Reviewer's report

Title: Do 'public health' parenting interventions work for children? A systematic review and meta-analysis of Triple P parenting programmes

Version: 1 Date: 9 July 2012

Reviewer: Frances Gardner

Reviewer's report:

General comments
This review has many substantial strengths; it asks a very important question about the Triple P system of parenting programmes, at a time when it is being disseminated rapidly round the world as a universal programme which is intended to confer substantial public health benefit. The question is original in that the review focuses on this public health benefit; moreover, I believe it is conducted to more rigorous standards, using PRISMA guidelines, than previous reviews on the topic. The paper is clearly written, and conveys a good deal of information quite concisely.

Major compulsory revisions
1. The review could cover rather better the issue of universal vs. targeted prevention. These issues could be addressed in a revised paper; this would add considerably to the value, clarity and originality of the review. The Triple P programme is disseminated as a population-wide system of parenting support, aimed at all parents, and operating both at a 'universal' level, and as an impressive 'stepped' series of more intensive interventions targeted at populations with increasing levels of need, including clinical populations. My impression from the literature, and from previous reviews of related interventions (eg Furlong et al’s 2012 Cochrane review), is that the evidence is strong for targeted parenting programmes, aimed at those showing clinical or subclinical levels of behaviour problems. However, it is much less convincing for ‘universal’ samples, for which several trials of Triple P and other interventions have shown no effect (eg Malti et al, 2011 Swiss trial). Thus I think it would be important for the review make better use of its extensive data to draw conclusions about the effectiveness of the programme for ‘universal’ vs. targeted samples. This is a really critical question to answer for informing policy, and will have a profound effect on the cost of dissemination. The authors would need to categorise the programmes more fully by level of need of children, then comparing effects in samples where all/ or the majority are above the clinical cut off, vs. those that are not. At the moment, this information is made explicit in table 1 only for some of the trials, and there is no subgroup analysis based on this factor. I know that for many of these trials, doing this is not easy, as there is often unclear reporting about inclusion criteria, however, one approach might be to inspect mean baseline scores on child behaviour problems. I note that this is also the first objective of the paper, to ‘identify characteristics of the populations’ in the trials;
without this information, this objective is not met.

The authors spend some time critically discussing the evidence for whole-population interventions, however, it may be that much of what is delivered in these populations is in fact programmes for targeted use (eg a high percent of Triple P level 4, aimed at high need children, appears to have been delivered in the South Carolina population trial); so although this is a valuable discussion of an important question, it does not address the somewhat different question of universal-level delivery of Triple P. In order to live up to the title of the paper, the authors should ensure that these concepts are used as clearly and precisely as possible throughout.

2. The discussion is generally of a high standard, however, I think could be a little more balanced- after all, many might argue that an average effect size on child outcomes of .6 is very good, and that since reporting standards were poor, perhaps for some studies the risk of bias could be described as ‘unknown’ rather than ‘high’, especially as the majority were published at a time when standards were not so high. The authors should include some further discussion about whether alternative interpretations of the data could be made.

Discretionary revisions

1. Eligibility criteria: studies had to have a ‘comparison group’, but is that all? Were there no further inclusion criteria based on methodological quality – did the authors intend to include studies that were not only non randomised, but which had comparison groups that were unmatched, or selected to be different? This should be clarified.

2. What definition of reporting bias was used? Was it only based on comparing abstract and results? This could be made a bit clearer, and any limitations of the approach drawn attention to. Later on, in the risk of bias section, the authors make many useful points about ‘selective reporting’, which seems to be the same as what Cochrane calls ‘outcome reporting bias’. Perhaps they could clarify the usage of these important terms, and make sure they are consistent with PRISMA/ Cochrane handbook terminology.

3. There are many very important recommendations for the field, in particular the need for better adherence to trial reporting guidelines. However, the final para of the conclusions should be revised to include a strong plea to journal editors and reviewers, and not just to authors. There are many psychology journals (particularly but not always, lower impact ones), which, despite publishing many trials, have not signed up to/ or do not adhere to CONSORT guidelines. It could be emphasised that authors who choose journals that do not adhere to these guidelines, and editors who choose not to adopt them, are doing the field, as well as their own work, a disservice.

4. The abstract conclusions should reflect a bit more accurately the content of the review. They refer only to the credibility of the results with respect to conflicts of interest. The results find high risk of bias / or poor reporting in many spheres, and this should be stated more fully in the abstract conclusions, rather than focusing
on the single issue of conflicts of interest. In the section on whole-population interventions, the authors should make it clear which trials under discussion are included in the review, and which (if any) are not.

Minor essential revisions:

In the 1st para of the conclusions, when the authors suggest that targeted programmes may work better, this should be backed up by recent high quality evidence, including Furlong et al. 2012 Cochrane review on group based parenting programmes.

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I have received funding but not fees for a keynote speech at a Triple P conference, from the conference organisers. I am involved in two small pilot trials of Triple P interventions.