Author's response to reviews

Title: Effective Elements of School Health Promotion across Behavioral Domains: A Systematic Review of Reviews

Authors:

Louk WH Peters (louk.peters@tno.nl)
Gerjo J Kok (G.Kok@psychology.unimaas.nl)
Geert TM ten Dam (G.T.M.tenDam@uva.nl)
Goof J Buijs (gbuijs@nigz.nl)
Theo GWM Paulussen (theo.paulusse@tno.nl)

Version: 2 Date: 4 May 2009

Author's response to reviews:

MS: 1978596831246401
Effective Elements of School Health Promotion Across Behavioral Domains: A Systematic Review of Reviews
Louk WH Peters, Gerjo J Kok, Geert TM ten Dam, Goof J Buijs and Theo GWM Paulussen

Dear editorial staff, dear Dr. Bucceri,

We were happy to learn that BMC Public Health, on the basis of the comments by three referees, will consider a revised version of our manuscript entitled “Effective Elements of School Health Promotion across Behavioral Domains: A Systematic Review of Reviews”. Please find enclosed the revised version of our manuscript, which has about 7779 words and 9 tables. As in the previous version, Tables 2-9 are in landscape format to accommodate the number of columns needed for presenting the results; these tables are included in one additional file.

We gladly acknowledge that the referees consider the manuscript to be important in its field and well-written. We have taken great care in revising and improving the manuscript and in responding to the referees. Below you will find our detailed point-by-point response. We first respond to the second referee, since this referee had the largest number of comments and was the only one who stated that major compulsory revisions are necessary.

Referee 2
This referee had 7 issues for major compulsory revision, which are addressed below.

1. The referee asked for information about the origin of the data extraction form. On page 7 we have now mentioned that the form was developed ad hoc for this review but was largely based on other sources. The items inquiring about the
focus and methods of reviews were derived from other systematic reviews of reviews [21,27], whereas the categories of effective elements were based on data extraction forms used in systematic reviews [e.g., 31,32].

2. The referee’s statement that the tool we used for the quality rating is recommended by the Cochrane handbook is incorrect. We used the Quality Assessment Tool for Reviews, which is specifically meant to be used in a review of reviews to assess the quality of included reviews, whereas the tool recommended in the Cochrane Collaboration handbook (Quality Assessment Tool for Quantitative Studies) is meant to be used in a review to assess the quality of quantitative primary studies. Nevertheless, in our original manuscript we did mention the latter tool in the last section of the Discussion, where we recommended that reviewers of primary studies use this tool. In the revised version, we have now added that this tool is recommended in the Cochrane handbook (p. 30).

3, 4, 5, 6. Four of the referee’s comments pertain to decisions with respect to the quality of reviews in the Methods (Analysis), Results, Discussion and Limitations sections. The referee questions the making of considerations of effectiveness irrespective of the quality rating (comment 3) and recommends to exclude weak reviews (comment 4) or at least support their inclusion with adequate references (comment 6). Also, the referee would like to see an explicit distinction between conclusions that are based on high-quality reviews and those based on low-quality reviews (comment 5).

In response to these comments, and in line with other reviews of reviews [21,27,29], we decided to consider only strong and moderate reviews as a basis for conclusions about effective elements (see p. 9). This decision led to major changes in the Methods (Analysis), Results, Discussion and Conclusion sections. We used weak reviews only in a supplementary and speculative way: for elements that had been examined by strong or moderate reviews in at least one domain, the results of weak reviews were used to indicate whether there is any suggestion that the element might be effective in other domains. This is mentioned explicitly in the Analysis (p. 10) and Discussion (p. 25) sections; also, throughout the Results section, whenever weak reviews are mentioned their status is consistently indicated as ‘supplementary’.

In comment 3 the referee considered our analysis to be too fuzzy. In response, we decided to apply two types of analysis: one based on interpretation, similar to the analysis in the previous version, and the other based on rules used in another review of reviews [26]. A main difference between the two types of analysis lies in the handling of conflicting results between reviews. In our view, if differential results arise, it is logical to try to come to an overall conclusion by interpreting the differential results and examining which factors may have caused them. Since the potential causes of differential results may pertain to all kinds of factors (e.g., differences in study designs, study samples included, follow-up periods examined, specificity of operationalizations, etc.) we did not set strict rules for interpretation. Although this may be regarded as fuzzy by some, we believe that interpretation may lead to more in-depth understanding; also, much
of the fuzziness is taken away by explicit discussion of the specific reviews involved and clear presentation of their characteristics and results in the tables.

To accommodate the referee’s critique, we included the second type of analysis, which did set strict criteria for sufficient, tentative and insufficient evidence and for handling conflicting results. In the Discussion section we discuss the results obtained with both types of analysis (p. 24-25): the results were comparable for five elements, and different for two elements. The differential results of the two types can be explained by the different approaches to handling conflicting review results (see p. 24-25).

6-7. The last two issues for major compulsory revision pertained to textual remarks with respect to the Conclusions section (comment 6) and the manuscript in general (comment 7). Since the other two referees considered the manuscript to be well-written and the conclusions to be well balanced, we decided not to revise the manuscript on these points.

Referee 2 had 7 issues for minor essential revisions.

1. The referee asks which government is indicated by the word ‘Nation’. However, Nation is the name of an author. We did not revise this, since in our view this was sufficiently clear in the original text, as we used the phrase ‘Nation and colleagues’ and gave the reference number.

2. In the previous text we used the terms sex and sexuality interchangeably. As suggested by the referee, we have now used the term sexual behavior as much as possible (p. 6). In some cases, mainly when referring to programs, we have used the term sexuality programs, and in the Results section the term sexuality reviews.

3. As suggested by the referee, we have specified the author who extracted the data (p.8).

4. The referee questions whether it is appropriate to use reviews about substance abuse as indicative for the behaviors of tobacco and alcohol use (p. 9). In the revised version this question is no longer relevant, since in response to comments made by referee 3 we decided to combine the categories of tobacco, alcohol and substance abuse into one substance abuse domain (see our response to referee 3 below).

With respect to the previous version we would like to add that we agree with referee 2 that tobacco- and alcohol-specific reviews are preferable for making inferences about these specific behaviors, but that we can also defend the inclusion of substance abuse reviews by pointing out that nearly all 15 substance abuse reviews included a focus on both tobacco and alcohol.

5. In line with the referee’s comment, we have now added the term ‘statistically significant’ where we previously only mentioned ‘significant’ (p.12, now p.13).

6. In line with the referee’s suggestion, we have now clearly mentioned that the results for price regulation in the substance abuse domain are from weak reviews (p. 21, now p. 22); this also followed from our general decision to use weak reviews only in a supplementary way (see also above: comments 3-6 pertaining
to major compulsory revisions). Elaborating on the element of ‘laws and policies’ in Table 8: we excluded the results for the specific elements ‘laws restricting sales to minors’ and ‘minimum drinking age’ from the text and Table 8, because collapsing the substance abuse behaviors into one single domain had the consequence that these results are only relevant for one domain. In the revised version, price regulation is the only policy element that is relevant for multiple domains (nutrition, substance abuse); therefore, we chose to mention only this specific element in the text and Table 8 instead of using the general term policies.

7. As requested by the referee, we aligned left the titles in Tables 2-9.

Discretionary revisions

1. We did not follow the referee’s suggestion to use author instead of assessor (Abstract), since the second assessor is not part of the team of authors.

2. We followed the referee’s suggestion to refrain from superlatives, such as vast and enormous (p. 4-5).

3. We removed the term ‘a fair share’, as we do not consider it to be very relevant. Specifying this issue, as referee 1 requested, would lead to a large increase in the already large number of words.

Referee 1

Referee 1 was very positive about the manuscript and did not have any issues for major compulsory revision.

The referee had 3 issues for minor essential revision.

1. The referee was unsure whether the terms ‘aspects’ and ‘elements’ were used interchangeably throughout the Results section and suggested to use only one term or explain the differences. Indeed, we used these terms interchangeably and have now consistently used the term ‘elements’ to refer to elements that are examined as ‘effective elements’. To elaborate: in cases where we use other terms (e.g. ‘characteristics’), we refer to review characteristics other than effective elements, such as inclusion criteria.

2. The referee had problems with a statement about parent involvement and sexuality programs, because it was considered to be speculative and to disrupt the text flow (p. 20, now p. 22). We have now deleted the statement. In fact, we no longer mention the sexuality domain specifically with respect to parent involvement, mainly to restrict the number of words, as other changes (particularly those in response to referee 2) led to an increase in the word count.

3. The referee was confused about the negative formulation of the element “intervention content that is not limited to dissemination of information” since it leaves the interpretation open to many other elements or processes. We can understand the referee’s confusion. The negative formulation stemmed from the fact that the element ‘knowledge-only approach’ (programs that only focus on knowledge transmission) was found to be ineffective, whereas the other elements in our list of effective elements were found to be effective. With the
negative formulation we tried to incorporate the ineffective element into a list of effective elements, but apparently, this is confusing. In the revised manuscript, we have consistently used the formulation that a knowledge-only approach is ineffective (p. 16, 25, 31), which we hope is more clear.

To elaborate: the void that was left by the negative formulation (‘not limited to’) may be filled by pointing to the elements of program content that contribute positively to effectiveness across domains: social influences/norms and cognitive-behavioral skills. Thus, whereas it is not effective to only focus on knowledge, it is effective to address social influences and cognitive-behavioral skills.

In comment 3, the referee also stated that there is a need to identify what the process is that would constitute an element of the intervention, and that content is not a process. It is unclear to us what the referee means by this statement, and consequently we felt unable to make any revisions in this respect. In our view, teaching or learning a particular program content (whether it pertains to knowledge of health consequences, social influences or cognitive-behavioral skills) can certainly be considered a process.

Discretionary revisions
1. As suggested by the referee, we stated the domains in the last paragraph of the abstract.
2. We have read the additional paper the referee suggested and, as it was relevant, we have included the reference.
3. The referee suggests to explain ‘fair share’ in quantitative terms (p. 10, now p.11). In light of the already large number of words we did not follow this suggestion, as it would cost many additional words and is not extremely relevant. Since the expression ‘fair share’ may evoke this question in readers and we are not able to answer it in the text, we have deleted that expression from the text. Nevertheless, we are of course willing to address this issue in this letter:

Of the 5 multiple behavior reviews, three focused entirely on school-based programs [16,24,79], the fourth included ‘mostly’ such programs [66] and in the fifth 67% of programs were school-based [35].

Of the 24 substance abuse reviews: fifteen focused entirely on school-based programs [23,25,32,34,37,40,41,43,44,61,69,71,72, 80,82], four were unclear about the percentage of school-based programs [54,59,60,64], and in the others the percentage ranged from 70% [42] to 96% [45].

Of the 17 sexuality reviews: five had an exclusive focus on schools [38,50,83,84,85], three were unclear [17,36,46,58], and in the remaining reviews the percentage of school-based programs was between 30% [47] and 89% [84].

Of the 9 nutrition reviews: three focused entirely on school-based programs (at least for data used for this review) [31,86,70], one was unclear [87], and in the remaining reviews the percentage ranged from 40% [52] to 93% [51,57].

4. We followed the referee’s request to replace the term ‘authors’ with ‘reviews’ (p. 11, now p. 12).
5. As the sentence about effect sizes and qualitative statements was unclear in meaning, we rewrote it. We have now made it clearer that the qualitative statements about effects in reviews (‘most reviews are cautiously positive’) are in line with quantitative results of meta-analyses and reviews (‘effect sizes are small but statistically significant’).

6. This comment, pertaining to problems with a particular part of a sentence about parent involvement in sexuality programs, is no longer relevant, as the whole sentence was deleted in the revision (see also comment 2 of minor essential revisions).

7. For reasons of readability, the referee suggests to write out the elements instead of giving the corresponding letters when referring to the list of effective elements on page 23 and 24 (now p. 24-26). Although we place great value on readability and have done our best to produce a paper that is easily readable, we have only partly complied with the referee’s suggestion, mainly to restrict the word count. In the revised version, we have retained the corresponding letters on page 24; since the list of elements is indented and the corresponding letters are used immediately below the introduction of the list, we believe that readers can easily identify which elements are meant. On page 26, the corresponding letters are used to indicate the similarities with the list of another review. Because two lists are involved here, we feel that writing the elements in full would complicate rather than facilitate understanding of the similarities between the two lists; also it would take more words to explain the similarities.

Referee 3
Referee 3 was very positive about the manuscript, had no issues for major compulsory revision and only a small number of minor essential and discretionary revisions.

Minor essential revisions
The referee is confused about the use of periods in Tables 3-9. We used a period for each review to indicate that the results for that particular review have ended. We did so to improve the clarity of results, especially since some results include numbers or phrases and it may not be clear to readers where the results of one review end and those of the next review begin. Perhaps ironically, the periods apparently confuse the referee. Nevertheless, we have decided to keep them. However, if the referee or the editor wishes that we remove the periods, we will of course do so.

Also, the referee does not understand the meaning of ‘country’ in Table 3. ‘Country’ meant that countries can have different prevention goals for alcohol in relation to goals for other substances such as tobacco (US: abstinence of all substances, Europe: moderation for alcohol, abstinence for tobacco), which may affect results of alcohol-specific versus multiple-substance approaches. This was explained in the text (p. 13). In hindsight, the ‘country’ result in Table 3 should not have been mentioned in the row about abstinence, but one row above about one-substance versus multiple-substance programs. In the revised version, we decided to collapse the behaviors of tobacco and alcohol into one substance
abuse domain (see below). Since we only include elements that have been examined in at least two domains, the element ‘one versus multiple substances’ was removed from Table 3, as it is only relevant for the substance abuse domain. This also means that ‘country’ is no longer mentioned in Table 3.

Discretionary revisions

1. The referee argues that the inclusion of substance abuse, tobacco and alcohol in separate columns in the tables may give extra weight to their evidence compared to sexuality and nutrition. In addition, she questions whether the evidence would be the same if substance abuse, tobacco and alcohol were in a single category to be contrasted with sexuality and nutrition. We followed this suggestion and decided to combine the substance abuse, tobacco and alcohol reviews into one single substance abuse category. This decision was based on three considerations: a) most review results in this domain were from the 15 substance abuse reviews, with fewer results from the 5 tobacco-specific and 4 alcohol-specific reviews, b) results of substance abuse reviews were used as indicative for the specific behaviors of tobacco and alcohol, and c) a smaller number of columns in Tables 3-9 may improve the clarity and length of these tables. The decision to combine the specific behaviors into one single substance abuse domain did not affect the conclusions about effective elements, mainly because of points a and b did not change.

2. As suggested by the referee, we have explained the quality ratings of the reviews in the notes section of Tables 2-9.

3. To our knowledge, the Quality Tool for Reviews is not available from the Internet as a separate document. However, the tool is included in two reviews of reviews that are retrievable through the web [21,27]. We have mentioned the URLs of these reviews in the reference list.

4. Abstract: Since we combined the substance abuse behaviors into one domain, we believe that we have adequately addressed the referee’s concern about the difference between the domains mentioned in the tables and those mentioned in the abstract.

We hope that the revised manuscript and our detailed point-by-point response are satisfactory to the referees and to the editorial staff of BMC Public Health. We have made sure that the revised manuscript conforms to the journal style.

We hope to hear from you soon.

Sincerely, on behalf of all authors,

Louk Peters
TNO Quality of Life
BU Prevention and Health Care
PO Box 2215
2301 CE Leiden