



Published in final edited form as:

Neuroimage. 2021 January 15; 225: 117484. doi:10.1016/j.neuroimage.2020.117484.

Response to Lynch et al: On measuring head motion and effects of head molds during fMRI

E Jolly^{1,+}, S Sadhukha², L.J. Chang¹

¹Computational Social Affective Neuroscience Laboratory, Department of Psychological and Brain Science, Dartmouth College

²Department of Psychology, New York University

We recently presented evidence indicating limited efficacy of custom-molded headcases in reducing head motion in two naturalistic experimental contexts - passive movie watching, and speaking in the scanner (Jolly et al., 2020). In a commentary on this work, Lynch et al (2020) present additional data that support the original findings of (Power et al., 2019) and raise several potential issues with our recent work. We appreciate the opportunity to address these criticisms and raise additional points that should be considered when interpreting these conflicting findings. We do not believe that their criticisms diminish the value of our work, but instead, along with this reply, help better elucidate the key factors researchers should consider to make the most informed choice about their own research protocols.

We begin by addressing the findings from their newly reported data from 2 participants (the authors C.J.L and J.D.P). These data suggest that for longer (14.4m) resting state functional magnetic resonance imaging (rsfMRI) scans of adult participants (31 and 39 years old) custom molded headcases (headcases) significantly reduce mean framewise displacement (FD) and the proportion of values with $FD > 0.2\text{mm}$. While we have no reason to dispute the validity of these findings there are several important factors that should be considered when interpreting their results.

1. Unrepresentative participants and potential demand characteristics

Given that both participants are experienced neuroimaging researchers who have self-disclosed being repeatedly scanned with and without headcases since 2018, it is fair to assume that they are unlikely to be representative of the majority of individuals that participate in neuroimaging research. Moreover these researchers have extensively published on head motion including the sole paper demonstrating the efficacy of headcases in reducing head motion (Power et al., 2019) and it seems reasonable to assume that they have stronger prior beliefs than the majority of neuroimaging researchers that headcases could reduce head motion. It is well known that expectations can impact behavior across a variety of contexts such as experimental behavior (R. Rosenthal, 1966), clinical outcomes (Ashar et al., 2017), and even hormonal modulations (Crum et al., 2011) and neurotransmitter release (de la Fuente-Fernández et al., 2001).

*Corresponding Author (eshin.jolly@dartmouth.edu), 6207 Moore Hall, Dartmouth College, Hanover, NH, 03755.

Expectancy effects can also extend to interpersonal contexts (Robert Rosenthal & Rubin, 1978), in which experimenters' (Doyen et al., 2012; R. Rosenthal, 1966), teachers' (Robert Rosenthal & Jacobson, 1968), and clinicians' (Chen et al., 2019; Luborsky et al., 1999) expectations can impact the behavior of others. Our argument is not that the authors intentionally engaged in any malfeasance, rather that it is well known that expectations can non-consciously impact behavior (e.g. subtly lying more still during headcase sessions). We believe that this is an important consideration when evaluating Lynch et al.'s new evidence supporting the efficacy of headcases, because both individuals participated in both experimental conditions (with and without headcases) and were unblinded to the experimental hypothesis.

2. rsfMRI paradigms often contain more head motion than task paradigms providing more opportunity for headcases to reduce motion.

Task paradigms and naturalistic viewing paradigms, in particular, have been noted to produce less head motion than rsfMRI paradigms because of participants' level of engagement (Huijbers et al., 2017; Vanderwal et al., 2015). Therefore these new data, in addition to data from Power et al (2019), are primarily representative of how effective headcases can be *in the absence of a task*. This limitation is noteworthy, as the findings from our work on the other hand, speak specifically to naturalistic paradigms that engage participant attention, such as movie-watching and speaking. We believe this point is crucial, because researchers who are choosing their own data collection strategy, should weigh evidence by how appropriate and representative it is of their particular research question.

Overall, we believe the data from Lynch et al (2020) are valuable because they do address speculations we made in our own work about the limitations of a primarily developmental sample in Power et al (2019). They also partially address speculations we made about scan length as a potential factor that can influence the efficacy of headcases. While it is encouraging to see that that headcases may benefit longer rsfMRI scans, we also demonstrate in our Supplementary Materials (Fig S6) that mean differences appear to favor headcases for scans up to ~10m, but that these differences disappear and trend in the opposite direction (i.e. headcase motion is worse) as the scan continues (minutes ~10-45). This raises the possibility that beyond a certain run length, headcases may be no more beneficial than other standard approaches to fMRI data collection.

We now turn to address the four main criticisms raised with our work to provide additional factors that should be considered along with these critiques:

1. Use of across-subject paradigms

In principle, we agree that within-subject paradigms are more effective at controlling nuisance factors that may influence between-subject designs assuming appropriate counter-balancing of order effects. However, collecting data of this type was not straightforward in our circumstances for 2 reasons:

- a. We originally collected Dataset 1 earlier than Dataset 2 and had no original plans to perform a set of comparisons examining the efficacy of headcases. It was only after collection of Dataset 2 and reading Power et al (2019) that we noticed we were in a unique position to provide data and analyses examining how headcases affected head motion for *naturalistic* paradigms. Enough time had passed between data collection of each dataset that recontacting and manufacturing headcases for participants from Dataset 1 was not feasible.
- b. Even if we were able to perform a within-subjects comparison using Dataset 1, we believe that re-watching the same TV show is a *fundamentally different experience* that has the potential to influence head motion. For example, anticipating particular narrative events would have likely caused participants to react differently. Recent work supports this claim demonstrating “neural anticipation” effects in which multivariate activity patterns shift backwards in time upon repeated viewing of the same movie (Lee et al., n.d.). These anticipation effects are unlikely to occur or even be measurable in rsfMRI paradigms. At the same time, while our choice of movie induced a large range of emotions (Chang et al., 2018), a stimulus whose content can potentially induce more motion on first watch (e.g. frightening participants with scary videos) would not be expected to induce the same level motion on a second viewing irrespective of headcase use. This is an important point because unlike rsfMRI paradigms where researchers would not expect stimulus-related anticipatory effects, repeated viewing of naturalistic paradigms *can* be experienced differently by individuals and therefore impact within-subject head motion differences (Lee et al., n.d.).

Our results provide a useful datapoint for the community to recognize that headcases are not *universally* beneficial. We argue that comparisons between *different* fMRI samples may actually be *more* representative of typical data collection paradigms that do not involve repeatedly scanning the same individuals. It is unlikely that a particular lab would have collected data using a specific sample of participants and then re-scanned *the same participants* with headcases for a new paradigm. Statistically, the mean reduction in head motion observed in Lynch et al, and Power et al, are primarily representative of the expected reduction in motion for a *particular individual* who was scanned multiple times, not *distinct samples* of participants who did or did not wear headcases. Our data can speak to that latter scenario.

2. Comparing framewise displacement derived from sequences with different sampling rates

We appreciate this feedback and it did not occur to list this as a caveat in our paper and agree that it makes the most sense to compare datasets with the *same* temporal sampling resolution. However, Lynch et al, neglect to note that our

original paper *did* include comparisons between datasets that *were matched* on every acquisition parameter (Dataset 1 and Dataset 2). Comparisons between these datasets, (*Viewing* only), were not significantly different across FD_{Mean} , FD_{Median} , or $\text{Spike}_{\text{Proportion}}$ even when we removed high motion volumes ($FD > 0.3\text{mm}$) (Figs 1 & 2, left column blue and cyan bars). Thus in a comparison between groups of participants matched on all acquisition parameters, and in which participants were *not* speaking aloud (criticism IV in Lynch et al), we still failed to detect a significant reduction in head motion for participants who wore headcases.

We also do not believe that differences in TR length between Datasets 2 and 3 can fully account for our results. In these comparisons we actually find significant motion in the direction *opposite* from what would be expected based upon how TR length impacts FD magnitude. Lynch et al argue that displacement over a fixed time period (1mm/s) would manifest as a *smaller* FD magnitude (1.5mm) at a shorter TR (1.5s) while this same motion would manifest as a *larger* FD magnitude (2mm) at a longer TR (2s). Yet, despite having a longer TR in Dataset 2 (2s) than Dataset 3 (1.5s), FD_{Median} and FD_{mean} (after excluding high motion TRs) were *higher* in Dataset 3. This would not be expected if TR were the primary driver of the results that we report.

3. Use of task paradigms that dynamically reshape the head within rigid molds

While we appreciate Lynch et al's argument, we humbly disagree with their belief that it is "not surprising that speaking with a rigid head mold negates the mold effect or even produces more motion than speaking within a rigid head mold." In our experience, several research groups (including us) have looked into purchasing headcases *specifically for* paradigms that may have a higher likelihood of increasing head motion (e.g. speaking). Neither the information presented in Power et al (2019) nor the Caseforge website itself, ever suggested to us or other other research groups that we have been in communication with, that such tasks are a "misapplication" of headcases. In fact, we are aware of more than one research group that has chosen to proceed with using headcases in fMRI experiments with a speaking component, *despite* the findings in our paper. Therefore, we find it a bit surprising that Lynch et al are so confident in their assertion.

Regardless, it is precisely for this reason that we believe our work makes a valuable contribution to the literature: if a lab was previously considering purchasing headcases for use in a motion-inducing paradigm, to our knowledge, our manuscript is the *only* contribution to the literature that suggests (with data) that they may not experience positive benefits. Further, our work offers several possibilities as to why this would be the case (e.g. restriction in all planes of motion that focuses motion onto the z-axis while speaking). We believe the existence of such evidence can ultimately help researchers make more informed choices about their own data collection procedures.

4. Mold fit matters and appears to have been problematic.

We do not contest that a subset of our participants experienced a problematic fit with their headcase, but note critically that *these participants were excluded from all analyses*. Thus, it is unlikely that lack of comfort played a role in our findings. We also appreciate Lynch et al, elucidating the amount of time spent on ensuring good fits of their headcases. We believe this is critical information that the neuroimaging community should be made aware of and note that is absent on both Caseforge's website and in Power et al, 2019. Labs should be made aware that it may take up to 6 months to perfect headcase fits *prior* to data collection. It is our belief that for many research labs and research questions, this endeavor may be impractical or infeasible. Therefore, a key goal of our work (and noted in our conclusion) was to help researchers *make the most informed choice* regarding their own data collection procedures. Demonstrating that headcases are anything but a straightforward and simple fix for head motion is extremely valuable to the community because it can help researchers avoid frivolously investing money without *also* investing a significant amount of time on iteratively improving their optical scans and manufactured molds. Our work helps illustrate a realistic time-effort-funding tradeoff scenario that researchers may very well find themselves in.

In summary, we greatly appreciate the efforts of the Caseforge team to develop innovative solutions to improve data quality as well as the opportunity to have a public platform to discuss our work. We highly value the Power team's clear dedication to this topic from their prior work and willingness to contribute a commentary on our paper. We have provided a summary of these discussed findings in Table 1. The pursuit of improving research methods rarely receives the same recognition as contributing novel scientific discoveries, yet provides the critical infrastructure to facilitate all scientific endeavors. By piecing together when, where, and why innovative methods such as Caseforge's custom head molds improve data quality provides an invaluable resource to researchers around the world and we hope that all of these efforts will help Caseforge refine their product to meet neuroimaging researcher's needs.

Acknowledgements

This work was supported by an award from the National Institute of Mental Health R01MH116026. The authors declare no competing financial interests.

References

- Ashar YK, Chang LJ, & Wager TD (2017). Brain Mechanisms of the Placebo Effect: An Affective Appraisal Account. *Annual Review of Clinical Psychology*, 13, 73–98.
- Chang LJ, Jolly E, Cheong JH, & Rapuano K (2018). Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *bioRxiv*. <https://www.biorxiv.org/content/10.1101/487892v1.abstract>
- Chen P-HA, Cheong JH, Jolly E, Elhence H, Wager TD, & Chang LJ (2019). Socially transmitted placebo effects. *Nature Human Behaviour*, 3(12), 1295–1305.
- Crum AJ, Corbin WR, Brownell KD, & Salovey P (2011). Mind over milkshakes: mindsets, not just nutrients, determine ghrelin response. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, 30(4), 424–429; discussion 430–431.

- de la Fuente-Fernández R, Ruth TJ, Sossi V, Schulzer M, Calne DB, & Stoessl AJ (2001). Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science*, 293(5532), 1164–1166. [PubMed: 11498597]
- Doyen S, Klein O, Pichon C-L, & Cleeremans A (2012). Behavioral priming: it's all in the mind, but whose mind? *PloS One*, 7(1), e29081. [PubMed: 22279526]
- Huijbers W, Van Dijk KRA, Boenniger MM, Stirnberg R, & Breteler MMB (2017). Less head motion during MRI under task than resting-state conditions. *NeuroImage*, 147, 111–120. [PubMed: 27919751]
- Jolly E, Sadhukha S, & Chang LJ (2020). Custom-molded headcases have limited efficacy in reducing head motion during naturalistic fMRI experiments. *NeuroImage*, 222, 117207. [PubMed: 32745683]
- Lee CS, Aly M, & Baldassano C (n.d.). Anticipation of temporally structured events in the brain. 10.1101/2020.10.14.338145
- Luborsky L, Diger L, Seligman DA, Rosenthal R, Krause ED, Johnson S, Halperin G, Bishop M, Berman JS, & Schweizer E (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, 6(1), 95–106.
- Power JD, Silver BM, Silverman MR, Ajodan EL, Bos DJ, & Jones RM (2019). Customized head molds reduce motion during resting state fMRI scans. *NeuroImage*, 189, 141–149. [PubMed: 30639840]
- Rosenthal R (1966). Experimenter effects in behavioral research. *Appleton-Century-Crofts*.
- Rosenthal R, & Jacobson L (1968). Pygmalion in the classroom. *The Urban Review*, 3(1), 16–20.
- Rosenthal R, & Rubin DB (1978). Interpersonal expectancy effects: The first 345 studies. *The Behavioral and Brain Sciences*, 1(3), 377–386.
- Vanderwal T, Kelly C, Eilbott J, Mayes LC, & Castellanos FX (2015). Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *NeuroImage*, 122, 222–232. [PubMed: 26241683]

Table 1.

Summary of referenced headcase findings

Data Source	Task Paradigm	Sample Size	Age M (SD)	Scan Duration	Headcase Improvement
Chang et al, 2018	Naturalistic	37	19 (1.7)	45.47	
Jolly et al, 2020	Movie Viewing	26	22.8 (4.7)	45.47	No
Jolly et al, 2020	Naturalistic	26	22.8 (4.7)	20.9 (6.5)	
Chen et al, 2017	Speaking	17	20.8 (2.1)	22.2 (8.6)	No
Power et al, 2019	Resting state	13	15 (5)	4.8	Yes
Lynch et al, 2020	Resting state	2	37 (2.8)	14.4	Yes
Etzel & Braver, 2020	Resting state	1	Not reported	10.7	Yes

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript