Cerebral laterality as assessed by functional transcranial Doppler ultrasound in left-and right-handers: A comparison between handwriting and writing using a smartphone.

Notes on the revised manuscript

We would like to express our gratitude to the reviewers for their insightful comments. The comments have been carefully considered by the authors, and we have taken great care to address them. The manuscript has been updated accordingly, and we have prepared a point-by-point response (see below) where we describe how we implemented the requested changes or provide further clarification on our decisions.

<u>Reviewer 1</u>

Reviewer's comment 1:

I can see that this proposal has already been through several rounds of review and meaningful adjustment. My only suggestion would be to add a hypothesis and study goal regarding the correlation/relationship between the behavioral handwriting skill/scores and the ultrasonography signal using your equipment. There is a lot of evidence that behavioral language skills do correlate with functional brain imaging as shown in these articles

https://pubmed.ncbi.nlm.nih.gov/29923097/

https://www.frontiersin.org/articles/10.3389/fnhum.2012.00119/full#:~:text=Notably%2C%20some%20high%20correlations%20between%20fMRI%20brain%20activations,uncontrolled%20underlying%20factors%2C%20such%20as%20age%20%28Lazic%2C%202010%29.

There is already a plan in place to study handwriting skills and obtain scores which could be used in this brain-behavior correlational analysis.

Authors' response:

We would like to thank the Reviewer for their positive comments on our State 1 Registered Report. We concur that brain-behavioural correlations are important for neuroscientific research. However, we respectfully note that we do not plan to specifically collect data on handwriting skill. In fact, we will instruct the participants that they will be free to write in the way they are most accustomed to (e.g., in capital or lower-case letters, stressing the words or not) and that we will not be taking correct spelling of words into account. The only behavioural measure we intend to collect is number of words generated, which does not necessarily reflect handwriting skill per se, but rather skill on a phonemic fluency task (word generation in response to a cue letter). This is a language skill measure and, in fact, the Reviewer does mention "behavioural language skill" in their suggestion along with "handwitting skill", so we are unsure if they would like us to use handwriting or language skill scores and correlate them with laterality indices (i.e., indices stemming from the ultrasonography signal).

In any case, extensive discussions took place amongst our team of authors concerning the possibility of adding a hypothesis and study goal regarding a possible correlation between number of words generated and laterality indices. The outcome of these discussion was a collective decision not to preregister such a hypothesis. The primary reason for this decision is our desire to keep this Stage 1 Registered Report clean and focused on our main hypotheses (which is to assess laterality for written language using a mobile device and explore possible handedness differences). Thus, given that the suggested hypothesis deviates from the our study's core scope, we have decided to explore and report the possible correlation between number of words generated and laterality indices, following the suggestion of the Reviewer (and as Gutierrez-Sigut et al., 2015 and Woodhead et al., 2018 have previously done), but treat this as an additional, unplanned, exploratory analysis. In essence, we are inclined to refrain from adding such a (preregistered) hypothesis at Stage 1, but will certainly report this analysis in our Stage 2 Registered Report. Thank you again for your valuable input.

Gutierrez-Sigut, E., Payne, H., & MacSweeney, M. (2015). Investigating language lateralization during phonological and semantic fluency tasks using functional transcranial Doppler sonography. *Laterality: Asymmetries of Body, Brain and Cognition*, *20*(1), 49-68.

Woodhead, Z. V., Rutherford, H. A., & Bishop, D. V. (2018). Measurement of language laterality using functional transcranial Doppler ultrasound: a comparison of different tasks. *Wellcome open research, 3*.

Reviewer 2

Reviewer's comment 1: Research question

I can see the paper has already been through several rounds of revision; I'm not sure whether the earlier rounds are visible – am new to PCI and didn't see them, but that may be as well as I can take a completely fresh look.

I'm sorry to raise a pretty fundamental question at this point, but I think we need a better rationale for the study. Currently it has rather a so-what feel to it. I feel it could be improved to make it more compelling, and so will share my specific thoughts on how it might be improved.

We are told: "Studying writing using a smartphone is important as more than 6 billion people owned a smartphone in 2021 and spent 2.5 hours using it per day."

But I can't grasp why the lateralisation of smartphone writing matters. There's also an ancillary question about lateralisation in left- and right-handers, which I think has more merit, but it is not well integrated with the rest of the rationale.

Suppose the first hypothesis were disconfirmed and it was found that smartphone writing was either more or less lateralised than handwriting. Would that affect how we theorise about cerebral lateralisation? If the study found that lateralisation was greater or less for smartphone writing than for handwriting, I am not sure how we'd interpret that, given the particular features of smartphone writing that differentiate the tasks.

Authors' response:

We thank Prof. Bishop for underscoring the importance of justifying the aims of our studies. In an effort to address the reviewer's concerns, we have substantially reworked our introduction in this revised version, adding more justification for our choices. Additionally, to maintain the manuscript's focus on cerebral lateralization, we have removed the paragraph discussing the localization of different brain areas involved in writing.

The first concern of the Reviewer is why the lateralisation of smartphone writing matters. To address this concern, we have replaced the first three paragraphs of our manuscript to highlight the importance of studying this question, as follows (the second new paragraph, is placed in the fourth paragraph on the introduction, for better flow):

Original paragraphs:

> *"The majority of humans have been found to exhibit left cerebral lateralization during oral language production (Knecht et al., 2000; Kherdr et al., 2002; Loring et al., 1990; Petit, Badcock, & Woolgar, 2020). A similar left cerebral lateralization pattern is found for written language production (Kondyli, et al., 2017; Menon & Desmond, 2001; Papadatou-Pastou et al., 2022; Planton et al., 2013). These findings, however, come from studies with handwriting tasks, a more dated method of writing given the rising popularity of electronic devices, such as the smartphone (Statista, 2022). This shift is not reflected in the neuroscientific literature as no study to date has investigated the cerebral laterality of language during writing using a smartphone.*
>
> *The use of smartphones has dramatically changed our communication. Since the first iPhone was introduced in 2007, people have increasingly adopted smartphones. The invasion of smartphones to our lives is evident in the report that over six billion people used a smartphone in 2021 (Statista, 2022). This is unsurprising given the great -and affordable- utility the smartphone offers, with capabilities of remote video calls to family and friends, access to distant learning for students, and remote work for employees.*
>
> *Another major utilisation of smartphones is texting. Texting is a novel method of writing using the touch-screen smartphone keyboard. Writing using a smartphone has similarities to handwriting, but it is a distinct skill. Handwriting is a complicated process that involves (i) fine motor skills to form letters and their appropriate placement on the paper (e.g., Mangen & Velay, 2010), (ii) the integration of visual and*

sensory feedback as the letter is written (e.g., Mangen & Velay, 2010), and (iii) access to long-term memory to recall the shape of the letter with all of its intricate details. In contrast, typing involves typing the key on the keyboard, which requires access to the long-term memory (to recall the configuration of the keys), and letter formation is done automatically. A surprising positive application for the smartphone keyboard technology came from Japan with the development of a 10-key layout (Lee, 2021). This layout allows the user to write Japanese faster than handwriting by typing a combination of the elements presented on her screen instead of forming the intricate Kanji characters on a paper (Lee, 2021). Given the widespread adoption of smartphones for texting, the distinct characteristics of writing with the smartphone, and the possible advantages that smartphone writing has over handwriting, comparing the neurophysiological aspects of it to traditional handwriting is worthwhile. Specifically, we will compare the cerebral laterality during handwriting and during writing using a smartphone."

New paragraphs:

(page 2, line 1)

"The majority of humans exhibit left cerebral lateralization during oral language production (Kherdr et al., 2002; Loring et al., 1990; Petit, Badcock, & Woolgar, 2020). This left cerebral lateralization pattern also manifests in written language production (Kondyli, et al., 2017; Menon & Desmond, 2001; Papadatou-Pastou et al., 2022; Planton et al., 2013). However, these findings on written language production predominantly derive from studies that employ handwriting tasks. Given the rising popularity of electronic devices, such as smartphones, for written communication, this approach may now be considered somewhat dated. To illustrate, over six and a half billion mobile phone subscriptions were in place in 2022 (Statista, 2023). In other words, despite the significant shift towards digital means of written language production, current neuroscientific literature doesn't reflect this change, with no study yet investigating cerebral laterality of language during writing using a smartphone.

(page 4, line 8)

Writing using a smartphone (i.e., texting using the touch-screen smartphone keyboard) has similarities to handwriting, but it is a distinct skill. Handwriting is a complex process that involves (i) fine motor skills for letter formation and appropriate placement on the paper (e.g., Mangen & Velay, 2010), (ii) the integration of visual and sensory feedback (e.g., Mangen & Velay, 2010), and (iii) access to long-term memory where letter shapes are stored. In contrast, typing on a smartphone keyboard involves pressing the keys on the touch screen, which does require access to long-term memory for key configuration, but in this case letter formation is done automatically. It is worth noting that there are also distinct advantages of smartphone writing compared to

> *traditional handwriting. For example, a positive application of smartphone keyboard technology came from Japan with the development of a 10-key layout, which allows users to type combinations of elements presented on the screen to write Japanese faster than handwriting the intricate Kanji characters on paper. Given the widespread adoption of smartphones for texting, the distinct characteristics of writing with the smartphone, and the possible advantages that smartphone writing has over handwriting, investigating writing using a smartphone seems worthwhile. Expanding our knowledge on language lateralization to address writing using a smartphone can provide insights into how the neurobiological basis of language functions compares between traditional means of writing and modern, digital means. As there literature on the overall neurobiological basis of writing using smartphones is very limited, studying the cerebral lateralization of written language production using smartphones indeed seems to be a good starting point. In addition, should the lateralization of handwritten language and writing using smartphones be found to be comparable, then the laterality of writing could be studied using smartphones, which allow for off-site data collection. Laterality data per se can be collected remotely, for example through remote EEG using EmotivPRO Builder (reference for EmoticPRO Builder) and EmotivLABS (Williams et al., 2023). ,.*

Because of the restructuring of the first paragraph of the introduction, the first sentence of the now second paragraph was adjusted slightly for better flow:

Original sentence:

> *Cerebral lateralization of language has mainly been investigated using functional magnetic resonance imaging (fMRI).*

New sentence (page 2, line 13):

> *The cerebral lateralization of written language has primarily been investigated using functional magnetic resonance imaging (fMRI).*

We would also like to note that our work on the laterality of writing using smartphones primarily addresses a basic research question, with a significant part of our motivation stemming from scientific curiosity.

The second concern that the reviewer raised is that our decision to compare left- and right-handers is not integrated with the rest of the rationale. To address this point, we have now removed the following paragraph from our manuscript:

> *"Cerebral lateralization of language processes is related to handedness. The proportion of left- to right-handers is around 1:9 in the general population (Papadatou-Pastou et al., 2020), while 90% of right-handers but only 70% of left-handers exhibit*

> *left-hemispheric language lateralization (Carey & Johnstone, 2014). An implication of that is that only 30% of left-handers (compared to 90% of right-handers) have their language processes located in the contra-lateral hemisphere to their writing hand (Carey & Johnstone, 2014), while the movements of a limb are predominantly controlled by the contra-lateral hemisphere. Moreover, stronger left-handed preference is correlated to a higher chance of atypical cerebral lateralization for language (Somers et al., 2015), while individuals with more atypical cerebral lateralization have a higher chance of being left-handers (Mazoyer et al., 2014).”*

and replaced it with the following paragraph (page 5, line 9):

> *“Consequently, handedness differences have not been previously studied in this specific context. Investigating handedness differences in the cerebral laterality of writing using a smartphone is important, as left-handers and right-handers are consistently found to differ in the cerebral organisation of other language functions, such as oral language production and comprehension (e.g., Knecht et al., 2000; Packheiser et al., 2020), but also in the context of written language via handwriting tasks, as discussed above (Kondyli et al., 2017; Papadatou-Pastou etal., 2022). Another reason why comparing left-handers to right-handers is deemed worthwhile in the context of the cerebral lateralization of writing, is that 90% of right-handers but only 70% of left-handers exhibit left-hemispheric language lateralization (Carey & Johnstone, 2014). An implication of that is that only 30% of left-handers (compared to 90% of right-handers) have their language processes located in the contra-lateral hemisphere to their writing hand (Carey & Johnstone, 2014), while limb movements are predominantly controlled by the contralateral hemisphere. Furthermore, left-handers account for approximately 10% of the population (Papadatou-Pastou et al., 2020), thereby necessitating consideration of their phenotypical variation to thoroughly understand the function of the human brain. Indeed, left-handers have been identified as a compelling and widely available, yet largely untapped, resource for neuroscientific studies (Willems, Van der Haegen, Fisher, & Francks, 2014). More recently, Bailey et al. (2020) quantified this observation, showing that only about 3% of participants in neuroimaging studies, from whom handedness data are available, are adextral. Bailey et al. (2020) thus also advocate for a more balanced consideration of handedness.”*

The reviewer also raised a concern about the potential interpretation of our findings with regards to Hypothesis 1 (“Suppose the first hypothesis were disconfirmed and it was found that smartphone writing was either more or less lateralised than handwriting. Would that affect how we theorise about cerebral lateralisation? If the study found that lateralisation was greater or less for smartphone writing than for handwriting, I am not sure how we’d interpret that, given the particular features of smartphone writing that differentiate the tasks.”). To address this concern we have replaced the following paragraph (last paragraph of the introduction before the hypotheses):

"The present study aims to investigate the cerebral laterality of writing using two different means of writing: handwriting and writing using a smartphone. We expect that the language network will be the same for the two methods of writing and thus the cerebral lateralization during those tasks will not show significant differences. We also aim to investigate possible handedness differences in the cerebral laterality of these two writing methods."

With the following paragraph (page 6, line 5):

"The present study aims to investigate the cerebral laterality of writing using a smartphone, by comparing it with the cerebral laterality of the best available benchmark task, namely handwriting. At present, there is no data available on other writing modalities. Oral language tasks, on the other hand, are not directly comparable due to differences like the absence of a motor component in oral language. Regardless, we here attempt to isolate the linguistic component of writing (both in the context of handwriting as well as in the context of typing in a smartphone screen), by using motor control tasks designed for the motor movement involved in each mode of writing. We expect that the language network underlying writing will be consistent across the two modes of writing and thus the cerebral lateralization during those tasks will not show significant differences. If our hypothesis is not supported, and we find differences between the lateralization of the two writing modes, then this could guide future studies using neuroimaging techniques with higher spatial resolution to localise the different networks subserving the linguistic component of the two writing modes. In addition, we aim to explore potential handedness differences in the cerebral laterality of writing using a smartphone."

Bailey, L. M., McMillan, L. E., & Newman, A. J. (2020). A sinister subject: Quantifying handedness-based recruitment biases in current neuroimaging research. *European Journal of Neuroscience*, 51(7), 1642-1656.

Carey, D. P., & Johnstone, L. T. (2014). Quantifying cerebral asymmetries for language in dextrals and adextrals with random-effects meta analysis. *Frontiers in Psychology*, 5, 1128.

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., ... & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, *123*(12), 2512-2518. Doi: 10.1093/brain/123.12.2512

Packheiser, J., Schmitz, J., Arning, L., Beste, C., Güntürkün, O., Oklenburg, S. (2020). A large-scale estimate on the relationship between language and motor lateralization. *Scientific Reports 2020* 10:1,10(1), 1–10. https://doi.org/10.1038/s41598-020-70057-3

Papadatou-Pastou, M., Sampanis, P., Koumzis, I., Stefanopoulou, S., Sousani, D., Tsigkou, A., & Badcock, N. A. (2022). Cerebral laterality of writing in right-and left-handers: A functional transcranial Doppler ultrasound study. *European Journal of Neuroscience*, 56(2), 3921-3937.

Reviewer's comment 2: Task analysis

The potential for confounds is emphasised by the way in which 'smartphone writing' is defined in the study: using the thumb of one hand to select letters. First, this means you have an inevitable confound with left- and right-handers using different thumbs. Second, I wondered how many people would be familiar with typing that way. Do you have to turn the phone to landscape direction or change settings to use it effectively? Don't many people who use thumbs use both thumbs? Essentially, if this is an unfamiliar approach to writing then it seems to introduce a confound, in terms of the need to learn how to do it. And, if the subjects are used to writing on a phone with a different method, then they would have to suppress their natural tendencies.

Another factor is that in conventional oral Word Generation, the words are generated during a silent interval, during which laterality is measured, whereas it seems like with writing tasks envisaged here, the writing starts immediately after the cue. So it's hard to know how far any difference seen in a written condition is just because the words are written immediately.

Pondering on these issues made me wonder why, if the main interest is in studying laterality in nonvocal word generation/spelling, the authors hadn't gone for the option of testing people who could type on a regular keyboard. The advantage of typing (assuming you select subjects who can touch type using all fingers) is that both left- and right-handers use both hands – presumably to a similar extent, given keyboard design

Essentially, the possible tasks can be broken down into stages as follows:

Conventional word generation: generate words, speak them after a delay.

Handwriting: generate words, convert word into letters in correct sequence (spelling), activate motor program to create written representations. This will inevitably introduce confound with handedness, given the movement of one hand will increase blood flow to the contralateral side. And indeed, if I've understood it correctly, in their preprint, Papadatou-Pastou did find R-sided laterality in left-handers on this task – though less extreme than for a nonverbal motor control task.

Smartphone with one thumb: generate words, convert word into letters in correct sequence (spelling), use one thumb to select letters from keyboard. This will also introduce confound with handedness. It may also introduce confound with task novelty.

Typing on keyboard: generate words, convert word into letters in correct sequence (spelling), use both hands to type response on keyboard. This avoids confound with handedness. If touch typists are recruited, task novelty should not be a factor.

(You could also use both thumbs used for smartphone keyboard in landscape orientation, but that may not be familiar for all users, so again there is a novelty confound).

So if you wanted to see how far word generation in a non-spoken modality was similar to conventional spoken word generation, then typing would provide a method that did not need a subtracted condition to control for lateralised hand movement. I realise that is not the question the authors have started with, but I'd personally be intrigued to see if those tasks were equivalent, and whether there was greater dissociation between tasks in left-handers. Compared with what is proposed by the authors, there would be far fewer confounds and so it could be easier to interpret in terms of brain organisational differences in left- and right-handers – which I do think is an important and tractable question.

So I think there would be value in a direct comparison of oral word generation and typed words on a keyboard, though you would need to modify the oral task to start immediately after the cue so that the timing of the two tasks is comparable. There is precedent for this: Gutierrez-Sigut, E., Payne, H., & MacSweeney, M. (2015). Investigating language lateralization during phonological and semantic fluency tasks using functional transcranial Doppler sonography. Laterality, 20(1), 49–68.
https://doi.org/10.1080/1357650X.2014.914950

Authors' response:
We thank the reviewer for suggesting further ideas for improving our proposed study. To address these ideas, we need to provide the reviewer with more context on the present study and its position within the broader landscape of research conducted in our laboratory.

As the reviewer has noted, we are currently focusing our studies on cerebral lateralisation of non-verbal modes of language. The three main modes being used in our everyday life are 1) handwriting, 2) typing using a computer keyboard, and 3) typing using a smartphone. Handwriting is a central component in all of our studies (either as the variable of interest or as a comparison task), while typing using a computer keyboard vs. handwriting is currently under investigation is an in-principle accepted Registered Report (please find the Stage 1 manuscript here: https://osf.io/9qmge/?view_only=62cd4305f2d24ea8aae8810bb659e407). In other words, we are already conducting a study investigating writing on a PC keyboard, as the reviewer suggests.

The reason we chose to study the two different methods of typing, using a PC keyboard and using a smartphone, in separate studies was to avoid excessively long data collection sessions and prevent overworking our participants. Currently, the fTCD testing time for the PC typing study runs for approximately 1 hour and 20 minutes. Therefore, it was not practically feasible to add additional tasks, such as an oral generation task or a smartphone writing task. This decision was made to ensure the feasibility and manageability of the study.

When it comes to oral language, previous studies coming from our lab (i.e., Kondyli et al., 2017; Papadatou-Pastou et al., 2022) have investigated oral language in comparison to handwriting. However, in the context of the present preregistration, we chose not to include an oral language task, aiming to keep the testing session within 1 hour and 20 minutes. Additionally, as explained above and now incorporated into the text of our manuscript, we believe that handwriting, rather than oral language, serves as the most suitable benchmark task. See page 6, line 5:

> *"The present study aims to investigate the cerebral laterality of writing using a smartphone, by comparing it with the cerebral laterality of the best available benchmark task , namely handwriting. At present, there is no data available on other writing modalities. Oral language tasks, on the other hand, might not be directly comparable due to differences like the absence of a motor component in oral language. Regardless, we here attempt to isolate the linguistic component of writing (both in the context of handwriting as well as in the context of typing in a smartphone screen), by using motor control tasks specifically designed for the motor movement involved in each mode of writing."*

Overall, other projects within our lab already compare oral language to handwriting as well as handwriting with typing using a PC keyboard. Therefore, in this new registration, we would like to focus on typing using a smartphone and compare it with handwriting. We also note that Reviewer 1 has approved the methodology. Given our rationale for the methodology and this approval, we'd prefer to keep the methodology as is.

With regards to the more specific concerns of the reviewer:

(a) Left- and right-handers using different thumbs acting as a confound: Indeed, left-handers will use their left thumb, while right-handers will use their right thumb. However, this will also be the case for the motor control task that participants will perform ("*Participants will be instructed to tap randomly on the smartphone keyboard using the thumb of their dominant hand after the cue letter is presented*", see page 11, line7). The activation that corresponds to the control task will later be subtracted from the activation that corresponds to the typing task, thereby isolating the linguistic component of language. Therefore, we do not expect this to be a confound.

(b) Typing with only one thumb and whether this is familiar: The reason for instructing all participants to hold the smartphone with one hand and type using their thumb is to ensure a consistent method of typing across all participants and to avoid introducing further confounds. This method was chosen because, compared to using both thumbs, using one thumb is the most comparable method to handwriting given that both methods involve using only one hand. We acknowledge that this method of writing might be novel to participants who are accustomed to using both thumbs when typing on a smartphone. To address this, we will assess their ability to type using a smartphones as an inclusion criterion. As described in our methods, participants who fail to reach a minimum word composition speed of 12 words per minute while holding their phones with their dominant hand and writing using their thumb will be excluded from the study. Below is the pertinent passage from the manuscript, which describes the inclusion criterion (page 8, line 1):

"Ability to operate a smartphone keyboard using their thumb on their dominant hand. We will test this by having participants perform a word composition task using a smartphone keyboard and measure their Words Per Minute (WPM) typing speed. For this word composition task, participants will be given 60 seconds to write as many words as they can think of as fast and accurately as they can. The total amount of correctly written words (correct orthography but without taking into consideration the stresses used in written Greek) will be measured and this will be the WPM score of each participant. We will set a baseline of 12 WPM because that translates to three words in 15 seconds which is the time window that our participants will have during the experiment to compose words starting from the cue letter."

c) We have now added to the manuscript that the device will be held in portrait direction (page 11, line 7).

d) Immediate word generation (not after a given interval): Indeed, the words are generated immediately when one employs writing tasks. This is the case for both the handwriting and the writing using a smartphone task, so there will be no confound there. Moreover, as the Reviewer also points out, recent fTCD literature has started to use whispering while generating words (and indeed we use whispering rather than silent generation in another registered report in our lab where we test oral language in comparison to handwritten language, see here for the osf.io page of the study: https://osf.io/56vnq/).

Reviewer's comment 3: Reliability and positive controls

There's one key problem with the design as it stands. There is a prediction of a null result. There are two reasons why two measures may not differ statistically: one is because they really are equivalent, and the other is because they are so unreliable that genuine differences are masked. I'm not all that familiar with Bayes factors, but I don't think they really solve that issue. Minimally, I think we need to have some planned analyses that consider the reliability of tasks. A classic approach to this is Bland-Altman analysis (see references). Although I have used that in some studies, I find the visual representation of information confusing, and think the same point can be made using scatterplots. We had exactly this criticism In our work on bilingualism, where a reviewer argued that a lack of difference between Lis in two languages could be just down to poor reliability, and we were able to counter that argument by reporting the data as scatterplots with individuals showing standard errors. See figure 3 here: https://wellcomeopenresearch.org/articles/1-15 for an example. I also recommend plotting the time course of the blood flow velocity (see figure 2 from the same paper), as this can establish whether the maximum differences between L and R in different tasks are occurring at a similar time point.

In fact, if you were to move away from the approach that requires control tasks, that provides an opportunity for an even stronger test, because you could just have two blocks for each of two tasks, and so you'd have an index of within-task test-retest reliability. The correlation between block 1 and block 2 for the same task establishes a lower bound for the cross-task agreement. And in effect this acts as a positive control – i.e. if you had low correlations for two runs of the same task, this would indicate that the methods were not suitable to show the effects of interest. This is a realistic concern, because there will be motor movements during the interval of word generation, and these could make the signal noisier than it would be with conventional oral WG.

Authors' response:

To address the reviewer's concerns about the reliability of our measures and following her suggestions, we have added a number of different ways to assess reliability. Let us note that we opted to implement a split-half reliability test instead of testing the correlation between the blocks of tasks because blocks may be confounded by time (one block following the other)t. To summarise, we will take the following steps (page 13, line 9):

> *"To address the reliability of our measures, we will:*
> *(i) provide a scatterplot showing individual mean LIs for all six LIs with horizontal and vertical error bars denoting standard errors, following Bishop et al. (2021), in order to visually inspect reliability.*
> *(ii) plot left and right hemisphere activation as a function of epoch time for all four tasks (namely Handwriting, Handwriting Control, Typing, Typing control) as well as for the two difference scores (Handwriting minus Handwriting Control and Typing minus Typing Control), again following Bishop et al. (2021), to visually establish whether the maximum differences between left and right channels in different tasks are occurring at a similar time point.*
> *(iii) perform a split-half reliability test separately for each task (namely Handwriting, Handwriting Control, and their difference score; Typing, Typing Control, and their difference score). If the correlation is below 60%, we will assume that our tasks are not suitable to show the effects of interest and not include these tasks in further analyses."*

We understand the reviewer suggested the correlation between blocks approach for the case when we do not use control tasks, which is not our case, as we will use control tasks. However, we believe this will be a useful guide with regards to the reliability of our measures even in our case.

Reviewer's comment 4: Predictions re handedness groups

I think there is real interest in questions about whether the pattern of lateralisation differs between left- and right-handers, which was one question looked at in a recent study by Parker et al (2022). The extension to a task involving writing does give this extra interest; I don't think that has been properly looked at previously. However, the sample size of 16 L handers and 16 R handers does seem too small to demonstrate a handedness effect on Word Generation, let alone a differential handedness effect on another task.

This is output from R's pwr package, using means from Parker et al on Word Generation to estimate the true effect:

n = 16

delta = 0.86

sd = 1.8

sig.level = 0.05

power = 0.3729685

alternative = one.sided

NOTE: n is number in *each* group

So there is only a 1/3 chance of showing a handedness effect on the well-established WG task.

Following Mazoyer et al, it may be possible to increase power by focusing on very strong left-handers – according to Mazoyer et al, they are more likely to have atypical laterality, so this approach could amplify the effect of interest.

This brings us to the next topic, the analysis.

Authors' response:

Following the reviewer's suggestion, we have increased the initial recruitment number to 20 individuals per handedness group (page 7, line 3) and our maximal number to 40 per group. We also increased the maximum amount of participants from 60 to 80 (page 7, line 11) We would like to clarify, though, that in our original manuscript we were not suggesting that we will find an effect with 32 participants (in total, 16 per group), but we were stating that we would follow the sequential procedure, which posits that it is unnecessary to begin analyzing before gathering data from at least 32 participants.

For Hypothesis 2 ("Right-handers will exhibit stronger left-hemispheric lateralization compared to left-handers for the linguistic part of writing using a smartphone") we have chosen a prior with a half-Cauchy distribution set with a width parameter of 0.94, which corresponds to Cohen's *d* for the lateralization difference of the linguistic component of writing between left- and right-handers in Papadatou-Pastou et al. (2022). Assuming that we expect the data to be in favour of the alternative hypothesis and that we might find a high variability in our sample size (between 2 and 3 standard deviations) we can expect to find a BF10 smaller than ⅓ with greater than chance probability with our initial minimum sample size of *n* = 16 (Phylactou & Konstantinou., 2022). Moreover, with these criteria, there is an above 70% chance of detecting evidence in favour of the alternative for our initial maximum sample size of 40 (20 in each group) and a greater than 80% chance for our updated maximum sample size of 80 (40 in each group; Phylactou & Konstantinou., 2022).

To further increase the likelihood of detecting handedness effects, and following the reviewer's suggestion, we altered our inclusion criteria for handedness. To that end, we have chosen to exclude ambidextrous participants. Specifically, we will exclude participants with an EHI score of 40-60%. The text now reads (page 8, line 11):

> *"Participants will have an EHI score of 0-40% (left-handers) or 60-100% (right-handers)."*

We would also like to briefly discuss the use of power analysis within the Bayesian framework. A distinction that needs to be made is that, unlike frequentist approaches, the Bayesian approach that we will implement for our study is not relying on statistical power but instead utilises a prior distribution for testing hypotheses. The Bayesian approach does not require power analysis because it does not rely on null hypothesis significance testing with fixed thresholds, such as *p*-values. Instead, Bayesian analysis provides a direct assessment of the evidence in favor of different hypotheses by evaluating the posterior probabilities. To put it another way, in Bayesian statistics, the focus is on updating prior beliefs about the null hypothesis using observed data, rather than making decisions based on *p*-values or significance testing. In other words, in Bayesian statistics the goal is not to reject or accept a null hypothesis but rather to update prior beliefs. Bayesian methods utilize other components, such as prior distributions and decision criteria like Bayes factors, to make decisions about how strong the evidence is. This approach allows researchers to make more nuanced interpretations of the data and draw conclusions based on the strength of evidence in the posterior distribution.

Please allow us to also justify our choice of the Bayesian approach in contrast to a frequentist approach. One important reason is the fact that this approach allows us to quantify evidence for the null hypothesis (see Hypothesis 1). The other reason is the fact that the Bayesian approach is considered good practice within the open science movement. For example, Bayesian methods allow for flexible modelling and updating of prior beliefs based on observed data. This flexibility promotes transparency as researchers can explicitly state their prior assumptions and update them with new evidence. By providing a clear and flexible framework, the Bayesian approach encourages transparency in research methodology and analysis.

Phylactou, P., & Konstantinou, N. (2022). Bayesian T-Test sample size determination: Reference tables for various bayes factor thresholds, effect sizes, sample sizes, and variance assumptions. *PsyArXiv.(10.31234/osf. io/jnp8c)*

Reviewer's comment 5: Statistical analysis

In effect, it looks as if the predictions would fit into a 2 x 2 within x between Anova design, with task as a within-subject variable and handedness between-subjects. This may be a rather old-fashioned approach these days when everyone is using linear mixed models. But I think it's worth considering something along these lines, whether implemented as Anova or something more sophisticated, because the significance of the interaction term would be testing the key prediction of more dissociation between lateralities for left-handers.

A practical problem for casting the analysis this way is that if we focus on the interaction between handedness and task, then the issues regarding power become even more serious: unfortunately interaction terms tend to require much larger sample sizes than main effects.

Authors' response:
We thank the reviewer for their comment and suggestion. We would like to provide a justification for our analyses of choice and offer to implement their suggestions in the exploratory part of our analysis.

We have chosen to perform two *t*-tests because the *t*-test directly compares the contrasts proposed in the hypotheses. This approach helps maintain a focused registration and provides evidence for our research question through simple analyses. Opting for a simpler model is crucial not only for the aims of a registered report but also for Bayesian analyses in general. In the Bayesian framework, more complex designs like mixed models can introduce greater statistical flexibility, potentially leading to false conclusions due to overfitting of the models (Blanchard et al., 2018).

Nevertheless, in response to the reviewer's suggestion, we will include a mixed effects model as an exploratory analysis of our data in our Stage 2 manuscript. Specifically, for the exploratory analyses, we will perform a mixed effects linear model with method of writing as the within participants variable and handedness group as the between participants variable. This analysis will also be performed within the Bayesian framework and any consideration with regards to overfitting the data will be discussed. We do not wish to register this analysis, to keep our registration focused, as mentioned above. Additionally, we want to highlight that our raw dataset will be uploaded to osf.io, allowing interested parties to conduct further analyses if they wish.

Blanchard, T., Lombrozo, T., & Nichols, S. (2018). Bayesian Occam's razor is a razor of the people. *Cognitive Science*, *42*(4), 1345-1359.

Reviewer's comment 6: Overview

I think there is the core of a good project in this paper, but the research question needs tightening up, and, unfortunately, I think that a much larger sample size will be needed to answer it convincingly.

Authors' response:
We would like to express our gratitude to the Reviewer for their insightful suggestions, which have significantly improved our manuscript. We have strived to tighten up the rationale of the study. Furthermore, we explain why our now larger sample size is sufficient under the Bayesian approach. We trust that these revisions have elevated the quality of the manuscript.