Author’s response to reviews

Title: A randomized controlled trial to examine the effect of two teaching methods on preschool children’s language and communication, executive functions, socioemotional comprehension, and early math skills

Authors:

Tove Gerholm (tove@ling.su.se)

Petter Kallioinen (petter@ling.su.se)

Signe Tonér (signe.toner@ling.su.se)

Sofia Frankenberg (sofia.frankenberg@buv.su.se)

Susanne Kjällander (susanne.kjallander@buv.su.se)

Anna Palmer (anna.palmer@buv.su.se)

Hillevi Lenz-Taguchi (hillevi.lenz-taguchi@buv.su.se)

Version: 1 Date: 21 Jun 2019

Author’s response to reviews:

Replies to reviewer comments:

We now marked all changes in yellow in the text itself, as we realize that the line-references given below might change depending on how you view the text.

Reviewer 1:

1) In order to understand better the paradigm (which is really complex and full of nested variables, as you also acknowledge), I decided to look for your cited previous work, which in the manuscript is referenced [69] [link: https://bmcpsychology.biomedcentral.com/articles/10.1186/s40359-018-0239-y], and I realized that you first published a "Study protocol” paper and the present one is one of the further "Research papers” discussing different findings from the big study. In my opinion this should be stated before in the manuscript (for example, at the beginning of the "Aims” paragraph of the "The aims, interventions, questions and hypotheses of the study" section), because to me it was surprising to discover it so late during the reading of the paper.
Reply: This is a very good point, we have now added a few lines on the Protocol in the suggested section, you’ll find it on lines: 209-211.

2) Regarding the sequential numbering of the citations, I found that after citation [68] (line 271 of the manuscript), it jumps directly to the citation number [70, 71] (line 293), and the citation [69] is found later in line 333.

Reply: Thank you for pointing this out, the error is now corrected.

3) I also noticed that EEG suddenly appeared in the Methods section ("A subgroup of children was sampled to participate in the EEG testing", line 531) without being addressed anywhere in the background or previous sections and, while I personally know ERPs and how they work, I think that other readers of the paper might want to know a bit more about the methodology, about what has been found in literature on this topic (ERP registered on children performing auditory attentional tasks), or at least why did you consider to use this methodology to add some important data in your study. It would be also easier for readers to understand your results about selective attention task. In fact, psychophysiological measures are not even mentioned in the research questions nor in the expected results, so I am not sure that it's clear why did you include EEG results in your report if you don't explain them. Citation [78] is one of the few references about electrophysiological measures in children, but it is only present in the References section and is not cited anywhere in the text. There are others [133, 134], but I cannot find in the text a part where you discuss about the components of the ERPs that you were looking at (N100, P200, P300?), you only talk generally about "late amplitude" and "difference between attended and unattended responses".

Reply: Point taken, we should be more clear about the EEG-part of the study. We addressed the issue in the current version by clearly stating that the auditory measure is part of the EF-measures and based on Neville et al.'s paradigm as described in the 2013 paper (reference number 4). You can check the reformulations in the Background section, lines: 79-80, 196-198 and in the section Intervention, line 237-238. See also reply to Reviewer 2, question 3.

4) I understand that maybe this was not central in your study, but since you mention it, I would expect some words about this part (or you could omit this result from your report here, since I saw that you're going to discuss it in a deeper way in a paper to be published in Frontiers). So maybe you could shorten part of the description of the protocol by
simply calling out to the previous paper where you describe it very well, so you can spend words for more detailed content about the specific topics addressed here.

Reply: Yes, maybe we could refer more back to the protocol but after having discussed this in the group we believe that it is cumbersome for readers to NEED to go back and read another paper to fully understand the present one.

5) I would also advice discussing only important results which add interesting information to our knowledge on the topic and refer to hypotheses or research questions, because the "Discussion" section is complex and difficult to follow.

Reply: One reason for introducing so many aspects of the (null) results, and discuss them thoroughly, is that we imagine that the audience potentially interested in this study come from very different backgrounds. While some are into EEG and statistical methodology, like yourself, others might have an interest mainly in educational aspects. With the hope of being inclusive in the address, we chose to keep the various angles in the discussion and hope you can see our reason for doing this.

Reviewer 2

1) Even though this is a big study that has been competently developed and conducted, a main criticism has to do with the lack of results. This brings into question the main assumptions of the study. However, this reviewer believes that studies with negative results should be published, provided that data comes from appropriate designs and are competently analyze. Using mixed model regressions seem an appropriate strategy for analyzing data here. However, the authors seem to obviate issues related to the statistical power of their sample that should be considered (and analyzed) in a revised version of the manuscript. Also, a Bayesian approach might shed some light on the null findings.

Reply: We have added discussions about power in terms of effect sizes and sample sizes related to intervention differences, and we think these additions present the null results in a more nuanced manner. First, we discuss raw differences between intervention groups in the Results section, subheading Intervention effects, line 815-825, then power is briefly mention in the first paragraph of the Discussion section (lines 885-887) and also in the Discussion section regarding ERP results (lines 994-1005).
2) Whereas some of the predictions are based in the results of prior studies (e.g., SES effects on language, EF and socioemotional comprehension), other seem rather based in authors' intuitions (e.g., correlations between a high score on language tasks pre-intervention and socioemotional comprehension; higher SES for Swedish children compared to children whose Swedish is less dominant,…). These hypotheses should be also grounded on prior findings.

Reply: Good point, and were possible we have grounded hypotheses on prior findings. As for the prediction of a relation between high language scores and socioemotional comprehension, it is based on the assumption that a child who understands instructions, story narratives and test leaders’ questions will also be better able to answer to these questions appropriately, than would a child with less language skills. We have now added references justifying this, see line 316-319. As for the prediction concerning children with Swedish as L1 having in general a higher SES, it is based on the demographic situation in the studied area as many refugee families have arrived in the last couple of years and therefor are not fully integrated in the community (regardless of educational background etc). These families will attain a lower SES based on income and occupation.

3) Regarding the EEG study, the authors state that active electrodes were re-referenced to the mastoids. They should be more precise while reporting this information: average of the mastoids or linked mastoids?. How many channels (on average) were interpolated? A 50% rate of epoch rejection is a very high one, did the authors correct for ocular artefacts?. Indeed, a lot of information is missed from this study: the number, duration and characteristics of the stimuli; the background for the expected ERP effects (why did the authors analyze only activity between 100 and 200 ms?, which component?, Why did they include all the electrodes in the analyses?, which are the basis for creating a selective attention variable after subtracting activity for attended and unattended responses?, Why did the authors suddenly analyze activity between 300 and 400 ms?. Finally, the discussion of these data on p. 56 and 57 is too vague (what kind of general demands in terms of processes?, What kind of differences between pre- and post-intervention sessions that were not intended?).

Reply: Good points, the original text relied too much on the protocol paper in this respect. We have added quite a bit more information on both pre-processing (see e.g. section Testing procedure, lines 568-569 regarding bad channels and interpolation; line 570 regarding ocular artifacts; rejection rate lines 586-590)) and the experimental paradigm (see e.g. lines 551-562). Regarding the high rejection rate per epoch: It is a rather high rejection rate, even for children in this age, were a high rejection rate is common. In previous published work with the same paradigm rejection rates are not always stated, but number of trials kept is. We have 82% of the number of trials in Coch et al (2005) when testing older children (6-8 years), and 42% of the
The high rejection rate is unfortunate but in some respects compensated by our very high number of participants, and two recording sessions. Having visited Neville’s lab, I think these differences reflect more time spent per participating child, more emphasis on sitting still during recording, and perhaps an environment with less distracting visual background. We opted for shorter and, for the child, less constrained sessions, and recorded in diverse settings at pre-schools, and the high rejection rate is the price we pay. On the other hand, having 80% of the trial numbers of Coch (2005) is not a catastrophic difference, and many ERP-studies use much lower number of trials for their averages.

Regarding which component we study, choice of electrodes and time window, use of difference waves: Most answers to these questions have to do with previous comparable studies such as Coch 2005, Stevens 2009 and Neville 2013. Children of this age typically do not show the P1-N1-P2 component complex typical for adults in response to auditory stimuli, instead there is typically a less articulated component described usually as ‘broad positivity’. This is especially so in complex sound-environments such as the present one (but also in multi-feature MMN-paradigms). This component, and differences in it due to attention, starts around 100ms and can continue a bit after 200ms. However, in the studies we compare with, two use 100-200ms and one 100-225ms, and in our pilot data this was the approximate timing of the peak and attention effect. The main points from this description is added to the manuscript in the Results section, subheading Auditory Selective Attention, at lines 702-705. There were some omissions regarding the description of the selective attention variable in the protocol article. One of them is failing to mention this time window at 100-200 ms. The other is which electrodes should be used. Choice of electrodes was motivated by comparisons with previous literature, especially Neville 2013 were the intervention effect is seen in parietal and occipital electrodes. This made us include these electrodes as well as frontal and central electrodes in the ERP ANOVA. Given the interaction between electrode site and attention effect, which reflect a maxima in frontal and central electrodes, we did use only these electrodes in the difference wave. The use of difference waves is standard ERP practices, and was specifically motivated by the need to adapt the ERP-measure to the same form of regression analysis which we wanted to use for its control variables and to have a comparable analysis of the ERP measure. In retrospect the regression analysis did not give very interesting results, but in our planning of the analysis this seemed like a good option.

The analysis of the late time window was indeed an unexpected, exploratory analysis. We discovered this effect in the ERP averages, and felt that the description of the data would not be complete without an analysis of this part, especially since attention effects for older children and adults can be negative with similar topography and timing. We failed to mention this analysis in the section Analytic strategy, but now it is added at line 619-620. Additionally, the rationale for the analysis of the late time window is clarified in the Results section, subheading Auditory selective attention (line 717-721)
The discussion of processing demands is clarified (Discussion, liner 1235-1238) and the part about differences between pre and post sessions is changed because it was not consistent with results (lines 1242-1244).

Editor’s comments:

There are parts of the manuscript which are basically the same of the previous one published here:  https://bmcpsychology.biomedcentral.com/articles/10.1186/s40359-018-0239-y. For instance, the 7 research questions which motivated the study. I think that the methodology is fine, but maybe they could focus on some aspects, remove duplicated things and add new information about things that they don't mention at all (for example, justify the use of EEG and ERPs in the study of auditory selective attention in children with a short discussion of what found in literature and so on).

Reply: We have added, hopefully clarifying, information on the EEG paradigm and why we used it, see replies to Reviewer 1 and 2 for specifics. As for the research question and some other duplicate information, we argue that the paper needs to be able to stand on its own. In particular, it appears odd to remove or refer readers to an earlier paper in regard to research questions. We also could not reformulate those, as we need to keep to the actual study layout. Maybe we did not understand the comments the way they were intended? In that case, please get back to us on how we can solve this duplication business.