nature portfolio

Peer Review File

Perturbing human V1 degrades the fidelity of visual working memory

Corresponding Author: Professor Clayton Curtis

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
This study examines effects of TMS over human V1 applied during the delay period of a memory-guided saccade task. The authors observe disruption to saccade accuracy for targets located in the retinotopic region of the stimulation, when TMS is applied either early or in the middle of the delay. A reduction in EEG alpha-band lateralization associated with working memory was also observed. The paper concludes that activity in V1 is necessary for working memory maintenance.

This is a straightforward study present in a mostly clear style, and the methodological elements of phosphene mapping and electrical field modelling are a valuable addition. However this method is unable to support the paper's key conclusion, for two main reasons:

1. Although the direct electromagnetic action of the TMS pulse may be confined to V1, the spiking activity it induces will be transmitted via synaptic connections to other areas of the brain. Because of this it is impossible to conclude from the present results that activity in V1 is necessary for working memory: it may be that the activity indirectly induced in other brain areas by V1 stimulation disrupts WM-related activity in those areas.

2. Although the basis for many of the classic primate experiments, it is debated to what extent performance in the memory-guided saccade task depends on working memory for the target location versus allocation of selective visual attention and/or oculomotor preparation. It could be one of these other processes that is disrupted by stimulation here (directly, or indirectly as in 1).

Other points

The presentation of EEG results is ambiguous: is the significant decrease observed in alpha-band power in the simulated hemisphere (as stated in the main text and Fig 2C legend) or in the alpha lateralization index (as implied by Methods).

The authors should acknowledge previous work by Eva Feredoes and colleagues demonstrating disruption of working memory for motion direction by TMS of visual area MT+, e.g.
Zokaei, N., Manohar, S., Husain, M., & Feredoes, E. (2014). Causal Evidence for a Privileged Working Memory State in Early Visual Cortex. The Journal of Neuroscience, 34(1), 158–162. https://doi.org/10.1523/JNEUROSCI.2899-13.2014

Based on the reported method for choosing TMS intensity, it seems likely that phosphenes were induced by stimulation on at least some trials. This should be discussed.

Interpretation of hypothesis tests, particularly those with null results, would be aided by reporting Bayes Factors alongside p-values (a package like JASP can calculate both).

The authors may be interested in this recent paper showing WM-related activity at the single neuron level in primate V1:
Huang, J., Wang, T., Dai, W., Li, Y., Yang, Y., Zhang, Y., Wu, Y., Zhou, T., & Xing, D. (2024). Neuronal representation of visual working memory content in the primate primary visual cortex. Science Advances, 10(24), eadk3953. https://doi.org/10.1126/sciadv.adk3953

Reviewer #2

(Remarks to the Author)
The authors used transcranial magnetic stimulation (TMS) over early visual cortex to examine its necessity for visual working memory. The authors report that TMS to visual cortex during a maintenance (Experiments 1 and 2) as well as an encoding interval (Experiment 2) impairs memory-guided saccade accuracy. EEG recordings in Experiment 1 are used to suggest that changes in alpha band oscillatory activity after visual cortex TMS may be degrading visual working memory.

Overall, this is an interesting and well-written manuscript describing some potentially compelling findings. With the following further clarifications, analyses, and ideally additional data, the manuscript may make an important contribution to the literature:

1) TMS does not have the resolution to stimulate V1 in isolation. As can also be seen in Figure 2, V2 is also extensively stimulated, which is why most TMS studies of this region refer to it as early visual cortex or V1/V2 in the literature. Referring to the region of stimulation solely as V1 throughout this manuscript can be misleading.

2) It remains unclear from these studies whether TMS to early visual cortex or the perception of the visual phosphenes affected memory-guided saccade performance. In other words, could visual stimuli in the shape of the perceived phosphenes and presented at the same location and time as the TMS-induced phosphenes produce similar memory-guided saccade accuracy impairments? An additional experiment without TMS but presenting similar visual stimuli at a corresponding interval that accounts for retina to V1 transmission time would more directly implicate visual cortex involvement in visual working memory and rule out this alternative interpretation of the data.

3) Further details on how memory-guided saccade accuracy was affected would be informative. Were the saccades more hypometric? Were the latencies, durations, velocities, etc. also affected?

4) The decrease in EEG alpha oscillatory activity after the TMS may be independent from visual working performance in this study. For example, the alpha changes could solely result from the loud and felt TMS pulses that may have affected alpha activity, which has been shown to be reduced with the allocation of attention. Given the short-lived effects of the TMS on alpha activity as illustrated in Figure 2C, this transient attentional allocation explanation may be more likely than having an indirect and outlived effect on working memory. As with the MGS data, further details may be informative, such as whether the magnitude of changes in alpha power after TMS were correlated with memory-guided saccade accuracy.

Minor comment:

There are two 'E's and no 'F' in the legend for Figure 1.

Reviewer #3

(Remarks to the Author)
This is a very interesting and well carried out study that addresses an important question in working memory research, namely are early sensory areas such as V1 causally involved in the maintenance of visual working memory? While there is considerable correlational evidence of a role for V1 from human fMRI research, and some supporting evidence from neurophysiological studies in monkeys, there are some controversies and skepticism that persist in the field.

The authors evaluated the effect of triple-pulse TMS, applied during the early encoding period or the later maintenance period, on the precision of spatial working memory in a memory-guided saccade task. The results from two experiments provide quite compelling evidence for the notion that TMS applied to V1 impairs the precision of spatial working memory, with similar effects observed for TMS applied during the early encoding period or during the delay period. These effects are retinotopically specific, further indicating that it is unlikely to be attributable a coarse effect of stimulation (e.g., increased arousal). TMS also modifies the strength of EEG alpha responses in a manner that is consistent with poorer attentional focus in the early visual cortex after the application of TMS.

Overall, the paper is well written and the results appear generally solid. One question of modest concern is the large number of participants that were excluded from the study (12 out of 27 excluded), and it appears that the same participants were run in Experiments 1 and 2. While it is certainly important to maintain good participant data quality, there can be concerns, not specific to the authors but a more general concern in the field, that the exclusion of subjects can sometimes lead to the biasing of statistical outcomes, especially if it is performed in posthoc manner not fully blinded to the experimental results of those participants. One way to address this would be to run this study (perhaps with small changes in experimental paradigm) in a newly recruited group of participants and demonstrate successful replication and generalization of these effects.

Another issue is that previously published study by Rademaker et al (2017) used a similar TMS paradigm, applying stimulation during the maintenance period of a working memory task, but reported an opposite direction of effects. The task was somewhat different (memory for orientation) and there were multiple items, but that study reported that stimulation to V1 actually improved the precision of visual working memory at the corresponding retinotopic location. That study also used a quite similar TMS protocol (triple-pulse TMS stimulation at 10 Hz) so it is not clear why a different pattern of results are

observed across the present study and this previous work. Is the difference in performance attributable to the difference in task (spatial working memory vs. orientation memory) or presence of multiple items? Or might it be that TMS applied to V1 causes a bias effect in terms of the saccade landing location? For example, if one analyzes the effect of TMS on corresponding vs. non-corresponding stimulus trials, is there evidence of a spatial bias in the former situation. One might imagine that TMS leads to either an attractive (or possibly repulsive) bias in spatial judgments, relative to the perceived phosphene location. It seems that this study would benefit from additional experimentation to address such discrepancies; otherwise, the field will consist of independent reports with apparently contradictory findings.

Minor comments

Regarding the widely distributed nature of visual working memory, the study by Ester et al., Neuron, 2015, would be appropriate to cite along with reference 4.

“The most surprising finding in working memory research, which has now been replicated
numerous times, is that the multivariate patterns of neural activity in primary visual cortex, V1,can be used to decode the contents of visual working memory 3,15.”
- Consider adding reference to study by Rademaker et al., 2019, which addresses the types of information that can be successfully maintained in V1 in the presence of distracting information

“Our findings bolster human neuroimaging studies that consistently report that the
visual properties of memoranda stored in working memory can be decoded from the patterns of activity in V1 during retention intervals 4,10.”
- Can cite other relevant references here (e.g., 3, Rademaker et al., 2019)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
In my previous review I criticized the manuscript's central conclusion that V1 is necessary for working memory maintenance, on the basis that TMS to area V1 may indirectly affect activity in other regions of the brain. The authors respond that TMS only has indirect effects on areas monosynaptically connected to the targeted area. Firstly, that answer seems to support my criticism rather than refute it. Secondly, it is obviously false: TMS of area V1 can induce phosphenes, which means that TMS of V1 can cause someone to say "Hey, I saw a flash of light!" But the muscles involved in speech do not have monosynaptic connections to V1. Perhaps the authors will reply that they meant to refer only to subthreshold TMS of V1 (too weak to produce phosphenes) as having indirect effects limited to monosynaptic connections but (a) if this is the case, the studies they cite for this claim certainly do not provide enough evidence for it, and (b) by their own admission they don't know that their stimulation was subthreshold. In any case, it clearly isn't possible for the authors to rule out indirect effects of TMS of V1 on other brain areas, so they can't conclude from their results that representations in V1 are necessary for memory maintenance.

I also registered concern that performance in the memory-guided saccade task might depend on other processes, for instance spatially focused attention, instead of or in addition to visuospatial working memory, and that it could be one of those processes that is disrupted here. The authors suggest the distinction between these processes is semantic, but I strongly disagree: spatial attention, saccade planning and working memory may be confounded in this particular task but they are not confounded in general. The key question is: is there a good reason to believe that their observation, disruption of performance by TMS of V1, will generalize to other tasks that we also consider tests of working memory, e.g. tasks where multiple locations need to be stored, or non-spatial features of items? Or will they instead generalize to tasks that involve maintaining covert attention at a location, or planning saccades? Without answers to those questions it is unfortunately not possible to draw the broad conclusions about V1 involvement in working memory that are claimed in this manuscript.

Reviewer #2

(Remarks to the Author)
The authors have adequately addressed most of my concerns raised in the previous round of reviews. Importantly, the inclusion of a control experiment to rule out an alternative explanation for the increases in spatial working memory errors makes the manuscript much more compelling. One outstanding concern that remains to be adequately addressed, however, is an alternative attentional explanation for the transient changes in alpha activity and working memory errors. Since many studies have shown lateralized alpha power changes after the allocation of spatial attention (e.g., Worden et al., 2000), the reported alpha changes and memory errors may be a consequence of lateralized attentional shifts caused by the lateralized

loud clicks and induced scalp sensations associated with the TMS. The reported alpha power changes (i.e., less lateralization after the TMS), as well as the very transient nature of them, seem entirely consistent with this attentional modulation account. Without a lateralized control TMS site, this alternative explanation cannot be ruled out and should at least be acknowledged.

Reviewer #3

(Remarks to the Author)
The authors have been responsive to the previous reviews, and have largely addressed my concerns. While the revised discussion section is not able to pinpoint the sources of the discrepancies across various TMS studies of visual working memory, it is good that these specific points about the existing literature are discussed and made clear to the reader. Overall, this is a well-conducted study that provides strong positive evidence of a causal role for V1 (or possibly other co-stimulated early visual areas) in working memory maintenance, so I am in favor of recommending publication of this work.

Version 2:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I am impressed by the authors persistence and adaptability in responding to my previous comments, but it is still the case that their method cannot support their strong conclusions about visual cortex housing WM representations.

The authors have added to the discussion a number of caveats to this conclusion, reflecting my own and other reviewers' concerns, although the headline claims in the abstract and elsewhere have not been changed and remain hyperbolic. I would say they have done enough to address my concerns about the use of a saccade task, but not the criticism that V1 stimulation may indirectly affect other brain regions. On this, the authors have acknowledged that the effects of suprathreshold stimulation of V1 extend beyond visual cortex, but defend their conclusion that V1/V2b are necessary to maintain WM representations with the compound argument that (a) their stimulation was subthreshold, (b) subthreshold effects of TMS are limited to monosynaptically connected regions and (c) V1/V2 are not monosynaptically connected to DLPFC, and the "overwhelming majority of connections between V1 to LIP and FEF are not monosynaptic". This latter statement suggests that the authors believe there may in fact be monosynaptic connections to LIP and FEF: if so, this is a flaw even according to their own argument. It is also clear that they did not take sufficient care to ensure their stimulation was subthreshold in the experiment.

However the weakest point is claim (b), which they support with two references: Bestmann et al. (2004) and Paus et al. (1997). Both of these papers imaged cerebral blood flow changes induced by TMS. However, in neither study was visual cortex targeted, and neither paper was designed to test whether indirect TMS effects were limited to monosynaptically connected regions, nor attempted to draw such a conclusion. Highlighting that imaging methods do not have the sensitivity to support such a conclusion, Bestmann et al. did not find evidence of BOLD modulation *in the area targeted by TMS* for subthreshold stimulation, despite finding indirect effects elsewhere in the brain. Without strong evidence that indirect effects of TMS are limited to visual cortex, their conclusion that WM representations must be held there is flawed.

Reviewer #2

(Remarks to the Author)

Version 3:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
I am satisfied that the revised manuscript provides a balanced account of the findings, and congratulate the authors on a fine study.

We thank the reviewers for their constructive comments and feel that their input has made our manuscript better. See below in *blue* our responses to the reviewers' comments and associated actions made in the manuscript.

Reviewer #1 (Remarks to the Author):

This study examines effects of TMS over human V1 applied during the delay period of a memory-guided saccade task. The authors observe disruption to saccade accuracy for targets located in the retinotopic region of the stimulation, when TMS is applied either early or in the middle of the delay. A reduction in EEG alpha-band lateralization associated with working memory was also observed. The paper concludes that activity in V1 is necessary for working memory maintenance.

This is a straightforward study present in a mostly clear style, and the methodological elements of phosphene mapping and electrical field modelling are a valuable addition. However this method is unable to support the paper's key conclusion, for two main reasons:

1. Although the direct electromagnetic action of the TMS pulse may be confined to V1, the spiking activity it induces will be transmitted via synaptic connections to other areas of the brain. Because of this it is impossible to conclude from the present results that activity in V1 is necessary for working memory: it may be that the activity indirectly induced in other brain areas by V1 stimulation disrupts WM-related activity in those areas.

Response: Thanks for the comment, as this is an issue that comes up often. The concept of diaschisis - a change in functioning of an area connected to a damaged part of the brain - has long been recognized as a potential source of concern in neuropsychological studies of brain lesions. As the reviewer rightly points out, it might also apply to TMS studies. However, there exists some really compelling evidence that this seems unlikely a problem in our study. For instance, the most problematic case would be if downstream higher-order areas like the frontal and parietal cortices were somehow affected by TMS applied to V1. First, neither V1 nor V2 have direct monosynaptic projections to DLPFC BA46 in the macaque at least (Anderson et al. 2011). The vast majority of connections between V1 and LIP and FEF are not monosynaptic (Anderson et al. 2011). For instance, monkey V1 is connected to LIP indirectly requiring several synapses (Blatt et al. 1990). Second, PET and fMRI studies of concurrent TMS suggest that distal areas affected are only those that are monosynaptically connected to the brain target of the TMS (Paus et al. 1997; Bestmann et al. 2004). Similarly, electrical microstimulation of macaque FEF without concurrent visual stimulation does not evoke BOLD activity in V1 (Ekstrom et al. 2008). Third, TMS perturbs spiking activity only within a 2mm radius, and synaptic potentials only within a slightly larger extent in the macaque (Romero et al. 2019). Together, we are confident that the effects we observed here are due to highly localized effects of TMS to V1 and perhaps V2d.

Action: We now add text in the discussion similar to the above text that we hope addresses this concern.

2. Although the basis for many of the classic primate experiments, it is debated to what extent performance in the memory-guided saccade task depends on working memory for the target location versus allocation of selective visual attention and/or oculomotor preparation. It could be one of these other processes that is disrupted by stimulation here (directly, or indirectly as in 1).

Response: The reviewer brings up an important question, to what degree are the effects of TMS to V1 due to disturbance in saccade planning or attention. Let us first respond to the oculomotor issue. If TMS to V1 impacted the motor metrics used by the oculomotor system (e.g., superior colliculus) to plan saccades, one would predict a general, not specific, distortion. However, our results are highly specific. We only find deficits when one is maintaining a location inside the phosphene field that is contralateral to the site of TMS stimulation. Moreover, the types of memory-guided saccade errors were not systematic; TMS impacted the gain and angle of saccade errors (again, only in the contralateral visual field). We've studied the oculomotor systems for decades and cannot think of an oculomotor mechanism that can explain these patterns of results. Second, let's consider the role of visual spatial attention. We agree that we could explain our results in terms of spatial attention maintained during the delay period that is focused precisely on the location of the target. We would bet that most people think of that as the definition of spatial working memory (e.g., (Awh and Jonides 2001; Jerde et al. 2012). Therefore, in order to sidestep semantic ambiguities, we hope the reviewer can understand why we think the memory-guided saccade task, which has been a workhorse for the study of working memory for 45 years, can be used to assess working memory functions. I will add that we are incredibly sympathetic to the reviewer's issue and hope that these words like "attention" and "working memory" can be replaced by neural network models that are not burdened by the uncertainty of our words. Indeed, neural network models for spatial attention and spatial working memory are almost identical in structure in many aspects.

Action: We now add text to clarify this issue.

Other points

The presentation of EEG results is ambiguous: is the significant decrease observed in alpha-band power in the simulated hemisphere (as stated in the main text and Fig 2C legend) or in the alpha lateralization index (as implied by Methods).

Response: We apologize for being unclear and thank the reviewer for spotting the issue. We observed an impact of TMS on alpha lateralization (Figure 2B) and we further explain this effect by showing an impact on alpha-band power for electrodes at the site of stimulation (Figure 2C). The basic data is shown in Supp Figure 3.

Action: We now better clarify exactly what we did in the Results, in Figure 2 caption, and in Supplementary Figure 3.

The authors should acknowledge previous work by Eva Feredoes and colleagues demonstrating disruption of working memory for motion direction by TMS of visual area MT+, e.g.

Zokaei, N., Manohar, S., Husain, M., & Feredoes, E. (2014). Causal Evidence for a Privileged Working Memory State in Early Visual Cortex. The Journal of Neuroscience, 34(1), 158–162. https://doi.org/10.1523/JNEUROSCI.2899-13.2014

Action: We now expand our discussion of our results in the context of previous TMS studies to visual cortex, including the suggested reference.

Based on the reported method for choosing TMS intensity, it seems likely that phosphenes were induced by stimulation on at least some trials. This should be discussed.

Response: We appreciate the reviewer's concern. We will note that the phosphene mapping was only feasible when the room was completely dark (i.e., black tape over LED lights on monitors and devices) and the monitor background is black and the contrast is low. When the monitor is brightened and its background set to gray, people rarely notice the phosphenes. If they reported seeing them, we reduced the stimulator output. Nonetheless, we performed a followup experiment suggested by Reviewer 2, in which we simulated visual phosphenes and found that it had no impact on working memory performance (see below for more details). Overall, we are now more confident that phosphenes are not acting as a distractor and mediating the effect of TMS to V1.

Action: We now describe in Methods (Phosphene Mapping) more about our procedures to deal with the potential awareness of phosphenes. We also now include the results from the control task.

Interpretation of hypothesis tests, particularly those with null results, would be aided by reporting Bayes Factors alongside p-values (a package like JASP can calculate both).

Response: We appreciate the suggestion. Since we use Fischerian statistics, instead of Bayesian, we can also provide effect sizes.

Action: We now include effect sizes so that readers can estimate the magnitude of our effects.

The authors may be interested in this recent paper showing WM-related activity at the single neuron level in primate V1:
Huang, J., Wang, T., Dai, W., Li, Y., Yang, Y., Zhang, Y., Wu, Y., Zhou, T., & Xing, D. (2024). Neuronal representation of visual working memory content in the primate primary visual cortex. Science Advances, 10(24), eadk3953. https://doi.org/10.1126/sciadv.adk3953

Action: Thanks, our lab did cover that paper in detail in our lab meeting when it came out. We now include it in the Discussion.

Reviewer #2 (Remarks to the Author):

The authors used transcranial magnetic stimulation (TMS) over early visual cortex to examine its necessity for visual working memory. The authors report that TMS to visual cortex during a maintenance (Experiments 1 and 2) as well as an encoding interval (Experiment 2) impairs memory-guided saccade accuracy. EEG recordings in Experiment 1 are used to suggest that changes in alpha band oscillatory activity after visual cortex TMS may be degrading visual working memory.

Overall, this is an interesting and well-written manuscript describing some potentially compelling findings. With the following further clarifications, analyses, and ideally additional data, the manuscript may make an important contribution to the literature:

Response: We thank the reviewer for the positive feedback on our study.

1) TMS does not have the resolution to stimulate V1 in isolation. As can also be seen in Figure 2, V2 is also extensively stimulated, which is why most TMS studies of this region refer to it as early visual cortex or V1/V2 in the literature. Referring to the region of stimulation solely as V1 throughout this manuscript can be misleading.

Response: Thank you for the comment. We agree that the effects of TMS cannot be limited to just V1. While our electric-field headmodels do suggest that targets we selected over V1 induces by far the most current in V1, we do see smaller effects in V2d.

Action: We have taken three measures to address this issue. First, we adjust our language throughout to be clear that we targeted V1 with TMS, meaning we directed the maximum focus into retinotopically defined V1, not other areas. Second, we adjust our language throughout to acknowledge that early visual cortex or V1/V2 may have been impacted by TMS. Third, we add a section in the Discussion that directly addresses the uncertainty in the V1 specificity.

2) It remains unclear from these studies whether TMS to early visual cortex or the perception of the visual phosphenes affected memory-guided saccade performance. In other words, could visual stimuli in the shape of the perceived phosphenes and presented at the same location and time as the TMS-induced phosphenes produce similar memory-guided saccade accuracy impairments? An additional experiment without TMS but presenting similar visual stimuli at a corresponding interval that accounts for retina to V1 transmission time would more directly implicate visual cortex involvement in visual working memory and rule out this alternative interpretation of the data.

Response: We too have thought about this as a potential confound and were compelled by the reviewer's comment to do the control experiment (note that Review 1 had a similar concern). Based on the description of how subjects described what the phosphenes looked like, we simulated this. The pseudo-phosphenes were grayscale noise patches presented within the area of each subject's phosphene field. The contrast was low and flipped three times at 20Hz at the timing of the TMS pulses (not applied here in the control) during the middle of the delay. If the awareness of the phosphenes drove our effects, one would predict that the pseudo-phosphenes would have a similar effect. It did not. Working memory

measures were completely unaffected, suggesting that the TMS and not phosphenes drove the impact on memory.

Action: We now include this control experiment, discussed in the Methods, Results, and included as a new Supp Figure 4.

3) Further details on how memory-guided saccade accuracy was affected would be informative. Were the saccades more hypometric? Were the latencies, durations, velocities, etc. also affected?

Response: This was a valuable comment and pushed us to dig deeper into other aspects of the saccade metrics that could have been perturbed by TMS. First, we now report that we did not observe any differences in peak saccade velocities or latencies between the conditions. This helps eliminate the possibility of speed-accuracy tradeoff caused by TMS or some slowing of saccades. We could not compute a reliable *main sequence* because the range of eccentricities were too small, limited by the phosphene fields experimentally determined. Second, using regression analysis we now decompose the MGS error into its radial (gain) and tangential (angular) components and find that both were affected by TMS equally. This is important because it suggests that the application of TMS to V1 causes errors that cannot merely be explained by some systematic motor factors. Instead, they point to induced memory errors that are highly specific to the PF.

Action: Performed further analyses of saccade metrics which are now summarized in Results and in Supplementary Figure 2, and described in Methods.

4) The decrease in EEG alpha oscillatory activity after the TMS may be independent from visual working performance in this study. For example, the alpha changes could solely result from the loud and felt TMS pulses that may have affected alpha activity, which has been shown to be reduced with the allocation of attention. Given the short-lived effects of the TMS on alpha activity as illustrated in Figure 2C, this transient attentional allocation explanation may be more likely than having an indirect and outlived effect on working memory. As with the MGS data, further details may be informative, such as whether the magnitude of changes in alpha power after TMS were correlated with memory-guided saccade accuracy.

Response: We thank the reviewer for bringing up this possibility. First, we do not think the lateralization of the alpha response could be caused by the clicking noise or even the physical sensation under the coil. It might evoke a general orienting response, but we cannot think of a mechanism by which that would impact the lateralized changes in alpha power. We could not find any evidence from previous studies of attention or working memory that suggest this might be a possibility. On the contrary, only a generalized, non-lateralized, orienting response has been reported in a recent study (Conde et al. 2019). Second, we fail to find any metrics of the EEG alpha power that predicts trial-wise MGS accuracy. We looked at this a number of ways and none of them seem to be correlated with WM accuracy. Thus, we are left only with the general effect that TMS has on alpha power and the general effect that WM accuracy is worse following TMS. We also scoured the literature looking for previous publications showing that some measure of alpha predicts WM accuracy on a trial-wise basis and have not found one. Third, we also explored group level effects. However, we found that subjects with larger TMS induced errors do not have any differences in lateralized alpha power. We are perplexed by the absence of any effects related to

the Reviewer's inquiry. Frankly, we are worried that the study was not properly designed to be sensitive to the relationship between lateralized alpha and memory accuracy. For instance, we imagine that we would need a lot more trials with EEG to find these effects given they are probably very small.

Action: Given that we only find null effects that are difficult to explain, we are hesitant to include these null effects as they could mislead readers.

Minor comment:

There are two 'E's and no 'F' in the legend for Figure 1.

Action: Replaced second E in Figure 1 to Figure 1F.

Reviewer #3 (Remarks to the Author):

This is a very interesting and well carried out study that addresses an important question in working memory research, namely are early sensory areas such as V1 causally involved in the maintenance of visual working memory? While there is considerable correlational evidence of a role for V1 from human fMRI research, and some supporting evidence from neurophysiological studies in monkeys, there are some controversies and skepticism that persist in the field.

The authors evaluated the effect of triple-pulse TMS, applied during the early encoding period or the later maintenance period, on the precision of spatial working memory in a memory-guided saccade task. The results from two experiments provide quite compelling evidence for the notion that TMS applied to V1 impairs the precision of spatial working memory, with similar effects observed for TMS applied during the early encoding period or during the delay period. These effects are retinotopically specific, further indicating that it is unlikely to be attributable a coarse effect of stimulation (e.g., increased arousal). TMS also modifies the strength of EEG alpha responses in a manner that is consistent with poorer attentional focus in the early visual cortex after the application of TMS. Overall, the paper is well written and the results appear generally solid.

Response: We thank the reviewer for the positive feedback on our study.

One question of modest concern is the large number of participants that were excluded from the study (12 out of 27 excluded), and it appears that the same participants were run in Experiments 1 and 2. While it is certainly important to maintain good participant data quality, there can be concerns, not specific to the authors but a more general concern in the field, that the exclusion of subjects can sometimes lead to the biasing of statistical outcomes, especially if it is performed in posthoc manner not fully blinded to the experimental results of those participants. One way to address this would be to run this study (perhaps with small changes in experimental paradigm) in a newly recruited group of participants and demonstrate successful replication and generalization of these effects.

Response: We appreciate the reviewer's comments. We will try to respond in several ways that we believe will mitigate these concerns.

First, in order to be sensitive to memory error, one has to have a measure that is sensitive and not biased by non-memory factors. We use memory-guided saccades for this very reason.
- We can measure errors up to the limits of the eye tracker (~0.0017 radians or 1/10th of a degree).
- We would be unaware of response mapping errors if using a button box, for example, where one has to learn and track the relationship between which button to press for which type of memory dependent response. In our experience, many errors are response mapping errors and experimenters are blind to them. Saccades made to a location in memory are natural and require a well practiced mapping.
- 2AFC procedures make a host of assumptions from which a simple saccade is free.
- The number of trials needed for 2AFC is much more than for MGS.

Those are some of the reasons we choose MGS. However, they come at a cost. If we cannot make high quality measures of gaze, we cannot use the subject as there is nothing reliable to measure. Often we do

not know whether a subject will be good until they are already well into the experiment. We can also report that the exclusion is done without respect to the memory performance of the subject. In fact, the gaze data is typically so poor that there is no way to estimate memory errors. Moreover, often subjects were excluded based on their poor data during the No TMS session. Thus, inclusion/exclusion was not impacted by the main experimental factor (i.e., noTMS vs TMS; or TMSinPF vs TMSoutPF).

Finally, we chose to focus on the reproducibility of our findings in the same population, and the generalizability of the findings across two different delay times when TMS was applied (i.e., early late). We feel that these attempts are at least equal to if not better than most TMS studies published. Nonetheless, we do agree with the reviewer that replicating these results in a whole new sample of participants would be useful. Indeed, we are doing that now in a new sample using slightly different task demands and applying TMS to a variety of areas in early visual cortex, as well as frontal and parietal cortex. These ongoing studies are beyond the scope of the current study in our opinion.

Another issue is that previously published study by Rademaker et al (2017) used a similar TMS paradigm, applying stimulation during the maintenance period of a working memory task, but reported an opposite direction of effects. The task was somewhat different (memory for orientation) and there were multiple items, but that study reported that stimulation to V1 actually improved the precision of visual working memory at the corresponding retinotopic location. That study also used a quite similar TMS protocol (triple-pulse TMS stimulation at 10 Hz) so it is not clear why a different pattern of results are observed across the present study and this previous work. Is the difference in performance attributable to the difference in task (spatial working memory vs. orientation memory) or presence of multiple items? Or might it be that TMS applied to V1 causes a bias effect in terms of the saccade landing location? For example, if one analyzes the effect of TMS on corresponding vs. non-corresponding stimulus trials, is there evidence of a spatial bias in the former situation? One might imagine that TMS leads to either an attractive (or possibly repulsive) bias in spatial judgments, relative to the perceived phosphene location. It seems that this study would benefit from additional experimentation to address such discrepancies; otherwise, the field will consist of independent reports with apparently contradictory findings.

Response: These comments raise important questions about why we found positive results, while other TMS studies have been all over the place. The few TMS studies of the role of visual cortex in WM have not been able to confirm or refute whether it is necessary for WM. For instance, TMS to occipital cortex only slows WM based decisions when applied after the delay period (Cattaneo et al., 2009). In one study, TMS worsened WM precision very early in the delay (100ms), but not later (400ms) or even during the presentation of the visual array (van Lamsweerde and Johnson, 2017). In another study, TMS had an effect in only 1 out of 6 conditions (high WM load at 400ms into the delay) (van de Ven et al., 2012). Finally, in another study 10Hz TMS improved WM precision (Rademaker et al., 2017). What is causing these inconsistencies? We speculate that there are a number of factors. First, these studies were designed to mainly test the role of visual cortex in encoding or early consolidation. Second, methodological issues might have hampered clear interpretations. Specifically, the retinal specificity of the effects was not well controlled, the timing of pulses was too early in the delay, the application of weak single pulses did not have robust effects, sensitive behavioral indices of WM performance were lacking, and measures of neural activity affected by TMS were not included. In our study, we actually think we benefited from the successes and mistakes in the previous studies.

Action: We now expand upon the section in the Discussion to put our results in a better situation with respect to these previous studies. Moreover, we have two followup TMS studies in progress that we think might shed more light on why there have been inconsistent results, but these are well beyond the scope of the current study that we feel stands on its own.

Minor comments

Regarding the widely distributed nature of visual working memory, the study by Ester et al., Neuron, 2015, would be appropriate to cite along with reference 4.

Action: Thanks, done.

"The most surprising finding in working memory research, which has now been replicated numerous times, is that the multivariate patterns of neural activity in primary visual cortex, V1,can be used to decode the contents of visual working memory 3,15."
- Consider adding reference to study by Rademaker et al., 2019, which addresses the types of information that can be successfully maintained in V1 in the presence of distracting information

Action: Done.

"Our findings bolster human neuroimaging studies that consistently report that the visual properties of memoranda stored in working memory can be decoded from the patterns of activity in V1 during retention intervals 4,10."
- Can cite other relevant references here (e.g., 3, Rademaker et al., 2019)

Action: Done.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

In my previous review I criticized the manuscript's central conclusion that V1 is necessary for working memory maintenance, on the basis that TMS to area V1 may indirectly affect activity in other regions of the brain. The authors respond that TMS only has indirect effects on areas monosynaptically connected to the targeted area. Firstly, that answer seems to support my criticism rather than refute it. Secondly, it is obviously false: TMS of area V1 can induce phosphenes, which means that TMS of V1 can cause someone to say "Hey, I saw a flash of light!" But the muscles involved in speech do not have monosynaptic connections to V1. Perhaps the authors will reply that they meant to refer only to subthreshold TMS of V1 (too weak to produce phosphenes) as having indirect effects limited to monosynaptic connections but (a) if this is the case, the studies they cite for this claim certainly do not provide enough evidence for it, and (b) by their own admission they don't know that their stimulation was subthreshold. In any case, it clearly isn't possible for the authors to rule out indirect effects of TMS of V1 on other brain areas, so they can't conclude from their results that representations in V1 are necessary for memory maintenance.

Response: We thank the reviewer for clarifying to us the significance of considering indirect effects of TMS. After reading more carefully in this area, we agree with the reviewer that we need to address this possibility. We now add a paragraph devoted to acknowledging that our results could be due to indirect effects of TMS on monosynaptically connected brain areas. We discuss and cite papers relevant to this issue and discuss the implications of indirect effects, if they were indeed the main driver of the increased memory errors. In brief, we acknowledge the possibility.

With regard to the issue of subthreshold stimulation, as now clearly detailed in our methods, participants did not report seeing phosphenes because of our carefully controlled experimental setup (eg, brightness of monitor prevents one from seeing phosphenes). And if they did, it was only on the first few trials, after which we lowered the TMS intensity and they no longer reported seeing phosphenes. Moreover, we conducted a psychophysical experiment where we simulated phosphenes, which had no impact on working memory behavior. These simulated phosphenes could have caused them to say “Hey, I saw a flash of light!” but none of our subjects did. Thus, we believe we do know that our stimulation was subthreshold (unless our participants were lying). Plus, with our psychophysical experiment we know that the indirect “cognitive” effects or awareness of visual distraction cannot explain our results.

Action: See first half of large paragraph in discussion on “caveats.”

I also registered concern that performance in the memory-guided saccade task might depend on other processes, for instance spatially focused attention, instead of or in addition to visuospatial working memory, and that it could be one of those processes that is disrupted here. The authors suggest the distinction between these processes is semantic, but I strongly

disagree: spatial attention, saccade planning and working memory may be confounded in this particular task but they are not confounded in general. The key question is: is there a good reason to believe that their observation, disruption of performance by TMS of V1, will generalize to other tasks that we also consider tests of working memory, e.g. tasks where multiple locations need to be stored, or non-spatial features of items? Or will they instead generalize to tasks that involve maintaining covert attention at a location, or planning saccades? Without answers to those questions it is unfortunately not possible to draw the broad conclusions about V1 involvement in working memory that are claimed in this manuscript.

Response: We respect the reviewer's opinion on the potential other cognitive/motor processes that might have been impacted other than working memory. We now add to a paragraph to the discussion that considers how covert spatial attention might mediate the effects we observed. We also cite and discuss a recent TMS study that used very similar methods that we use here from one of my colleagues, Marisa Carrasco. They find that endogenous attention deployed covertly to a spatial target is not impacted by TMS to V1. Frankly, we found these results surprising as we had assumed that one would find the same results. However, as the reviewer predicted, these can be disassociated. Thus, we tip our hats to the reviewer and think the inclusion of these new findings really help with the cognitive specificity of the effect.

Moreover, the metrics of the saccade errors are not systematic. If TMS to V1 somehow had an impact on the oculomotor system's planning of the saccade, one would predict that the effect would be systematic (for instance if one considers a perturbation to the map in the superior colliculus). However, we find that the TMS induced memory errors are non-systematic - both the gain and angular components of the saccades were impacted by TMS. Plus, saccade reaction times and peak velocities were unaffected by TMS. Therefore, we think this is pretty good evidence that the effects were not motor in nature, but are specific to memory.

Action: We now include the results from a new data analysis of saccade errors in the results. We also include a discussion of this issue in the second half of the "caveats" paragraph.

Reviewer #2 (Remarks to the Author):

The authors have adequately addressed most of my concerns raised in the previous round of reviews. Importantly, the inclusion of a control experiment to rule out an alternative explanation for the increases in spatial working memory errors makes the manuscript much more compelling. One outstanding concern that remains to be adequately addressed, however, is an alternative attentional explanation for the transient changes in alpha activity and working memory errors. Since many studies have shown lateralized alpha power changes after the allocation of spatial attention (e.g., Worden et al., 2000), the reported alpha changes and memory errors may be a consequence of lateralized attentional shifts caused by the lateralized loud clicks and induced scalp sensations associated with the TMS. The reported alpha power changes (i.e., less lateralization after the TMS), as well as the very transient nature of them, seem entirely consistent with this attentional modulation account. Without a lateralized control TMS site, this alternative explanation cannot be ruled out and should at least be acknowledged.

Response: First we thank the reviewer for their feedback throughout the reviewer process and appreciate the impact their comments have made on the quality of our manuscript. In response to their lingering comments, we now acknowledge and discuss what role spatial attention might have in our results. See the response to Reviewer 1 above. Moreover, we would like to add that spatial working memory, not only spatial attention, invokes a similar lateralized EEG alpha band response (Foster et al. 2016). Although we believe that the fact that both attention and working memory share some common mechanisms, as evidenced by the similarity in alpha topography, but also fMRI based decoding interchangeability (Jerde et al. 2012). Finally, we agree with the reviewer that the click and perhaps somatosensory effects could impact EEG signals. However, because the coil placement is almost literally at the center of the back of the head, and the click sound is almost straight back, we do not think these central, non-lateralized signals would invoke enough lateralized EEG effects to be detectable. For instance, the closer two targets are to the vertical meridian in working memory studies the more difficult it is to detect lateralized effects in EEG and fMRI.

Action: See the second to last paragraph in the Discussion, in which we discuss the attention issue.

Reviewer #3 (Remarks to the Author):

The authors have been responsive to the previous reviews, and have largely addressed my concerns. While the revised discussion section is not able to pinpoint the sources of the discrepancies across various TMS studies of visual working memory, it is good that these specific points about the existing literature are discussed and made clear to the reader. Overall, this is a well-conducted study that provides strong positive evidence of a causal role for V1 (or possibly other co-stimulated early visual areas) in working memory maintenance, so I am in favor of recommending publication of this work.

Response: First we thank the reviewer for their feedback throughout the reviewer process and appreciate the impact their comments have made on the quality of our manuscript.

Reviewer #1 was still not satisfied with our last revision. Below we respond point-by-poi nt (in blue font) to each of the Reviewer's remaining issues (verbatim in black font). We feel that the only remaining issue was that the Reviewer wanted us to temper our interpretations such that they were not "hyperbolic." We have done so throughout the manuscript. Our changes are marked in the manuscript.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

I am impressed by the authors persistence and adaptability in responding to my previous comments, but it is still the case that their method cannot support their strong conclusions about visual cortex housing WM representations.

We thank the reviewer for appreciating our enthusiasm for our research. Below, we detail how we have tempered, once again, our interpretations of the results. We are encouraged that the reviewer has no issues with the quality of our experimental design, our analyses, and the vast majority of our conclusions. The only remaining issue is the degree to which we are allowed to interpret our results in the most straightforward way possible, while also acknowledging potential caveats. We strongly hope the Reviewer will respect our ability to interpret our results, at least in the discussion, where that is usually permissible, especially since we highlight the alternative interpretations advocated by the Reviewer.

The authors have added to the discussion a number of caveats to this conclusion, reflecting my own and other reviewers' concerns, although the headline claims in the abstract and elsewhere have not been changed and remain hyperbolic. I would say they have done enough to address my concerns about the use of a saccade task, but not the criticism that V1 stimulation may indirectly affect other brain regions. On this, the authors have acknowledged that the effects of suprathreshold stimulation of V1 extend beyond visual cortex, but defend their conclusion that V1/V2b are necessary to maintain WM representations with the compound argument that (a) their stimulation was subthreshold, (b) subthreshold effects of TMS are limited to monosynaptically connected regions and (c) V1/V2 are not monosynaptically connected to DLPFC, and the "overwhelming majority of connections between V1 to LIP and FEF are not monosynaptic". This latter statement suggests that the authors believe there may in fact be monosynaptic connections to LIP and FEF: if so, this is a flaw even according to their own argument.

First, we have modified the only conclusion in the abstract. Note that the conclusion now only states that the results are "consistent" with distributed models of working memory and removes reference to "necessary" and "critical."

Before: *"These results change the question from whether early visual cortex is necessary for working memory to what mechanisms it uses to support memory. Moreover, they point to models in which the mechanisms supporting working memory are distributed a*cross brain regions, including sensory areas that here we show are critical for memory storage."

After: *"These results are consistent with models in which the mechanisms supporting working memory are distributed across brain regions, including sensory areas."*

It is also clear that they did not take sufficient care to ensure their stimulation was subthreshold in the experiment.

We disagree with this characterization as we used careful methods to ensure that TMS was subthreshold and did not invoke phosphenes. However, no method is perfect and a couple of subjects reported seeing infrequent phosphenes only at the beginning of the first block of the experiment. This was extremely rare and can in no way account for our results that were consistent across blocks, sessions, and subjects. Moreover, when this rare event happened, we immediately reduced the TMS output such that no phosphenes were observed. We take pride in the rigor of our work and these built in procedures demonstrate that we did indeed take sufficient care to ensure subthreshold stimulation. Additionally, we also took care to make sure that we were delivering sufficient strength of TMS to V1 such that we would have a reliable impact on memory behavior. We could have dialed the output way back, but would probably have found no effect on WM. Indeed, many of the null effects reported with TMS are likely because of low stimulator output. Finally, recall we performed a followup control psychophysical task that demonstrated that visible pseudo-phosphenes cannot account for our behavioral results.

However the weakest point is claim (b), which they support with two references: Bestmann et al. (2004) and Paus et al. (1997). Both of these papers imaged cerebral blood flow changes induced by TMS. However, in neither study was visual cortex targeted, and neither paper was designed to test whether indirect TMS effects were limited to monosynaptically connected regions, nor attempted to draw such a conclusion. Highlighting that imaging methods do not have the sensitivity to support such a conclusion, Bestmann et al. did not find evidence of BOLD modulation *in the area targeted by TMS* for subthreshold stimulation, despite finding indirect effects elsewhere in the brain. Without strong evidence that indirect effects of TMS are limited to visual cortex, their conclusion that WM representations must be held there is flawed.

We apologize for the confusion associated with this sentence. We agree with the reviewer and we removed the sentence about neuroimaging and indirect effects of TMS, as well as the two citations. In that paragraph, we now acknowledge that our effects could be due to indirect effects of TMS to V1. We do this despite the fact that there is no credible evidence that indirect effects of TMS have demonstrable effects on behavior. Nonetheless, the Reviewer is right that it could no matter how unlikely be possible.

With regard to our title, *"Perturbing human V1 degrades the fidelity of visual working memory"*, in case the reviewer has a problem with it, note that it is not an interpretation but a simple statement about what we did and what we measured.

Similarly, in the first sentence in Discussion, we write:
"In summary, TMS targeting human V1 caused a deficit in working memory that was spatially restricted to the portion of the visual field perturbed by the electrical field induced by the TMS pulses."
Here, note that again we simply state that when we applied TMS to V1 it impacted working memory in a spatially specific manner. This is a factual statement and does not involve the interpretation of the effect, and thus should not be an issue.

In the last sentence of the first paragraph of the Discussion, we removed the words "direct causal" and "necessary", and changed to "plays an important role" to again temper our conclusions.

"Bolstered by these experimental procedures, our findings provide ~~direct causal~~ evidence that early visual cortex ~~is~~ plays an important role ~~necessary for~~ in working memory storage, beyond simply the visual encoding or consolidation of stimuli into memory."

With regard to the last sentence in the Discussion, we have once again tempered our interpretation. Note how we omit reference to "necessary" in the text. We replace it with not only another reminder of the "indirect effect" caveat, but with diluted text like, "play an important role" and "contribute". We also situate this one piece of evidence in the context of dozens of neuroimaging and electrophysiology reports consistent with our result. Namely, almost 20 years of neuroimaging research has found that patterns of activity in V1 can be used to decode the contents of memory. And neuronal electrophysiology recorded directly from monkeys has demonstrated that V1 activity represents the contents of memory. So, it should not be terribly surprising that TMS to V1 impacts working memory. The most surprising thing is that it has taken so long for someone to provide clear TMS evidence, as we have.

Before: *“Indeed, we demonstrate that the neural mechanisms housed in V1 (and perhaps V2d) are necessary to maintain accurate working memory representations. Therefore, the question is no longer if visual cortex is necessary for working memory, but now it is by what mechanisms does it contribute to working memory?”*

After: *“Assuming indirect effects do not account for our results, they are consistent with models in which the neural mechanisms housed in V1 (and perhaps V2d) play an important role in maintaining accurate working memory representations. Given the accumulation of evidence from neuroimaging, electrophysiology, and TMS, we argue that the question should no longer be about if visual cortex plays an important role in working memory, but instead should be about by what mechanisms does it contribute to working memory?”*