**JoVE Revision**

**Editorial comments:**

The manuscript has been modified by the Science Editor to comply with the JoVE formatting standard. Please maintain the current formatting throughout the manuscript. The updated manuscript (55596\_R0\_102116) is located in your Editorial Manager account. In the revised PDF submission, there is a hyperlink for downloading the .docx file. **Please download the .docx file and use this updated version for any future revisions**. The updated manuscript is also attached.

**• JoVE is unable to publish manuscripts containing commercial sounding language, including trademark or registered trademark symbols (TM/R) and the mention of company brand names before an instrument or reagent. Please remove all commercial sounding language from your manuscript (text and figures). Examples of commercial sounding language in your manuscript are: cheerios, etc. All commercial products should be sufficiently referenced in the table of materials/reagents. Please replace all commercial sounding language in your manuscript with generic names that are not company-specific.**

We have removed all commercial sounding language. For the Table of Materials, we contacted EGI and they do not provide catalogue numbers for their products and their website is not designed to easily browse their different products. Therefore, we included the specific names of all EGI materials so researchers can contact EGI to purchase the products we used.

**• Please remove the trademark symbols (TM/R) from the Table of Materials.**

Trademark symbols have been removed.

**• Please use h for hour(s), min for minute(s) and s for seconds throughout the manuscript (including figures and tables).**

We have made this adjustment.

**• In the Long Abstract (150-300 words), please include a statement that clearly states the goal of the protocol. For example, “This protocol/manuscript describes…”**

We have added a statement that clearly states the goal. We write, “This protocol describes a paradigm that is designed to comprehensively assess infant EEG activity in both social and nonsocial contexts as well as tease apart how different types of social inputs differentially relate to infant EEG.”

**• Please ensure that all text in the protocol section is written in the imperative tense as if you are telling someone how to do the technique (i.e. “Do this”, “Measure that” etc.). Avoid usage of phrases such as “could be,” “should be,” and “would be” throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a “Note”, however, notes should be used sparingly and actions should be described in the imperative tense wherever possible.**

**• Please re-write the following in the imperative tense: 1.3.4, 2.2.1, 2.2.2, 2.2.3, 2.3.2, 2.3.3, 2.3.4, 2.3.5, 2.3.5.1, 2.3.8.2, 5.1**

We have adjusted all text in the protocol section so that it is written in the imperative tense and have adjusted all points noted above.

**• Line 153- Ethics statement: Please mention the name of the University (review board).**

This information is now in the manuscript.

**• Step 1.2.4: Please note that this step cannot be filmed. Please remove the highlighting.**

We have removed the highlighting.

**• 1.3.1: How is the infant’s head measured?**

The infant’s head is measured with a soft measuring type. We have added this information to the protocol.

• **1.3.9: How is the electrode impedance assessed? How is the electrolyte solution administered?**

Electrode impedance is assessed with the EEG recording software. The electrolyte solution is administered using the pipettes (which were filled with electrolyte solution) and squeezing a few drops of solution onto the electrodes with poor contact. We have added this information to the revised manuscript.

**• 1.3.10: How is the information saved?**

The way that this information is saved likely varies by EEG system. In Netstation, there is a button to “save” impedance information. We have added this information to the protocol.

**• 2.1.1: Please provide a reference for recording data.**

For this study, we use equipment from Electrical Geodesics and followed their specifications for EEG recording. We can cite their manual, but did not do so previously because of the limitation of including manufacturer names in the manuscript. Please let us know how to proceed.

**• 2.2.2: Please re-write in the imperative tense or make this a Note.**

We have made this information a note.

**• 2.2.4: Please make this a Note.**

This is now a note.

**• 3.3: How is this done?**

Our paradigm is designed such that our EEG recording software (Netstation) automatically closes after the paradigm is completed. The recording file is automatically saved during this process. We have adjusted this in the protocol.

**• Section 4: Please provide a reference for EEG data processing.**

We have added references.

**• 4.2: If this step is to be filmed, please add stepwise detail on how to segment the raw data and exclude artifact. Please note that data analysis using only graphical user interface can be filmed. Alternatively, provide a reference and remove the highlight.**

We have removed the highlight for this step. We included a reference for the “EEG Data Processing” section, which is applicable to all sub-steps.

**• 4.6: This step (calculation) cannot be filmed.**

We have removed the highlight for this step.

**• Section 5: This section cannot be filmed as there are no action items. This section should be moved to the Results section.**

We have moved this section to the results section.

**• Section 6: This section does not use a graphical user interface. We are unable to film coding and calculations. Please un-highlight this section and provide appropriate references. If possible, please provide the codes used in your studies as a Supplemental code file (via Editorial Manager).**

We apologize for not making this section clear. We used a behavioral coding software, which does have a graphical user interface. The researcher watches the video of the infant during EEG recording and manually assigns behavioral “codes” to where the infant was looking. Codes refer to not computer code, but to different classifications of looking behavior. In other words, the software is a way to log events during the video. For example, in this study some of the codes/events included: “infant looking at left screen”, “infant looking at right screen”, “infant looking at experimenter.” Therefore, the researcher would watch the video and when the infant looked at the right screen, they logged appropriate code/event and when the infant looked at the experimenter, they logged the appropriate code/event. The end result is a list of behavioral looking codes/events that are matched with the video. The software then allows for analysis of this coded data/events so it is possible to assess how long the infant spent looking at different things.

**• After you have made all of the recommended changes to your protocol (listed above), please re-evaluate the length of your protocol section. There is a 10-page limit for the protocol text, and a 3- page limit for filmable content. If your protocol is longer than 3 pages, please highlight (in yellow) 2.75 pages or less of text (which includes headings and spaces) to identify which steps should be visualized to tell the most cohesive story of your protocol steps. See JoVE’s instructions for authors for more clarification. Remember that the non-highlighted protocol steps will remain in the manuscript and therefore will still be available to the reader.**

We have highlighted 2.75 pages of text.

**• If you are re-using figures from a previous publication, please obtain explicit permission to re-use the figure from the previous publisher (this can be in the form of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."**

We have uploaded a word document including an email that says we do not need permission to use the adapted figures.

**• Figures 2-5: Please provide these figures as tiff or pdf files.**

We have attached the figures as pdf files.

**• Please define the error bars (SD, SEM, etc.) in the legend.**

We have defined the error bars as standard errors in the legend.

**• Please include a supplemental figure legend (below figure 5).**

Are you referring to a figure legend for the supplemental files that include a “List of the stimuli used in each block” to accompany the paradigm as well as the stimuli files? If so, we have added this below figure 5. We are happy to adjust if this is not what you are looking for.

**• Please ensure that your discussion covers the following in detail and in paragraph form: 1) modifications and troubleshooting, 2) limitations of the technique, 3) significance with respect to existing methods, 4) future applications and 5) critical steps within the protocol.**

We have made sure the discussion covers all topics listed.

**• References: Please abbreviate all journal titles.**

We have abbreviated all journal titles.

**• JoVE reference format requires that DOIs are included, when available, for all references listed in the article. This is helpful for readers to locate the included references and obtain more information. Please note that often DOIs are not listed with PubMed abstracts and as such, may not be properly included when citing directly from PubMed. In these cases, please manually include DOIs in reference information.**

We have added DOIs to the references, however the following did not have DOIs: Nichols et al., 2005; Jones et al., 1997.

**• Please take this opportunity to thoroughly proofread your manuscript to ensure that there are no spelling or grammatical errors.**

We have proofread the manuscript.

**• Grammar:**

**-Please revise the protocol text to avoid the use of any pronouns (i.e. "we", "you", "your", “our” etc.).**

**-Step 2.3.5: responds contingently, continued appearing.**

**-2.3.5: Please correct the run-on sentence.**

**-2.3.7: directs the infant’s**

We have corrected all grammar comments.

**Reviewers' comments:**

**Reviewer #1:**

***Manuscript Summary:***

**This manuscript presents a new method for measuring infant EEG in social and nonsocial contexts. The study has several strengths. First, live presentation is used, which has been demonstrated previously to be critical in maximising social brain activity. The authors present several conditions varying in key parameters (e.g. social input, language etc), to allow greater finesse in interpretation than has been possible previously. This area has also traditionally suffered from a lack of standardization (as the authors point out); whilst standardised social/nonsocial videos are becoming more available, live paradigms are still very variable.**

*Major Concerns:*

**1. The set-up (with two screens) is perhaps challenging for some labs to accomplish - it might be worth discussing any possibility of simplification with just one screen?**

Thank you for this suggestion. Simplifying the paradigm to be used with just one screen is possible and we have added this modification to the discussion (see p.14, lines 699-705). The main consideration in using one or two screens is whether the researcher would like to look at concordance of the infant following the experimenter’s directions to look at the left or right screen. If concordance is a main question, then two screens are necessary to assess whether infants are following the experimenter’s instructions. We were interested in assessing whether infants complied with the experimenter’s directions, so we included two screens. Thus, we were able to compare how often the infants looked at the correct screen (the one the experimenter directed their attention to) and the incorrect screen. If the researcher is interested in how much infants are looking at the one screen, it would not be possible to attribute this to joint attention. However, if interests do not involve concordance with gaze, then a single screen would be fine to use.

**2. The authors could expand the relevance of this paradigm to studies of populations at high risk for social difficulties, eg. ASD, where resting state is often looked at with sub-optimal paradigms.**

We appreciate this suggestion. We have expanded the relevance of this paradigm to populations at high risk for social difficulties (see p.16, lines 790-797).

**3. Consider suggesting that EEG is also analysed during epochs of visual attention to the social or nonsocial content of the probe (not just overall during the condition)- this may allow investigators to more specifically pull out social brain activity.**

Thank you for this recommendation. We have added as a modification that it would be possible to analyze the data depending on whether the infant is looking at the screens or the experimenter (see p.14, lines 707-709). However, the joint attention and social engagement conditions still have social content, even when the infant is looking at the screen. For example, during the joint attention condition, if the infant is looking at the same screen as the experimenter, this is part of the social experience of joint attention. Also, in both of these social conditions, the infant hears the experimenter’s voice. Therefore, even if they are not looking at the experimenter, there is still social content to the interaction.

**4. Muscle/eye movement noise should be considered - this will be important in conditions where lots of shifting attention is expected (e.g. joint attention), This will contaminate higher frequencies and needs to be explicitly considered.**

We have added a step to the protocol, which outlines our artifact rejection parameter (see p. 9, lines 466-468).

**5. The interpretation of higher 4-6Hz power as 'less activation' is puzzling - typically theta power is interpreted as greater being greater activation. This should be discussed in the manuscript, with references to support the interpretation and selection of different frequency bands.**

This information can be found in the “Representative Results” section before we discuss the EEG power results (see p.12, lines 589-592).

**6. The power values look really high - can the authors contextualise these with respect to previous work, in terms of what investigators should expect? Usually I would expect more like 2ish for ln(power) in the alpha range, and 3ish in the theta range. Sometimes very high values can reflect substantial artifact so this needs to be considered.**

Our values are consistent with past research (Tierney, Gabard-Durnam, Vogel-Farley, Tager-Flusberg, & Nelson, 2012) that used comparable EEG recording and processing methods to compute EEG power (high-density net referenced to Cz and then re-referenced to the average reference). There is much variability in infant EEG research in terms of the EEG system used for recording as well as processing choices such as artifact parameters and choice of reference, which can affect EEG power values. Further, there have not been enough studies with identical methods to establish concrete norms of expected values. In the representative results section, we explain that our results are consistent with data processed in the same way, but that values may vary depending on the processing methods used (see p.13, lines 623-627).

***Minor Concerns:***

**7. Segment lengths should be specified.**

We have now specified the segment lengths we used in the protocol (see p.9, lines 462-464).

*Additional Comments to Authors:*

I would like to commend the authors for working on this paradigm - this is a challenging area and I think this study makes an important contribution.

**Reviewer #2:**

***Manuscript Summary:***

**N/A**

***Major Concerns:***

**N/A**

***Minor Concerns:***

**What ages is this paradigm most appropriate for? Do the different social and non-social conditions need to be modified due to infant age?**

Thank you for this question. Although we have only tested this paradigm on infants ranging from 11-14 months, we anticipate that this paradigm would be appropriate for 6 months through 24 months of age and that the conditions would not need to be modified. When considering the youngest appropriate age, the main consideration is that infants are old enough that they are able to follow gaze and thus are able to engage in joint attention. Six-month olds have established joint attention (Morales et al., 2000; Morales, Mundy, & Rojas, 1998), so this would most likely be the youngest age that this paradigm is appropriate for. On the older side, 12-24 months is the age range when joint attention is particularly important for language learning (Markus, Mundy, Morales, Delgado, & Yale, 2000; Mundy et al., 2007). If using this paradigm with younger ages, it is possible that the blocks may need to be shortened to maintain infant interest. We have added this information to the discussion (see p.14, lines 690-697).

**In the representative results, a sample size of 70+ infants was analyzed. What sample size is appropriate for this task?**

The sample size appropriate for this task will vary depending on the research questions being asked and the analyses planned. The main consideration is expected attrition, thus accounting for how many infants will likely have useable EEG data. Of the 85 infants with whom we successfully recorded EEG during the paradigm, 73 had useable data in at least one condition (85.88%). Within these 73 infants, 78.1 % (57 infants) had useable EEG in the nonsocial condition; 71.20 % (52 infants) had data for the language-only condition; 91.80 % (67 infants) of infants had useable data in the joint attention condition; and 87.85 % (63 infants) had useable data in the social engagement condition. Thus, researchers would need to take this level of attrition into account when deciding on their sample size for EEG analyses.

**Title on cover page is different from title on 1st page.**

We apologize for the inconsistency; we have corrected the title so that it is the same on all pages.

**In the analyses sections, were multiple comparisons controlled for? Or were only the 4-6Hz and 6-9Hz frequency bands analyzed?**

Only 4-6 Hz and 6-9 Hz frequency bands were analyzed, as these bands have been widely used in infant research (Calkins, Fox, & Marshall, 1996; Henderson, Yoder, Yale, & McDuffie, 2002; Marshall, Bar-Haim, & Fox, 2002; Mundy, Card, & Fox, 2000). Within a given model, we used Bonferroni corrections.

**EGI is specifically mentioned and used in this paradigm, would this paradigm be appropriate for other EEG systems?**

Yes, this paradigm would be appropriate for other EEG systems. We have added a “note” in the protocol that addresses this point (see p.10, lines 487-489).

*Additional Comments to Authors:*

N/A

**Reviewer #3:**

***Manuscript Summary:***

**The major strength of this manuscript and proposed methods video is the nicely articulated argument for disentangling the neural correlates of social behavior in infants by using four carefully designed conditions rather than what may be a biased baseline: (1) nonsocial, (2) joint attention, (3) language-only, and (4) social engagement. The authors show that their paradigm examining EEG power using a within-subjects design in a sample of 73 12-month-olds is able to discern differences in neural activity across these four conditions. My suggestions are directed at clarifying steps in the procedure, and adding additional potential applications for the protocol.**

**Procedural:**

**- The authors could consider adding some detail about techniques for recruiting infants for EEG studies that they have found to be successful, such as how EEG is discussed with parents and any incentives used.**

Thank you for this suggestion. We have added a section to the protocol outlining techniques for recruiting infants for EEG research (see p.4, lines 171-190). We have also added additional details to the protocol on how to discuss EEG with parents when they are in the lab for a visit and incentives that are helpful (see p.5-6, lines 228-268).

**-Was a pointing-only condition considered? I understand it may be difficult to squeeze in a fifth condition type, but perhaps it could be a suggestion for future work, as pointing cannot be parsed out from other elements of joint attention.**

This is an excellent idea. We did not include a pointing-only condition in this initial protocol, given that there was a limit to how many conditions we could include and still get the majority of infants through the paradigm. However, we certainly agree that this would be a helpful future direction to parse how infant EEG power may vary by the specific elements of joint attention. We have added this to the discussion of future directions (see p.16, lines 776-779).

**- Section 2.3.8 starting on line 305 suggests that the experimenter only needs to talk behind the screen. However, Table 1 has L/R directions of where the experimenter should look in the language-only condition. Please reconcile.**

You are correct that the experimenter is talking from behind the screen, we apologize for the confusion. We have removed the L/R directions from the language-only condition.

**Application:**

**- It may not be necessary for this type of paper, but the authors could state why they targeted 12-month-olds for their pilot. I would be interested in their thoughts as to whether this procedure would work in younger infants in particular. More broadly, what is the age range that the technique could be or should be used?**

We targeted 12-month olds for our pilot, as this age is a time when joint attention is particularly important for language development (Markus, Mundy, Morales, Delgado, & Yale, 2000; Mundy et al., 2007). Therefore, we thought that assessing EEG during joint attention at this age would be especially relevant.

We have added this information to the introduction (see p.3, lines 136-140). Reviewer two also had a question about what ages would be appropriate to use for this protocol, see our detailed response above on p. 6. In sum, this protocol is likely appropriate for 6 month olds-24 month olds.

**- The authors could consider some brief discussion on how to expand their method from dyadic to triadic social interactions.**

We are not sure what type research questions the reviewer has in mind, but we are happy to include any suggestions you have. One consideration may be the logistical difficulty of a triadic interaction, as it would likely involve an additional experimenter in the booth. We worked to make the experience of each infant as similar as possible, which may be more difficult in a triadic interaction.

**- The authors may want to take a look at Gonzalez et al. (2016) Front. Psychol. 7:216, which is focused on motor, rather than social, development but discusses using EEG measures like power as well as coherence and mu desynchronization to make broader connections across domains, and could be used in support of their concluding remarks on p.14.**

Thank you for this helpful reference. We have added a discussion of the importance of assessing EEG using multiple measures in multiple domains to broaden our understanding of infant development (see p.16, lines 765-771).

**Minor Comments:**

**- p.1, line 45: should be "In the nonsocial condition…"**

**- p.4, line 183: I would also add 'and change' the infant before starting.**

**- p.6, line 264: For better readability, I would suggest one sentence with the open curtain instruction and a new sentence with the closed curtain instruction.**

**- p.11, line 499: "on the net" is repeated in the sentence**

**- p.19: Step 8 is missing a closing bracket in block 5 (Joint attention).**

We have corrected all of the minor comments cited.

*Major Concerns:*

N/A

*Minor Concerns:*

N/A

*Additional Comments to Authors:*

N/A

**Reviewer #4:**

***Manuscript Summary:***

**The paper describes an electroencephalography (EEG) experiment where infant participants' brain activity was recorded in conditions with or without social interaction. Brain activity was indexed with EEG power and was shown to depend on the social interaction in the experiment. The study is valuable in providing novel methodology and evidence for infant brain research using realistic face-to-face communications to study the development of social neurocognition. A video-based demonstration will greatly add to the transparency and replicability of the study as well as to the application of the method to further research using social settings in EEG acquisition. I suggest that the paper will be accepted after minor revision according to the points detailed below. The Introduction gives adequate background and motivation for the present study. The hypotheses are stated and related to the differences between experimental conditions.**

**The protocol is methodologically sound, adheres to ethical standards and is described in great detail. However I was somewhat concerned about the temporal order of experimental conditions. Starting from line 272: "Blocks follow the sequence: social engagement, nonsocial, joint attention, language-only, joint attention, nonsocial, social engagement, and language-only." Does this mean that the order of conditions was the same for all participants? If so, I think this may have influenced the results and could constitute a confounding effect. I would strongly recommend permuting and balancing the order of conditions across participants.**

We appreciate this point. However a random order of the blocks would not allow for maximum useable EEG data for two key pragmatic reasons: (1) Our pilot testing showed that it was necessary to alternate between social (when the experimenter was present-joint attention and social engagement) and nonsocial (when the experimenter was absent-nonsocial and language-only) blocks to ensure that infants stayed engaged throughout the paradigm and (2) to maximize getting useable data in each condition, the first four blocks need to include all four conditions in case infants are not able to complete the protocol. Based on the challenges of conducting EEG research with infants, especially with a long paradigm, those constraints are important. However, we agree that within those constraints, the order of blocks could be permuted across participants. We have added this information to the manuscript (see p.14, lines 682-688).

*Major Concerns:*

I have no major concerns.

***Minor Concerns:***

**1. Account of movement artefacts. The experiment included conditions involving face-to-face communication and social engagement. I would believe that such conditions can arouse more bodily movement and head movement in the infant participants than for example watching stimuli on a computer screen. The authors further state on Line 531 that "infants were fussier during the conditions where the experimenter was behind the curtain". Do the authors have protocol for assessing the presence of movement during EEG acquisition and the contribution of such movement to the EEG power? Should video-based quality control be included in the "Protocol" section of the paper?**

You are absolutely correct that it is important to cut out data when there is movement artifact. To detect and reject artifacts, we used an artifact threshold (*e.g.,* data is rejected when it exceeds a certain threshold in micro-volts), which is common practice (Bell, 2002; Brito et al., 2016; Tierney et al., 2012; Welch et al., 2014). We have now included our artifact rejection criteria in the protocol (see p. 9, lines 466-468). As you noted, infants were fussier during the conditions when the experimenter was behind the curtain and this data was cut out by our automatic artifact rejection as infants had less useable EEG data in these conditions compared to the conditions when experimenter was present (see our response to your next point for more details).

**2. Number of averaged EEG data frames. How many data frames were averaged per condition in the study? How did this differ across conditions?**

Each of the eight blocks lasted 2.5 minutes and each condition was repeated twice. Therefore, EEG was recorded during 5 minutes for each condition. This data was then segmented into 30-second epochs, which were subject to artifact rejection. For an infant to have useable data in a given condition, the infant had to have at least one useable/good 30-second epoch. The average number of useable epochs for each condition was:

-Nonsocial condition = 2.60 epochs (78.08 seconds) on average

-Language-only condition = 2.75 epochs (82.60 seconds) on average

-Joint attention condition = 4.19 epochs (125.75 seconds) on average

-Social engagement condition = 3.95 epochs (118.36 seconds) on average

Thus, there was more usable data in the social face-to-face conditions (joint attention and social engagement) compared to the nonsocial conditions (nonsocial and language-only). The amount of usable data was tested as a possible covariate and was unrelated to EEG power values. We have added this information to the manuscript (see p.12, lines 594-598).

**3. Abstract. "In then nonsocial condition, infants view objects on computer screens." Should be "In the nonsocial condition…"**

We apologize for this typo, we have corrected it in the revised manuscript.

*Additional Comments to Authors:*

N/A

**Reviewer #5:**

***Manuscript Summary:***

**The authors outline an infant EEG paradigm designed to pinpoint changes in EEG activity related to specific aspects of social interactions: language, face-to-face interactions, and 'joint-attention'. They discuss the importance of having multiple conditions that are tightly controlled in order to 'tease apart' these various elements of social functioning, and the associated underlying changes in EEG.**

***Major Concerns:***

**\*The apparent variability in the social-engagement condition is concerning. It is designed to assess the property of 'face-to-face' interaction. However, the description discusses 'suggestions' for how to maintain infants' engagement that include songs, games with hand actions (itsy-bitsy spider), and 'talking warmly', peekaboo. There is tremendous variability in the auditory and visual experience of the infant across the suggestions mentioned for this condition, (song vs. speech, hand actions/gestures vs. not, potential differences in facial expressions of emotion such as neutral vs. happy) making interpretation of the eeg activity during this condition difficult in and of itself, and thus making its use as a 'control' condition for a joint attention condition much less effective. The authors' main argument is about the importance of tightly controlled conditions in order to 'tease apart' the activity specific to different aspects of social interaction. With such variability within a given condition, the confidence in the 'tightly controlled' aspects of each condition, and thus the aim of the experiment, is undermined.**

We see your concern. Our goal in this condition was to keep the infant engaged and focused on the experimenter’s face. Our description may have not made it clear that each social engagement block was standardized to include songs with hand motions, peek-a-boo, and positive affect. In piloting, we found that a sequence including multiple songs with hand motions and peek-a-boo was necessary to maintain infants’ attention (as opposed to just one song or just peek-a-boo). In other words, to sustain infant attention on the experimenter’s face for the duration of the 2.5-minute block, it was necessary to maintain novelty. Further, the experimenter was trained to always maintain the same level of positive affect throughout the block. The experimenter was contingent and did respond to the infant because we wanted contingent interaction. While it would be possible to have only one context (*e.g.,* only songs) during the block, our goal was for the infant’s experience of social engagement (*e.g.,* focused on the experimenter’s face and always hearing the experimenter) to be controlled across subjects. We have adjusted the protocol to make it clear that in each block of social engagement, the experimenter should use songs with hand motions and play peek-a-boo (see p.8, lines 374-385). Our paradigm and results show that context matters. If researchers were interested, they could zoom in more and further parse the social engagement condition into songs versus peek-a-boo, to see if there were differences in neural activation.

**\*I see an issue with 'tight control' across conditions on a global scale in this paradigm. Joint attention is indeed a complex characteristic that includes, as the authors point out, language, the presence of at least one other social agent, and some entity that engages the agents' attention. But there are many more subtle yet potentially critical aspects of this complex social interaction that may be driving eeg activity and a difference from a non-social condition, that would need to be 'controlled' in order to make specific claims about what aspects are associated with underlying changes in EEG. Subtle differences such as direct or averted eye-gaze, gaze that alternates between infant and object, etc., could be critical in eliciting changes in the EEG. The additional conditions did not systematically account for these alternative yet vitally important aspects of joint attention, and thus it is difficult to determine what any differences in EEG that occur b/w joint attention and non-social conditions really indicate.**

Your point is well taken. We agree that joint attention is a complicated social interaction. Joint attention is defined as coordinated attention with another person with respect to an object or event (Mundy et al., 2000) and we wanted the condition to be a realistic experience for the infants. We recognize that there are multiple aspects of joint attention that could be driving the difference between the joint attention and nonsocial conditions. This paradigm is a first step in understanding what broad social factors (*i.e.,* language and face-to-face interaction) could underlie differences between the joint attention and nonsocial conditions. Given the time constraints of maintaining infant attention for long periods of time while also getting useable EEG data, we did not have the capability to add in additional conditions to tease apart the different components of joint attention. We agree that there are other more subtle differences, such as alternating gaze and the role of pointing, that we did not elucidate in the current paradigm and that would be important to explore in future studies. For example, other studies could include a condition of the experimenter only alternating gaze between the infant and screen and a condition of the experimenter only pointing at the screen. These conditions would allow a further parsing of how different components of joint attention may relate to differences in neural activation. We have added this as a direction for future research (see p.16, lines 776-779).

**\*Relatedly, the authors write in lines 350-351: "joint attention includes language, face-to-face interaction, and the presence of joint attention". How is the 'presence of joint attention' actually operationalized? Is it possible that there was 'joint attention' in the social engagement condition? If both baby and experimenter were looking at the experimenter's hands as they made a 'spider crawl up a waterspout', how can it be determined that no 'joint attention' occurred in the social engagement condition?**

Joint attention was operationalized as the experimenter directing the infant’s attention to objects on the computer screens and commenting on the objects. Specifically for each trial, the experimenter made eye contact with the infant, then turned in a pre-specified direction, looked at the appropriate screen, and pointed at the object while commenting on the object. The experimenter then alternated gaze between the infant and object until the end of the trial. Joint attention was only present in the joint attention condition and was not present in the social engagement condition. We agree that in a normal social interaction, joint attention is likely to occur. However, in our social engagement condition, the experimenter kept her eyes on the infant and was trained to only look at the infant throughout the block. Thus, the experimenter did not look at their hands and kept their attention directed only at the baby so as not to create this confound. We have clarified this point in the protocol (see p.8, lines 374-385).

**\*How is it ensured that mothers' reactions to the conditions are not influencing the infants' reactions? Infants are incredibly attuned to their mothers' body language, tension etc. and sitting on the mother's lap puts the infant in close contact with the mother. Unless mothers' vision and hearing have been blocked, it is highly possible that mothers' own reactions to the conditions will influence the infant, and thus the infants' eeg.**

This is a challenging issue. We agree that it is possible mothers’ reactions relate to infants’ reactions and thus their EEG. Based on the length of our paradigm and the goal of getting useable EEG data for every infant in each condition, we needed mothers to be visually aware of their infant so that they could quickly prevent the infant from pulling on the net. Thus, we did not have mothers wear visors for the pragmatic reason of maintaining data quality. We did instruct mothers not to say anything to their infants or socially interact. It would be a reasonable modification to have the mothers wear headphones and this is in the discussion of the manuscript (see p.15, line 715). We agree that having infants on their mothers laps introduces a confound. However for our set-up, this was necessary to ensure useable data throughout the paradigm as we find that when infants sit in a high chair at this age, they are less tolerant to the length of the paradigm. We have addressed this point in the discussion (see p.15, lines 713-719). It is an interesting empirical question how mothers’ reactions may influence infant EEG and future studies could code the mothers’ body language to assess this question.

**\*I am confused about the analytic approach outlined, as well as the conclusions drawn from analyses. Seeing lower versus higher eeg power in certain scalp regions does not necessarily lead to the conclusion that brain regions underlying that portion of the scalp were 'more active' during a given condition. Volume conduction of the EEG signal makes it very difficult to determine where in the cortex signals are generated, and thus conclusions about functions of the eeg scalp activity based on functions of certain brain regions (e.g., frontal cortex as important for attention and orienting; lines 455-456) are tenuous at best.**

**Additionally, higher versus lower power is not necessarily inherently meaningful, especially when different frequency bands are considered. An interpretation of what higher and lower power means across the conditions in this experiment is also made more difficult by the less convention analysis of raw power. In more conventional event-related analyses, power is calculated as a change from non-task related EEG to task-related eeg in each condition.**

We agree that caution is needed when drawing conclusions about EEG and that it is difficult to know exactly where EEG scalp activity is generated from in the cortex. We have adjusted the representative results section to clarify that we are referring to “regional scalp power” (see…). However, the EEG power recorded from scalp regions in our study is consistent with fMRI studies (Mundy & Jarrold, 2010; Redcay, Kleiner, & Saxe, 2012; Schilbach et al., 2009; Williams et al., 2005). In addition, our interpretation of attributing functional significance based on regional EEG patterns is in line with common practice (Bell & Diaz, 2012; Fox et al., 2001; Gonzalez, Reeb-Sutherland, & Nelson, 2016; Grossman, 2015; Henderson et al., 2002; Mundy et al., 2000; Mundy & Jarrold, 2010; Paulus, Kühn-Popp, Licata, Sodian, & Meinhardt, 2013). We recognize that a caveat is needed and we have added this issue as a limitation in the discussion (see p.15, lines 737-740). Pairing EEG with fMRI or fNIRS would be an important future step to better understand links between scalp activity and underlying regions and we have added this to our future directions (see p.16, lines 770-771).

In response to your point about higher versus lower power, our interpretation of lower 4-6 Hz and 6-9 Hz EEG power indexing greater activation is consistent with infant EEG research, including studies that recorded EEG in only one condition (Calkins, Fox, & Marshall, 1996; Davidson, 1988; Henderson, Yoder, Yale, & McDuffie, 2002; Mundy, Card, & Fox, 2000). Moreover, as we were comparing four conditions, using ANOVAs allowed us to directly compare the mean level of ln(4-6 Hz and 6-9 Hz power) across the conditions, information which would be hidden by using change scores, as a change score could reflect higher power in one condition or lower power in a different condition. We agree that change scores are the convention in some event-related research, such as the mu rhythm (Cuevas, Cannon, Yoo, & Fox, 2014; Fox et al., 2015). However, our analytic approach of using ANOVAs with condition as a within-subjects factor is common practice in infant EEG research that records EEG in multiple global conditions, as opposed to second by second changes in event-related EEG (Bell, 2002; Davidson & Fox, 1982; Jones, Venema, Lowy, Earl, & Webb, 2015; Orekhova, Stroganova, Posikera, & Elam, 2006; Santesso, Schmidt, & Trainor, 2007).

**\*The difficulties in infant eeg paradigms are often addressed with many subtle details in setup and administration. For example: the type of toys used to distract infants, the timing of those distractions as an experimenter moves to execute different aspects of EEG application, the specific instructions given to parents to help them be calm during application and thus be able to keep infants calm, the ways in which parents restrict infants from touching the EEG cap. These elements of the methods are presented, but on a broad and vague scale. This is unfortunate given that these specific details often make the difference between successful and unsuccessful eeg application and recording in infant studies. For a methods journal, I would think these details would be particularly relevant.**

This is a good point, thank you for the suggestion and we are happy to include more information. We have expanded the relevant sections of the protocol by adding all relevant details including: how we explained EEG to the parent (see p.5-6, lines 228-266); the types of toys used (see p.5, lines 219-220); who used the toys to distract the infant (see p.6, lines 242-243 and 258-259); what instructions were given to parents before net application (see p.5-6, lines 228-266); and instructions for the parents for during EEG recording (see p.5-6, lines 228-266).

**\*Relatedly, analytic details may be beyond the scope of the journal, but the specific thresholds and other parameters selected for artefact rejection that determine what "good" epochs are included in analyses are critical details that novice infant EEG researchers would want and need to know. Recommendations on whether any infants or blocks should be rejected based on not enough 'good' data could also be included.**

Other reviewers had similar questions about the artifact rejection threshold. We have now included this information in the protocol (see p. 9, lines 466-468). To clarify, each block was segmented into 30-second epochs and artifact rejection was performed on each epoch. For an infant to have included EEG data for a given condition, they had to have at least a minimum of one useable 30-second epoch of data in that condition. However, most infants had more useable epochs of data for each condition. On average, infants had 2.60 useable epochs (78.08 seconds) in the nonsocial condition; 2.75 useable epochs (82.60) in the language-only condition; 4.19 (125.75 seconds) useable epochs in the joint attention condition; and 3.95 (118.36) useable epochs in the social engagement condition.

**\*It is also critical to know how portions of the eeg in which infants are not paying attention to the desired stimuli are dealt with. Analyzing EEG data only during which the infant is attending to the specified stimulus is a critical aspect of accurate and interpretable infant EEG work. Were blocks dropped in which a certain percentage of the block included infants' being distracted? These are additional critical details necessary for appropriate replication and analysis of the paradigm.**

Our research goal was to understand the larger role of recording context and to see broadly what the infant EEG power differences are based on overall context. We therefore designed our conditions to more tap differences in global recording context, as opposed to parsing EEG data based on exactly what the infant is doing. We segmented each block into 30-second epochs and ran artifact rejection on each epoch. Blocks were therefore excluded based on artifact rejection parameters only. For the nonsocial and language-only conditions, the main consideration was that the conditions be nonsocial, which was accomplished by having the experimenter hidden from view (behind the curtain). For the joint attention and social engagement conditions, we validated that infants were engaging in those conditions as intended by coding their looking behavior. For this paradigm, we were also interested in whether individual differences in looking behavior related to infant EEG power. We included the infant looking measures in the EEG models and there was no relation between the looking measures and infant EEG power in the joint attention and social engagement conditions. This may be because infants were for the most part engaged in the conditions as intended, as we set up the conditions to maximize engagement.

Your suggestion could also be of interest. For example, a researcher could cut out data when the infant is looking away during the social engagement condition. We have added this as a modification in the discussion (see p.14, lines 709-711).

***Minor Concerns:***

**\*EEG "power" is not defined up front. This is not necessarily a known construct and should be explained for readers upon its first use.**

We have now defined EEG power upon its first use (see p.2, lines 71-72).

**\*From the methods description, I think the paradigm setup is such that categories are identical across each block, and that though the specific color of a category item may be different in each block, the overall collection of colors across the 10 items within a block is identical (e.g., blue flower, green glove in block 1; green flower, blue glove in block 2). The point about identical collections of colors in every block should be made explicit so the reader understands the extent to which the viewing experience is matched across conditions/blocks.**

We have added this information to the protocol (see p.7, lines 329-335).

**\*Steps 1.3.3 - 1.3.10 are things that one would find in a training manual for their specific EEG system. So I wonder how useful it is to include here. Indeed, the instructions for EEG net preparation appear specific to EGI, or other liquid saline-based passive electrode set-ups. These instructions will not apply to active-electrode gel-based systems. It would be useful to make it more clear that the instructions are for one specific type of eeg system, or else make the instructions more general so as to apply to other types of systems as well.**

Thank you for this suggestion, reviewer 2 had a similar point. We have added a note to the protocol that explains our paradigm was developed using a high-density system with liquid saline-based electrodes and that the steps could vary depending on the EEG system used (see p.10, lines 487-489).