

Peer Review File

Article information: <https://dx.doi.org/10.21037/mhealth-22-13>

Reviewer & Comment #	Comment	Reply	Changes in text
A1	Sample sizes were extremely small	We agree. We added a statement in our limitations section.	Limitations, Line 375-378: Though we initially recruited 100 study participants as planned for this pilot, many of those initially recruited in a clinic setting did not continue with the study outside of the clinic setting resulting in a smaller sample size than originally planned.
A2	Retention was extremely low	We agree. This problem was already included in our limitations section lines 378-379.	No changes
A3	Attrition was differential	We agree. This problem was already included in our limitations section lines 378-379.	No changes
A4	Analyses were not all run as intent-to-treat but rather as-treated	We did conduct an intent-to-treat analysis including all those who received the EDL app in the intervention group regardless of their actual use of the app. We think this reviewer may have thought otherwise because we were not able to collect data on all 100 participants who were initially recruited in the clinic setting and because of some of the wording we used in the discussion.	Methods, Lines 196-198: All analyses were conducted as intention-to-treat, meaning that those who were randomized to receive the EDL app were analyzed in the intervention group regardless of their actual use of the app. Results, Lines 232-236: We conducted intention-to-treat analyses; however, low participation in the data collection visits outside of the clinical setting prevented us from including all 100 of those who initially agreed to participate in the study because we lacked data from those who did not participate in the data collection visits. Throughout discussion: We changed 'received the EDL app' to 'randomized to receive the EDL intervention'.

A5	Randomization did not appear to be successful in the first place (given very different levels of the outcome across randomization groups at baseline)	We did use an acceptable, computer-based algorithm for randomization. The baseline characteristics of the 100 study participants show that they were comparable. The issue arose with the lack of data that we were able to collect in subsequent data collection visits. Since this was a pilot study, it is a learning that we would apply to future, larger studies. The methods already described the use of a computer-based algorithm. We expanded on the limitations of our study and recommendations for future studies in our limitations section.	See changes described above for comment A4. Discussion, Lines 381-384: For future studies, we would consider routinely conducting data collection directly before or after well child care visits in the clinic setting when possible given the difficulty we experienced with collecting data outside of the clinic setting. This difficulty resulted in small sample sizes, baseline differences for the outcome of interest between the EDL intervention and control groups, and differential attrition.
A6	It isn't clear how their qualitative analysis was performed: type of analysis, double-coding, list of interview questions; How were the interviewees selected and did the authors continue interviewing until they reached saturation?	We edited the manuscript to provide additional information regarding the recruitment of interviewees and qualitative analysis. We have included a copy of the interview guide and would like to have it included as an appendix if that is an option.	See Qualitative Analysis section, lines 206-225.
A7	How many of the interviewees used the app and how much?	We appreciate this suggestion and included some data regarding the interviewees use of the app.	Results, Lines 276-283: Among the 13 study participants who received the EDL app and agreed to participate in semi-structured interviews, all 13 set a reading goal (they were prompted to do this when they first downloaded and opened the app) but only 5 used the app enough to earn one or more rewards. Among the 5 who earned rewards, 4 received 4 or more rewards (300 or more points) indicating regular engagement with the app, and 1 received only 1 reward (50 points) indicating limited engagement with the app. The remaining 8 of those interviewed did not receive any rewards (<50 points) indicating minimal engagement with the app.

A8	Some of the endpoints for this study seem confusing. The READ measure makes a lot of sense, since the intervention targeted reading behaviors. But the PIDA and PVR measures seem to assess things that were not targeted by the intervention. If the authors expected the intervention to affect these latter outcomes: why? More information on the theory there would be helpful.	The content of the app focused on 4 categories of child-development promoting behaviors: talking, reading, playing, and praise. In addition to the reading log, parents/caregivers could earn points for doing talking, playing and praise activities. We hypothesized that the PIDA (parental involvement in developmental advance) and PVR (parental verbal responsivity) scores may change because the app did target these behaviors measured by these scales.	We clarified that the app emphasized talking, reading, playing, and praise. -Description of EDL content and features, lines 112-114 and 129-131. -Description of Measures, lines 177-180.
A9	The way sample members were selected seems potentially difficult for analysis, at least if I am understanding it correctly (though I may not be!). In particular, participants were selected from a clinic that already participates in the Reach Out & Read and Healthy Steps programs. It seems as though it is possible that folks going to these clinics are already being served by a reading support program and thus have no real motivation or need for the app. This would both limit generalizability to parents of children who attend a clinic with existing reading support and also might artificially deflate the effect of the intervention (since a lot of people in both the treatment and control group may already be "treated" by a reading support program). Do the authors have any information on whether sample members were already participating in these programs? Or are these programs considered "developmental therapies" and these people were excluded?	We edited our manuscript to clarify that those who received ROAR and/or Healthy Steps and our reasons for not excluding them. Our discussion already included a suggestion that a mobile health intervention to promote child development should be used in combination with other interventions.	Recruitment and Participants section, Lines 143-146: We did not exclude children who received Reach Out and Read books and/or Healthy Steps, because we aimed to conduct a pragmatic trial that reflects the 'real life' context for families. We wanted to understand if the EDL app could augment the care already provided by pediatric primary care clinics.

B1	I question the term “gamification” in this study. Based on the interpretation I gathered from the methodology and findings of the study, the author(s) were administering a “Points and Reward Tracking system”, not a gamified system. I recommend rewording to align with what was conducted in the study. Although the definition provided for Gamification may be supported by goal setting, logging, points, and rewards, the general understanding of “gamify” in general population includes an additional element of direct engagement and interaction with the content on the screen. I recognize the direct engagement/interaction for this study was the # of minutes a person was reading to/with their child, but this does not stimulate dopamine and the desire to “want more” of the stimulant such as a more interactive “gamified” experience.	We agree with this comment. We removed references to gamification and clarified that the app included a points and reward system.	We removed references to gamification throughout the manuscript.
B2	Have you considered analyzing age in relation to attrition to see if there is an underlying mechanism of disconnect between one’s age and the task at hand?	This is a great suggestion. However, we did not measure the age of the primary caregiver. We did measure the age of the child participant, but all children were between 12 and 15 months (within 3 months of each other) at study enrollment.	No change
B3	Check the document throughout for an extra space between sentences.	Done.	Done.
B4	Pg. 8, line 263-272. I question the accuracy of this statement. Particularly from line 267-272. To indicate there was an increase between baseline and follow-up you would need the to compare the same # of baseline participants. The increased percentage may only be increasing because you are now only comparing those that were more vested in the study from the beginning. Recommend rewording this section.	Agree. We reworded this section.	Results, lines xxx: Among those for whom we had both baseline and follow up STIMQ subscale scores, we found that the PIDA and READ scores increased from baseline to follow up for both groups (READ p=0.001 EDL, p=0.004 IP; PIDA p=0.004 EDL, p<0.0001 IP).
B5	Pg. 9, Line 280-287. How many persons provided conversational feedback?	We edited the manuscript to clarify. Also see responses to Reviewer A’s questions regarding the qualitative analysis.	Results, lines 276-283: Among the 13 study participants who received the EDL app and agreed to

			participate in semi-structured interviews, all 13 set a reading goal (they were prompted to do this when they first downloaded and opened the app) but only 5 used the app enough to earn one or more rewards.
B6	Pg. 9 App feedback. Did you administer a Likert Scaled questionnaire in addition to the Q&A session? If so, please include this information. If not, this is a component necessary and needs to be added to future studies.	We did not use a Likert scale. As noted above, we did clarify the qualitative methods that we used. In the future, we will consider using Likert scale type questions as suggested by this reviewer.	See changes made in response to Reviewer A.
B7	Also consider gathering information on the software version of the smartdevice (iPhone 4 is going to be less compatible with app software platform than a person using an iPhone X or newer). The age of the smartdevice may have cause many of the issues expressed by participants in this study.	This is a great suggestion. Will do this in the future.	No changes needed.
B8	Pg 12, Line 392 – add “and” between ...college education, “and” did not...	We revised this sentence.	Discussion, Lines 394-396: However, we also found that the majority of the low-income caregivers with less than college education whom we recruited for the study did not engage with the app.
B9	Pg 12, Line 392-395. The first sentence, “Common barriers...” is left hanging and needs additional context. These last to sentences read oddly and need to be reworked.	We revised these sentences.	Discussion, Lines 396-399: The most common barrier to engagement was difficulty accessing the app. Some of the features that caregivers who did access the app found to be most be useful could be delivered via text messaging. Therefore, focusing on text messaging may be a solution to overcome the access barriers to using an app.
B10	Pg 17, Table 1. Fix the word Latino under Ethnicity.	The use of Latinx was purposeful. Our research group is moving toward using gender-neutral language. While there is not consensus among the Spanish speaking community, Latinx or Latine are the preferred gender-neutral terms. Our team, which includes native	No changes needed. We are happy for the editors to change this if the journal has a specific preference.

		Spanish-speakers, was most comfortable using 'Latinx'.	
B11	Pg 18, Table 3. Format this table for consistency of layout of the data output.	Done.	Done.

C1	It might be helpful to engage parents in using the app earlier (i.e., at first visit), and have a follow-up visit aimed at troubleshooting issues that are interfering with engagement.	We agree with statement and will consider for futures studies.	No changes needed.
C2	Given the target group for this intervention, it may be beneficial to have multiple family members engage in using the app or integrate a social element within it.	We agree with statement and will consider for futures studies.	No changes needed.
C3	Providing a phone and/or data plan for the parent to use specifically for the study could help reduce the number of access and technical issues. Data plans that run out at the end of the month or slow connectivity until the user pays for more data may have been a barrier to consistent app engagement.	While we agree with this statement, our goal was to conduct a pragmatic trial that reflects the real-life context of the families most likely to benefit from this intervention. Therefore, we chose not to provide a phone or data. We added a sentence to the discussion to explain this decision.	Discussion, Lines 399-403: We considered providing phones and/or data plans to study participants to decrease m-health access barriers, but chose not to because of our desire to conduct a pragmatic trial that reflected the 'real life' context of the families we thought were most likely to benefit from the intervention.
C4	Use of text message to deliver some of the content, as the authors suggest, also seems very promising.	Thanks for this comment.	No changes needed.