervention. Contextual factors are school variables, individual variables, level of engagement/uptake etc. If the protocol is for a process evaluation rather than a service implementation evaluation, then the effectiveness outcomes need to be at the student level. How do the child's individual characteristics influence intervention effectiveness (change in depression?) eg baseline mental health, age, academic level, family structure, family affluence etc?

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

1. Thanks for the opportunity to review this paper. It discusses an exciting intervention and it is great to see a thorough process evaluation; this is so important in school-based studies to understand why something worked or did not and the role of context. It is good to see a logic model and that this was tested against pilot data. However, there were some points of clarity/missing information that I would like to see in a published version:

The authors reference the MRC framework but there is no mention of mechanisms of impact in the logic model, nor in the aims. I would expect to see one aim to be identifying the mechanisms of change of the intervention.

Response. Thank you for the positive feedback about our process evaluation and logic model. We adapted the MRC framework to focus on how an evidence-based cognitive-

behavioural intervention (i.e., SPARX) works within a specific context of delivery (i.e., schools). The efficacy of SPARX has already been established in prior RCT research (i.e., see Perry, Werner-Seidler et al., 2017, doi:10.2196/jmir.8241) and the mechanisms of action have been included in the Logic Model in the 'Therapeutic Component' section.

However, we have not included mechanisms of impact as an aim or outcome for several reasons. First, as mentioned, the efficacy of this intervention is established. Second, this protocol describes a hybrid type 1 approach, with a focus on implementation process factors and outcomes in the context of an effectiveness trial. This means that variation in how SPARX is delivered and implemented at the individual and school-level are the in the focus of our evaluation, rather than how mechanisms driving the effect of the intervention.

2. The aims section could be clearer, for example the authors mention a secondary aim is to evaluate individual level factors likely to impact outcomes but isn't this included under aim 2 as it includes 'individuals' in the bracket?

Response. Thank you for picking this up. You are correct in that our aim of individual factors are included under aim 2. We have revised this section to ensure that our aims are presented more coherently, clearly, and accurately. This section on pages 9-10 now reads:

Overall, the objective of this process evaluation is to understand how SPARX is implemented and delivered in schools, and to identify systematic differences and variation in delivery. The specific aims are:

1. To evaluate the reach (including completion), uptake, and acceptability of the intervention (school- and student-level).

2. To understand the contribution of contextual factors (e.g., characteristics of the outer/inner setting, intervention, individuals) on:
    - School-level fidelity to the implementation strategy. For example, different schools will likely provide different levels of study support based on available resourcing.
    - Implementation outcomes (intervention reach, uptake, acceptability), as assessed from the perspectives of school staff, teachers and students. For example, young people's openness to receiving mental health material via an app will likely impact intervention acceptability and completion.

3. To examine the impact of school-level variation (in implementation fidelity and outcomes) on clinical effectiveness outcomes at the school- and student-level. The differing ways in which schools support and deliver the intervention will inevitably impact its effectiveness, and this evaluation will assess these differences (3, 22). For example, degree of support from senior school leaders who would be expected to understand that participating in study activities is a priority for the school will likely impact effectiveness at the cluster level.

This process evaluation has been designed to capture important information from teachers, school staff and students at both the school and individual level, which may ultimately impact the effectiveness of the intervention on clinical mental health outcomes for students. Findings will provide insight into factors which support and/or hinder the implementation of digital universal mental health programs in school settings. Knowledge gained from this process evaluation will help to inform the development of a model and guide for how to best deliver digital mental health programs to young people in schools.

3. Fidelity is sometimes considered an implementation outcome (e.g. line 11 page 13) and sometimes is separated from the implementation outcomes, for example in the aims. I don't think it is an outcome but a feature of implementation which may explain the outcomes. But it would be good for the authors to clarify either way.

Response. We appreciate that you have raised this point and understand your concerns. In our study, we are referring to fidelity to the implementation strategy, rather than to the intervention itself (which in our study, is standardised due to its digital app-based delivery). We have conceptualised fidelity to the implementation strategy as a feature of implementation that may explain the outcomes (for example, as specified in the aims and within the logic model). We have made the following changes to ensure consistency in our descriptions throughout the manuscript:

- We have moved reference to fidelity from the section on implementation outcomes (RE-AIM framework) to the section on implementation processes (CFIR) on page 11. This section now reads (see final sentence for reference to how we have conceptualised fidelity):

The CFIR identifies five major domains, including (i) the outer setting, which includes the social, political and economic context that the organisation in which implementation is occurring exists; (ii) the inner setting, which includes features and characteristics of the organisation such as leadership and relative priority; (iii) the characteristics of individuals, which includes organisational staff knowledge and attitudes about the intervention and their role and identification within the wider organisation; (iv) the characteristics of the intervention itself and; (v) implementation processes, which includes the ways that the intervention will be delivered in a given context (including fidelity to the implementation strategy).

- In the Methods section, Table 1, page 15, we have moved "Fidelity to the implementation strategy" to the "Implementation Processes" section instead of the "Implementation Outcomes" section.

- In the Measures section on page 22, we have moved "Fidelity to the implementation strategy" to "Implementation Processes" instead of "Implementation outcomes".

4. It is not clear how many schools/participants will take part in the process evaluation? Is it all of them for the quantitative data? But how many for the qualitative interviews?

Response. In our submission, we specified that all school staff from intervention schools and all study facilitators will be asked to complete the online surveys (quantitative data) and will be given the option to complete the interviews (qualitative data) – see page 25. As this is a prospective protocol, we do not yet know how many schools, school staff, and study facilitators will take part in the study/consent to take part in the interviews. This means that we cannot provide clear estimates of sample sizes. However, we anticipate that there will be approximately three school staff, on average, who will be involved in the study from each school, which we have now specified in the Procedure section (details below). On this note, the dates for delivery have been slightly adapted to accommodate a 6-month COVID-19 related delay (also see Trial Status on page 32). Further, given the need for achieving a representative sample in the interview data, we are mindful of sampling a representative range of schools (e.g., metropolitan, regional, and rural). To achieve this, we may have to use purposive sampling.

The updated Procedure section on Page 25 now reads:

The FPS has two waves of delivery (October-December 2020; April-July 2021). Process evaluation data will be collected at each wave. All school students will complete relevant survey questions at the

same time as completing primary measures for the FPS. All study facilitators will be asked to provide informed consent for their data to be used for research purposes following the compulsory face-to-face training session they attend with the research team at the Black Dog Institute. Data provided by facilitators will be from two surveys, one completed immediately following the training and another completed after their final school visit. All school staff from intervention schools (on average, three from each school, estimated number of intervention schools = 100) who were directly involved in the study will be invited to complete one survey following the six-week SPARX intervention period. All surveys will be completed online.

After completing their respective online questionnaires, all school staff and study facilitators will be given the option to participate in a 60-minute individual interview (completed either in person or remotely). Purposive sampling will be used to capture a range of diverse school settings and experiences. Face-to-face interviews will be held in a quiet room on school grounds or at the Black Dog Institute. Virtual interviews, which may be required due to COVID-19 restrictions, will also be conducted. All interviews will be audio-recorded and transcribed verbatim. All contact with study facilitators and school staff, including the semi-structured interviews, will be made by research staff who have had no previous contact with them during the trial. This independence minimises the risk of bias and demand effects.

5. It is not clear when the interviews will take place.

Response. Thanks for raising this point. We have now clarified when the interviews will take place in the Data Collection Methods and Participant Groups for Process Evaluation section on page 19. We have revised Table 2 to specify when each data collection method, including interviews, will take place:

Table 2. Summary of data forms (and collection point) provided by each of the participant groups

	Questionnaire	Individual interview	Digital analytics
Year 8 students	✓ (post-intervention)		✓ (ongoing)
School staff	✓ (post-intervention)	✓ (post-intervention)	
Facilitators	✓ (before first school visit and after final post-intervention visit)	✓ (after final school post-intervention visit)	

We have also made minor changes in the paragraph text on page 19 to ensure clarity about timing of data collection.

6. The process evaluation is very quantitative, and I am curious as to why there is no qualitative work planned with the students to triangulate their data, particularly to understand acceptability of the intervention. There may be good reason for this but it would be good to include the reasoning behind why staff interviews and not students.

Response. This is a valid point, and one that we deliberated over when designing the study. We agree that it would be ideal to interview students about their experiences and beliefs about the acceptability of the intervention. However, we decided against collecting in-depth qualitative data from students given the complexity of the study and already high level of demand placed on students (e.g., engaging with the study apps, completing the online surveys at multiple time points). Moreover, SPARX has previously undergone extensive acceptability evaluation with young people in New Zealand and Australia (e.g., https://pubmed.ncbi.nlm.nih.gov/25659116/), including an evaluation conducted by our team in advance of a previous RCT (focus group outcomes unpublished). Therefore, on balance, given the previous data suggesting this app is acceptable,

together with the minimising the burden we place on participants, we opted not to conduct student-level interviews. Instead, we will be collecting information from students about their perceived acceptability of the intervention through the short online surveys collected at post-intervention time point (this is detailed in Table 1 on page 15 under “Acceptability/appropriateness”), which reduces the burden on students’ time. Further, our proposed methodology still enables us to answer our research questions about the role of schools and school staff in mental health program implementation and support.

To clarify this point about student data collection, we have included a section in the limitations section on page 29. This section reads:

Second, given the complexity of the study and high demand placed on students (e.g., engaging with the study apps and completing the online surveys at multiple time points over five years), we are not collecting in-depth qualitative data by way of interviews. Instead, we collect information about students’ perceptions of the intervention (e.g., acceptability) through short self-report questions in the online survey.

7. I suggest the authors double check that they are following Braun and Clarke as some of the terminology suggests they are planning as ‘codebook’ TA which is different to Braun and Clarke’s ‘reflexive’ TA:
https://www.tandfonline.com/doi/full/10.1080/14780887.2020.1769238

Response. We have changed our description of reflexive Thematic Analysis on pages 26-27 to ensure that it is consistent with Braun and Clarke’s approach. This section now reads:

An iterative and reflexive approach will be used to analyse the data, incorporating themes from the data together with topics covered in the interview guide. Two coders will independently engage in a familiarisation phase before generating codes and initial themes for a subset of the data. These codes and themes will be reviewed and discussed by the two coders, with refinement occurring via an interactive process. A senior qualitative analysist will also review the first-stage coding framework and scheme before all transcripts are coded by the first coder. Refinement will continue to occur via an interactive process until final codes and themes are realised and defined across the whole data set.
8. Observations are mentioned in methods but not in the analysis section.

Response. We apologise for this oversight. We have removed any mention of “observations” from the methods section, including on page 19, and page 22. Further, in Table 1 on page 14, we replaced Qualitative feedback based on informal observation with Interviews with study facilitators.

9. While I accept the authors’ claim that this is the first process evaluation for digitally delivered MH school-based interventions, there are plenty of other process evaluations of school-based MH interventions. This should be acknowledged, otherwise suggesting this study will be a model or road map for other school-based MH studies implies this is the first process evaluation of any school-based MH intervention.

Response. This is a worthwhile point to clarify in our manuscript and it was not our intention to present our work as the first process evaluation in broad school-based mental health interventions. We have specified on pages 6-7 that although other process evaluations of school-based mental health interventions have been conducted, this is the first evaluation of a digitally delivered intervention. We also provide a rationale for why process evaluations of digital interventions are important, given that some of the concerns are likely unique to this method of delivery. This section now reads:

Process evaluations of face-to-face mental health programs have been documented, some of which have been delivered in school settings (e.g., 17). However, we were unable to find any process evaluation descriptions in the literature that evaluate digital mental health programs in school settings. Conducting process evaluations specific to digital methods of delivery are important because the contextual barriers and facilitators, as well as concerns around fidelity of the intervention, are likely to have unique characteristics. For example, facilitator or training in intervention delivery will be different in supporting the use of an automated program relative to a face-to-face program. This process evaluation is an initial step towards addressing that gap.

Reviewer 2

We would like to preface our replies to reviewer 2 addressing a conceptual issue that was raised across comments. Our protocol paper documents the process evaluation within a hybrid type 1 design, which means that implementation process factors and outcomes are evaluated in the context of an effectiveness trial (i.e., the Future Proofing Randomised Controlled Trial). In this protocol, we focus on the contextual factors within schools that affect the implementation of the evidence-based intervention (SPARX). We also consider how these factors, in turn, affect student-level outcomes. The protocol for the effectiveness trial, which specifically reports on clinical outcoms for students at an individual level, has already been published, and we reference this on page 8 of the introduction (https://pubmed.ncbi.nlm.nih.gov/31898512/). Therefore, we have not provided extensive details about student-level clinical outcomes. We used the MRC framework to guide our approach, making adaptations as necessary to ensure focus on contextual factors – rather than the efficacy or mechanisms of the intervention per se.

1. The protocol is well-written and reflects significant expertise in what is an important and rapidly emerging area of research. I feel that some refinements in the aims would allow for more focussed analyses, but I accept that guidelines for process evaluations are broad and open to interpretation. My comments are really reflections on the complexity of the issues.
Response. We are glad to receive your positive feedback. In line with your comment, as well as from Reviewer 1, we have revised our aims (see pages 9-10) to improve their clarity for readers.

2. There seems to be some confusion between implementation analysis at the school level (service implementation?) and a process analysis which should seek to explain intervention implementation and intervention impact ie as stated in the MRC guidelines a process evaluation needs to understand how an intervention works, for whom and in what context. The authors say that they are following MRC guidelines for the process evaluation of a complex intervention. However, the overall study aim is to understand how SPARX is implemented and delivered in schools, and to identify systematic differences and variation in delivery between clusters (individual schools). This feels more like a service implementation aim.

Response. We adapted the MRC framework to guide our approach. We agree with the reviewer that our overall study aim is to understand how SPARX is implemented in schools, and to identify school-level variation in delivery/support. In terms of the MRC framework, we see this fitting under the umbrella of context – there is evidence documenting the efficacy of SPARX, the mechanisms of impact are known; our study is about understanding whether (and how) SPARX can be delivered at scale in schools and still be effective for students.

On page 9, we have specified more clearly that we used the MRC framework as a guide:

We adapted the MRC framework for complex interventions to focus on effectiveness of an evidence-based intervention within a specific context of delivery (i.e., schools).

3. Whilst it's important to understand the implementation in terms of knowing what has been delivered, a process evaluation in the context of an RCT also needs to try to explain the impact of the intervention(ie whether or not it was effective in reducing depression). In this way a process evaluation can inform future service implementation outside of an RCT. This protocol seems to over emphasise the implementation at the school level. In-fact what is described as the implementation strategy in the logic model (support from school) feels like part it should be conceptualised as part of the intervention package.
Response. Our evaluation will consider the clinical impact of the intervention at the student level (e.g., on depression). This is documented in the RCT trial protocol, which is published elsewhere: Werner-Seidler, Huckvale et al., 2020, doi:10.1186/s13063-019-3901-7. In the process evaluation protocol documented here, we focus on the contextual factors that influence those effects, to provide a more nuanced understanding of how the school context affects implementation (and in turn, develop conclusions about whether schools are a feasible context for delivery of digital mental health programs at scale).
The reviewer also raises an interesting point about the packaging of the intervention. We acknowledge that differentiating implementation strategies from the intervention in our study is challenging. Our decision separate these out was driven by our objective of exploring whether our digital intervention approach in schools would be sustainable at scale. Based on our teams' expertise in school-mental health, we anticipate that the implementation supports and high levels of resourcing usedn our study may not be available or feasible for schools to continue at the end of the trial, meaning limited sustainability in the long-term. Therefore, we are interested in identifying how these supports and resourcing vary across schools, and how this school-level variation affects the uptake (and effectiveness) of SPARX by students so that we can draw some conclusions about what schools would need to do in order to retain the program without the support of the research team.

4. This specific study aims are to
a) To evaluate the reach (including completion), uptake, and acceptability of the intervention (school- and student-level).

Comment: The reach considers how well the intervention was accessed by the targets of the intervention ie what proportion of children received the intervention, completion is fidelity of delivery. The strategies to deliver the intervention eg reminders and staff support are arguably part of the complex intervention. The school is part of the context in which this happens. School uptake presumably reflects issues around participating in a randomised trial which are unrelated to the reach of the intervention itself.

Response. We agree with this interpretation. Differentiating the research trial processes from the implementation processes is difficult in a complex trial such as the Future Proofing Study. This is a factor that we will consider during the qualitative feedback component in order to see what parts were perceived by stakeholders as necessary, feasible and appropriate (or not).

We have specified that qualitative interviews will explore this complexity on page 24:

For school staff, questions focus on: motivations and expectations about the intervention and study processes more broadly; knowledge and beliefs about the intervention; relative advantages of the intervention; self-efficacy; barriers and facilitators affecting the delivery of the intervention, including fidelity; appropriateness and acceptability of the intervention; and recommendations for future implementation.

b) To understand the contribution of contextual factors (e.g., characteristics of the outer/inner setting, intervention, individuals) on school-level fidelity to the implementation strategy and on Implementation

outcomes (intervention reach, uptake, acceptability), as assessed from the perspectives of school staff, teachers and students.

Comment: The intervention isn't a contextual factor. I think here the aim is to look at contextual factors influencing level of uptake/ child engagement with the intervention (number of modules completed?) – would be good to specify more clearly how this is measured? I think that the school support for the intervention is a contextual factor but some aspects of what is called the implementation strategy I think are part of the intervention.

Response. We use the CFIR to identify and explore determinants that effect implementation. Within the CFIR, "characteristics of the intervention" are conceptualised as a contextual determinant, and so we use this terminology within our approach.

Measurement of contextual factors are comprehensively outlined in Table 1, however, we have added additional information about modules completed in the following sections:

1. Table 1, page 15: Digital analytic data including usage (app downloads, installs, opens), completion rate (number of modules completed) and time spent using SPARX.
2. Pages 22-23: SPARX app usage data from students will allow for the assessment of app use (downloads, installs, and opens), completion (number of modules completed, out of a total of seven)…

The final comment about school support feeds into the difficulty in differentiating implementation strategy from intervention, which we have responded to in more detail in 4 a) and b).

c) To examine the impact of school-level variation (in implementation fidelity and outcomes) on clinical effectiveness outcomes at the school- and student-level.

Comment: How is school level clinical effectiveness defined? Here the authors need to look at contextual factors influencing the effectiveness of intervention. Contextual factors are school variables, individual variables, level of engagement/uptake etc. If the protocol is for a process evaluation rather than a service implementation evaluation, then the effectiveness outcomes need to be at the student level. How do the child's individual characteristics influence intervention effectiveness (change in depression?) eg baseline mental health, age, academic level, family structure, family affluence etc?

Response. Student-level effectiveness outcomes, as well as the influence of individual characteristics, are already documented in the RCT protocol paper that we have specified in comment 3.

School-level clinical effectiveness is defined as changes in clinical outcomes (e.g., self-reported depression, anxiety) for different schools. We will examine the extent to which different contextual factors, such as school level variables (e.g., location, socio-economic status), levels of engagement and others, affect these outcomes at the school-level. We have more clearly defined school-level clinical effectiveness in Aim 3 on page 10 to aid clarity, which now reads:

To examine the impact of school-level variation (in implementation fidelity and outcomes) on clinical effectiveness outcomes at the school- and student-level. School-level clinical effectiveness is defined as changes in clinical outcomes (e.g., self-reported depression) for different schools. The differing ways in which schools support and deliver the intervention will inevitably impact its effectiveness, and this evaluation will assess these differences (3, 22).

VERSION 2 – REVIEW

REVIEWER	Judi Kidger University of Bristol UK
REVIEW RETURNED	13-Nov-2020

GENERAL COMMENTS	Thank you for the opportunity to review this revised version. I feel the added points of clarification have addressed my concerns.

REVIEWER	Cris Glazebrook University of Nottingham, UK
REVIEW RETURNED	11-Nov-2020

GENERAL COMMENTS	I think it is clear from the authors' responses to the reviewers' comments that this is a protocol for a process evaluation of intervention implementation within the context of a randomised trial. It appropriately covers exploration of the level of intervention delivered, reach, fidelity and contextual factors influencing delivery and uptake. As such it makes a valuable addition to the literature. Mechanisms of impact (what influences the effectiveness of the intervention ie reduction in depression) appears to be outside the scope of this study. This aim is not covered in the summary table, depression scores are not listed as a data source and the authors clarify that factors influencing effectiveness are covered by the trial protocol. The scope of the process evaluation should, therefore, be made clear in the abstract and paper. Additionally, there are a few minor amendments needed to accurately reflect the study aims. For example, the statement "How these factors relate to trial effectiveness outcomes will also be assessed" needs to be removed from the introduction and aim 3 removed from the aims. Alternatively, any exploration of mechanisms of impact (ie in this case the relationship between intervention implementation and effectiveness) needs to be more clearly specified eg in the summary table. Will all 100 schools be included in the analyses? Ideally the analysis should incorporate the child's perspective on how the intervention implementation affected the impact of the intervention.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1

1. Thank you for the opportunity to review this revised version. I feel the added points of clarification have addressed my concerns.

Response. We appreciate the suggestions that you provided in the initial round of revisions and are glad to hear that we addressed your concerns.

Reviewer 2

1. I think it is clear from the authors' responses to the reviewers' comments that this is a protocol for a process evaluation of intervention implementation within the context of a randomised trial. It appropriately covers exploration of the level of intervention delivered, reach, fidelity and contextual factors influencing delivery and uptake. As such it makes a valuable addition to the literature.

Response. We appreciate your feedback and the time spent reviewing our revised manuscript.

Mechanisms of impact (what influences the effectiveness of the intervention ie reduction in depression) appears to be outside the scope of this study. This aim is not covered in the summary table, depression scores are not listed as a data source and the authors clarify that factors influencing effectiveness are covered by the trial protocol. The scope of the process evaluation should, therefore, be made clear in the abstract and paper.

Response. We outlined at the outset of the design section on page 11 that we use a hybrid type 1 design. A hybrid type 1 design involves testing a clinical intervention while gathering information on its delivery or implementation during the effectiveness trial. In this type of design, the focus is on collecting process data, or contextual information, about what affects the implementation of an intervention in the real world. In this protocol, we are reporting only on the implementation outcomes and how these relate to the effectiveness of the intervention. The direct mechanisms of impact will be reported separately in the primary outcomes paper (see protocol: https://pubmed.ncbi.nlm.nih.gov/31898512/), and in this respect, is beyond the scope of the current protocol. To clarify this, as suggested by the reviewer, we have revised the abstract on page 2 to also specify the design type, such that the first sentence of the "Methods and Analysis" section now reads:

Using a hybrid type 1 design, a mixed methods approach will be used with data collected in the intervention arm of the Future Proofing Study.

2. Additionally, there are a few minor amendments needed to accurately reflect the study aims. For example, the statement "How these factors relate to trial effectiveness outcomes will also be assessed" needs to be removed from the introduction and aim 3 removed from the aims. Alternatively, any exploration of mechanisms of impact (ie in this case the relationship between intervention implementation and effectiveness) needs to be more clearly specified eg in the summary table.

Response. We thank the reviewer for this suggestion and have now included a sentence in Table 1, page 16, that refers to the relationship between intervention implementation and effectiveness. This sentence reads:

How might the school-level variation (in implementation fidelity and outcomes) affect clinical effectiveness outcomes (e.g., self-reported depression)? (Aim 3)

4. Will all 100 schools be included in the analyses?

Response. Yes, we anticipate that 100 intervention schools will provide quantitative data. In the "Data analysis" section on page 26, we specify that:

Survey questionnaire data (from approximately 100 schools) will be exported into data analytic software for analysis.

We specify on page 25 that purposive sampling will be used for qualitative data collection.

5. Ideally the analysis should incorporate the child's perspective on how the intervention implementation affected the impact of the intervention.

Response. This is a valid point, and one that we deliberated over when designing the study. We agree that it would be ideal to gather information about implementation from students. Ultimately

however, we decided against collecting in-depth implementation data from students given the complexity of the study and the already high level of demand placed on students (e.g., engaging with the study apps, completing the online surveys at multiple time points over five years). We will, however, be collecting information from students about their perceived acceptability of the intervention through the short online surveys collected at post-intervention time point (this is detailed in Table 1 on page 15 under “Acceptability/appropriateness”), which reduces the burden on students’ time.