



DEPARTMENT OF  
PSYCHOLOGICAL AND  
BRAIN SCIENCES

INDIANA UNIVERSITY  
Bloomington

June 28<sup>th</sup>, 2018

Dear Editor and Reviewers,

Thank you for your careful consideration of our manuscript. We have revised the draft according to your comments and will address them in the pages below.

Further, per the author's request, we have removed Dr. Chen Yu from the list of authors for this manuscript.

Sincerely,

Jeremy I. Borjon  
Sara E. Schroer  
Sven Bambach  
Lauren K. Slone  
Drew H. Abney  
David J. Crandall  
Linda B. Smith

Address correspondences to:  
Department of Psychological and Brain Sciences  
Indiana University  
1101 E 10th Street  
Bloomington, IN 47405-7007  
jborjon@iu.edu

### **Editorial comments:**

Changes to be made by the Author(s):

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version.

*This version has been thoroughly proofread for spelling and grammar issues.*

2. Figure 4: Please capitalize the first word in the labels on the x and y axis (i.e., Screen, Objects).

*The change has been made.*

3. Please rephrase the Introduction to include a clear statement of the overall goal of this method.

*We reworked the introduction to make the overall goal of this method clearer, specifically lines 62-64. "The goal of this method is to capture the infant experience from a first-person view in order to understand the visual environment available to them throughout development."*

4. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents. For example: Looxcie 3, Watec (WAT-230A), Supercircuits (PC207XP), KT&C model VSN500N, SereneLife HD Clip-On, Conbrov Pen TD88, Mvowizon Smiley Face Spy Button, the Narrative Clip, McCam, LENA™, LENA Foundation, etc.

*All commercial language has been removed from the manuscript.*

5. Please revise the protocol so that all text in the protocol section is written in the imperative tense as if telling someone how to do the technique (e.g., "Do this," "Ensure that," etc.). The actions should be described in the imperative tense in complete sentences wherever possible. 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 concise and used sparingly. Please include all safety procedures, etc. Please revise 1.1-1.8, 2.1, 2.2, 3.1-3.3, as well as their sub-steps accordingly.

*We revised the protocol to ensure it is in the imperative tense and removed all "should, could and would" statements.*

6. 2.4 (including its sub-steps): The Protocol should contain only action items that direct the reader to do something. Please move the discussion about the protocol to the Discussion.

*We have moved this portion of the Protocol to the Discussion.*

7. Lines 249-253: Please move details of the methodology to the Protocol.

*We moved Lines 249-253 to the Protocol.*

8. Line 266: Please include a space between all numbers and their corresponding units: 30 Hz, etc.

*A space has been included in line 266 and we confirmed the rest of the manuscript is to the same standard.*

9. As we are a methods journal, please revise the Discussion to explicitly cover the following in detail in 3-6 paragraphs with citations:

- a) Critical steps within the protocol
- b) Any modifications and troubleshooting of the technique
- c) Any limitations of the technique
- d) The significance with respect to existing methods
- e) Any future applications of the technique

*The discussion has been revised to explicitly cover the steps mentioned.*

10. Please remove the embedded Table of Materials from the manuscript. The uploaded xls/xlsx file is sufficient.

*The embedded Table of Materials has been removed.*

11. References: Please do not abbreviate journal titles. Please include volume and issue numbers for all references.

*We changed the journal titles to their full length.*

### **Reviewers' comments:**

#### **Reviewer #1:**

Manuscript Summary:

This paper describes recommended equipment and protocol for recording the egocentric view of toddlers and infants by using a head-mounted camera. The protocol should be helpful for other researchers to design and conduct experiments to study the development of child's visual environment set out of a laboratory setting.

Overall, the paper is well written and provides useful information for those who plan to carry out experiments to analyze egocentric views of infants and toddlers with head-mounted cameras. I only have minor comments to be addressed before the paper is published.

1. A part of the manuscript is highlighted (L99-L104, for instance). What does this mean?

*The highlighted portions indicate which parts of the manuscript will be filmed by JoVE's production team.*

2. The reported protocol includes different types of information. Some are guidelines for choosing equipment and designing experiments (e.g., 1.1.1 and 1.1.2), while others are more like observations and know-how from the authors' own experiments (e.g., 1.7 and 2.1). It would be better if the description of the protocol is reorganized such that guidelines and observations can be more easily understand without confusion.

*Thank you for this suggestion. We went ahead and moved any discussion of the method to the Discussion section.*

3. Some items such as 3.1, 3.2.2, and 3.3 seem to be obvious and may be omitted.

*We have omitted the items noted. We further removed the Representative Analysis section from the manuscript and replaced it with a Representative Method.*

4. In L110, "56.7 grams of less": The number is very specific. How did you get this?

*This number was taken from our own internal documents. We decided to include a reference for the recommended weight of a head camera and changed the recommended weight to 30 grams, which is in line with the reference provided.*

## **Reviewer #2:**

This paper provides a set of guidelines for researchers interested in collecting first-person video data with infants in lab and home settings. These techniques are difficult and the authors are pioneers in this field. Overall this paper could make a solid contribution to the field, opening this exciting area up for more research.

Intro -  
No issues.

Guidelines (general) -

The formatting of the guidelines is not intuitive (the list is hierarchical but visually appears "flat") - will this be resolved in the final draft?

*The final formatting of the document will be completed by JoVE. For the purposes of the final submitted drafts, this format is in line with their requirements.*

Why does this section start with a statement about the IRB?

*We have removed the statement about the IRB from the final manuscript.*

Why is under the heading "Protocol" (rather than, e.g. guidelines?) The three sections ( camera selection, data collection, and data analysis) provide guidelines for many studies and are thus not a protocol per se. The statements here should be integrated with the section on "representative results". Additionally, the section on representative results should be broken down into methods and then results.

*Thank you for these recommendations. We considerably rearranged the Protocol. We removed the Representative Analysis section and replaced it with Representative Methods. Statements that were more indicative of a discussion of the Protocol were moved to the Discussion.*

A number of the guidelines simply state a recommendation without enough description of how to actually follow the recommendation. For example, 1.2 (examples of what kinds of issues might come up, or more generally, some more general description of what camera features will be "easy enough" vs too hard) , 1.4 - add that they should be tested? Or are there some technical specifications to look for ?

*We reworded Section 1 to be more specific in recommendations for the head-camera. We did our best to avoid "judgement-call" language that was not explained more thoroughly. We also added a section to remind the experimenters to test the cameras before the experiment begins.*

Also: 1.1.1: Could the authors provide some specific examples /considerations for how cameras can be secured. Describing the way the four cameras in figure 2 were attached (and lessons learned) would be plenty

*We have revised Figure 2 to more clearly show the ways in which head cameras can be secured. The top row of Figure 2 demonstrates actual use of the cameras while the bottom row is a close-up shot of the camera on a mannequin.*

2.4.5. and more specifically - 2.4.5.2: It's not clear what the authors are referring to re: building redundancy into the cap or how this could be achieved.

*We have removed the vague language from this section and replaced it with more direct language.*

Comments on specific sections:

Section 1 (camera selection). I read through the section on head cameras surprised that there was no list of recommended cameras. I think point 1.8 should be brought up at the beginning of this section or there should be a mention of it at the beginning, e.g.: There are numerous small, lightweight and portable cameras readily available for purchase, and we provide a list of recommendations below.

*Per JoVE standards, all commercial language has been removed from the manuscript. Specific product models have been included in the Table of Materials.*

Under point 1, figure 2 doesn't seem super useful as is - cameras are all very small and its hardly possible to make out the differences. Could the images be cropped and expanded so that more details on the cameras and head-mounts are made visible?

*Thank you for this suggestion. We have revised Figure 2 to address your concerns. It now more clearly shows the way the head-cameras have been attached.*

1.8. Can the authors provide a description of the shared features of the listed cameras and what differentiates them from cameras not being recommended to use? E.g. is it just their size? Have these been tested for durability under heavy use? are they easy for parents to use? etc  
Battery requirements are not mentioned. Any considerations here?

*Thank you for suggesting to include battery requirements. We have added a section (1.7) to speak directly to this issue. Unfortunately, a thorough comparison of different camera models is outside the scope of this manuscript.*

Section 2 (data collection)

2.2 refers to three experimenters but 2.3.1.1. refers to a parent. Are there four adults putting the cap on? If no, then 2.2 should be rephrased to "adults"

*We revised Section 2.1-2.2 and their subsections to make this process clearer. The parent's role in this process is mainly to maintain their child's temperament and distract the child. The majority of the work should be done by one or two experimenters. The third experimenter is monitoring the image of the head camera.*

The angle of the camera is brought up multiple times 2.3.3.1 and then 2.4.2 - can the authors make a single heading in which to describe motivations/ goals and tradeoffs about angles (viewing hands, down vs. up, etc)

*As the angle of the camera is more a discussion point of the method, we moved these sections to the Discussion section.*

Section 3 (data analysis)

This section is a bit weak compared to the others. E.g. the authors mention computer vision techniques in the intro but not here. Can there be a few points about this added here, including examples of what has already been done in the literature and what is likely/unlikely to be possible given the nature of first person video?

*We have removed the Data Analysis section from this manuscript and replaced it with a Representative Method section. This new section gives a better idea of the process by which to collect head camera data in a laboratory setting.*

Add 3.1.1 - move the descriptions of downsampling and considerations for downsampling (lines 266-269) here

*We have removed this section from the Revised Draft.*

Re: 3.2.1. - is there software available to do the bounding boxes? Could this be provided to the readers?

*There are multiple bounding box software solutions, some of which can be coded in programming languages such as Matlab. Per JoVE's policies we do not mention commercial products in the Revised Manuscript. We do not provide a software to the readers.*

3.3. "more detailed analyses" should be explicitly linked to the section which follows on representative results - otherwise the point does not make a useful recommendation

*We have removed the Data Analysis section from this manuscript and replaced it with a Representative Method section.*

Representative analyses:

Figure 4a - for the histograms how are they using a t-test? On the means? This statistic isn't clear

*We revised this section to be clearer. Lines 214-217: "An independent t-test comparing the number of scenes with a given number of objects between the child and the parent revealed this child had a greater number of scenes with fewer objects in view compared to their parent"*

Discussion:

The three Sentences about gaze vs. head angle (323-326) seem out of place in the discussion - it's already been made above, and the next sentence (Head cameras allow us to investigate the aggregate of the scenes available to a child) is clear enough.

*Thank you for the suggestion. We removed the sentences in question.*

Minor points

P3 line 82 - the second three here seems to be a typo: (only 13% of frames 3)

*Thank you for catching this typo. It has been fixed.*

P3 line 84: Run on sentence. Also missing a period. Thus, if the question of interest is the visual content available to infants (not the dynamics of momentary attention) head-mounted cameras do a good job capturing the visual environment most of the time, providing a window into their everyday experiences

*Thank you for catching this run-on. We have edited the lines appropriately.*

2.1.2 should have a citation.

*This section is no longer in the Manuscript.*