
FINAL REPORT

S-R-compatibility task: Effect of stimulus size on horizontal response location – A replication study

AUTHORS: ÖZENOGLU, Ö., BAX, T., EWERT, M., BUB, N.

CORRESPONDING AUTHOR: ÖZENOGLU, Ö. (OOEZENOGLU@UOS.DE)

Abstract

The *A Theory of Magnitude* (ATOM) model proposes that time, space, number, and other dimensions are linked through a magnitude system within the parietal lobe (Walsh, 2003). One of the model's predictions is that there are intrinsic reciprocal interactions across magnitude dimensions.¹ If this were the case, manipulating one dimension should correspond with an interfering perception of the other dimension. Indeed, both neurophysiological and behavioral research provides evidence in favor of the ATOM model and its prediction. In behavioral research, the literature contains various studies that observed so-called SNARC effects (Spatial-Numerical Association of Response Codes, Dehaene et al, 1993) and size-congruency effects (e.g., Henik et al., 1982). A recent study by Wühr and Seegelke (2018) investigated yet another – rather unexplored – combination of ATOM's implications, namely whether there are compatibility effects between physical stimulus size and spatial response location. Specifically, the study results suggest that small objects are associated with the left side and large objects with the right side. As cumulative evidence in favor of the S-R compatibility effect would extend ATOMs framework, we considered a direct replication attempt as meaningful. In our experiment, participants responded to a small or large stimulus with either their right hand or their left hand while we varied the S-R mapping condition within participants. Our results support the hypothesis of Wühr & Seegelke (2018) that large objects are associated with the right side and small objects are associated with the left side. However, the observed effect was not more pronounced in right-hand responses. Nevertheless, our replication supports the existence of a general magnitude code, as proposed by ATOM, i.e., that there is an association between physical size (SIZE) and horizontal response location (SPACE).

Keywords: ATOM; compatibility; congruency; stimulus size; response position; SNARC; replication

¹ Put differently, ATOM proposes that we have an internal representation of “magnitude” across axes (e.g., horizontal or vertical), where magnitude increases or decreases towards the poles of an axis. For instance, if we compare the numbers 1 and 100, we say that 1 is smaller than 100, and 100 is larger than 1. The theory further states that other dimensions (such as space, physical size, etc.), may influence the perceived magnitude of, say, the numerical size.

This study was preregistered before data collection started to assure open and replicable scientific results. Please refer to the [preregistration](#) for more information. All materials are available in our [GitHub repository](#).

Introduction

This introduction is heavily inspired by Wühr & Seegelke (2018), and likewise the “Data analysis” and “Results” sections by Franke & Roettger (2019).

Walsh (2003) proposes in his *A Theory of Magnitude* (ATOM) model that time, space, number, and other dimensions are linked through a magnitude system within the parietal lobe of the human brain. In the subsequent years after his proposal, ATOM has been substantially supported by neuropsychological and neurophysiological data, leading to the idea of overlapping brain structures for the processing of time, space, and magnitude information in the human parietal cortex (e.g., Cohen Kadosh et al., 2007; Kaufmann et al., 2008; see Buetti & Walsh, 2009, for review). According to this theory, time, space and numbers are 1) influenced by each other and 2) processed by a common magnitude system. There has been growing evidence for each of these interaction effects, namely spatial-numerical effects (SNARC effect), spatial-temporal (STEARC² effect), and temporal-numerical (TiNARC³ effect). In the following paragraphs, we will take a closer look at SNARC and SNARC-like effects. To investigate these effects, researchers explored the interactions between NUMBER AND SPACE, NUMBER AND SIZE, and SIZE AND SPACE.

The first compatibility effect between numerical size (NUMBER) and horizontal response location (SPACE) was demonstrated by Dehaene, Dupoux & Mehler (1990). Results from later research showed that large numbers preferentially elicited a rightward response while small numbers triggered a leftward response (Dehaene et al., 1993). There have been several explanation attempts for these results. The most prominent idea is that these horizontal brain mappings of small numbers on the left and large numbers on the right result from writing, reading, and counting habits (e.g., Fischer, 2008).

The interactions between number magnitude (NUMBER) and physical size (SIZE) have been demonstrated in several experiments, especially through the size-congruity effect: When participants are presented with two numbers varying in numerical and physical size, their response is faster when the irrelevant physical size is congruent instead of incongruent with the to-be-judged numerical magnitude (Besner & Coltheart, 1979). This effect also occurs vice versa, i.e., when the irrelevant numerical size is congruent instead of incongruent with the to-be-judged physical stimulus size (Henik & Tzelgov, 1982).

This brings us to the last interaction: physical size (SIZE) and horizontal response location (SPACE). Some recent studies showed a SNARC-like compatibility effect between the *conceptual* size and horizontal response location. For example, Ren, Nicholls, Ma, & Chen (2011) demonstrated that participants left responses were faster to names of small objects (like “apple”) as compared to large objects (like “mountain”), while right responses were faster to names of large objects compared to small ones. This was one of the rare studies investigating SIZE AND SPACE, given that Ren et al. (2011) only looked at the *conceptual* size (see Experiment

² STEARC: spatial-temporal association of response codes

³ TiNARC: temporal-numerical association of response codes

4 of their study). In Experiment 2 by Ren et al. (2011), they slightly changed their experimental set-up: Now they compared *physical* and not only *conceptual* size by using filled circles of different sizes instead of words. The results uncovered a statistically significant compatibility effect for right-hand responses, meaning: Right-hand responses had a faster reaction time for large stimuli than for small stimuli. However, there were no significant results for left-hand responses. Nevertheless, these results indicate that the association between magnitude and horizontal response location may not be related to directional writing, reading, and counting habits, as the most prominent theory holds.

To consolidate the findings of Ren et al. (2011), Wühr & Seegelke (2018) conceptually replicated Experiment 2 of Ren et al. by conducting two experiments of their own. In Experiment 1, Wühr & Seegelke (2018) used a classic S-R compatibility task which required a right or left response to a single stimulus in each trial. The stimulus was either small or large, and the experiment was designed to test whether there exists a difference in response times to large or small stimuli, dependent on the side of the response (left or right). In this study, we aimed to directly replicate Experiment 1 by Wühr & Seegelke (2018).

Hypotheses

In accordance with experiment 1 by Wühr and Seegelke (2018), we addressed the following research hypotheses:

- I. Response times are faster in the compatible mapping condition than in the incompatible mapping condition.
- II. The S-R compatibility effect is more pronounced in right-hand responses.

Besides investigating the existence of a stimulus size-response location compatibility effect as addressed with hypothesis I. above, Wühr and Seegelke (2018) found that this effect is more pronounced in right-hand responses. Since what counts as “more pronounced” is ambiguous, and the authors did not elaborate on their interpretation, we decided to split this hypothesis in two. We reason that if the compatibility effect is more pronounced in right-hand responses, right-hand responses should be faster in the compatible mapping condition, and slower in the incompatible mapping condition compared to left-hand responses. Therefore, the discrepancy between the reaction times through manipulation should be larger for right-hand responses than for left-hand responses.

We tested the following hypotheses:

1. Response times are faster in the compatible mapping condition than in the incompatible mapping condition.
2. The compatible S-R mapping yields faster reaction times for right-responses than for left-responses.
3. The incompatible S-R mapping yields slower reaction times for right-responses than for left-responses.

Methods

Participants. 51 volunteers (19 female, 31 male, 1 no answer) with a mean age of 28 years (range 18-61 years) participated in the Online-Experiment. The premises for taking part in the experiment were either normal or corrected-to-normal visual acuity and good command of English. Participants needed to do the study on a laptop or desktop computer – phones and tablets were not allowed. These three premises were communicated in the study invitation, which was sent out via email and social media. Every participant was asked to take part in the experiment only once and was able to cancel the experiment at any time. Participants were neither informed about the purpose of the reversed mapping conditions nor were they informed about the objective of this study. There was no direct contact between participants and experimenters. The data were analyzed by the experimenters.

Materials. There were two imperative stimuli across the experiment: One small square (2 x 2 cm) and one large square (4 x 4 cm). In each trial, participants saw either one of the squares in the center of the screen and judged whether it is the smaller or the larger one. In both practice and main trials, the same two stimuli were shown in random order. In every trial, the square was shown in front of a light grey background (hex #f8f8f8). We created both stimuli by ourselves. They are available [here](#).

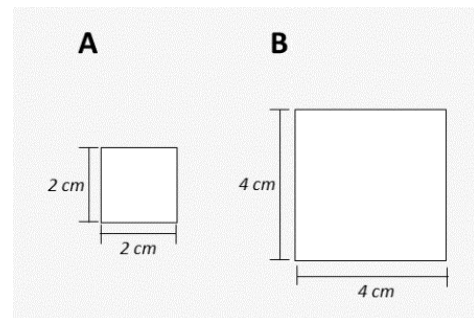


Figure 1: Two stimuli throughout the experiment. In each trial, either square A or square B is shown at the screen center (without height indication).

Study design. The study is a within-subjects design with two factors. The two factors are “mapping condition” (levels: compatible, incompatible) and “correct response location” (levels: right, left). In each trial, participants either saw a small square (2 x 2 cm) or a large square (4 x 4 cm) at the screen center. Their task was to judge whether the square is the smaller or the larger one (forced binary choice). In the compatible mapping condition, participants should have responded to the small square by pressing the “q” key and should have responded to the large square by pressing the “p” key. In the incompatible mapping condition, it was vice versa: Participants were asked to press the “q” key if they see the large square and the “p” key if they see the small square. In our experiment, we chose those keys because – as opposed to the original study’s keys (“tabulator” and “backspace”) – they do not differ in size. Moreover, both keys are on the same horizontal axis on the keyboard. We thought that it would not make a difference to the experiment as both keys are still associated with either side. The instructions were supplemented with an [image](#) of both stimulus sizes, similar to Figure 1.

Procedure. The experiment consisted of six parts:

- I. introduction & instructions
- II. practice phase (first S-R mapping)
- III. main test phase (first S-R mapping)

optional pause

- IV. practice phase (second S-R mapping)
- V. main test phase (second S-R mapping)
- VI. post-experiment questionnaire

In the first practice phase (II), participants completed ten trials with the first S-R mapping (2 stimuli x 5 repetitions). In the second practice phase (IV), there were 20 trials to complete with the second S-R mapping (2 Stimuli x 10 repetitions). Both main phases (III and V) consisted of 60 trials each (2 stimuli x 30 repetitions).

The relevant manipulation, namely reversing the mapping conditions, is within participants. In the first block (II. and III.), participants were not aware that a later switch in conditions will take place. After finishing the first main test phase, participants were informed about the reversed mapping and instructed to press the keys accordingly. As opposed to the original study, the order of mapping conditions (compatible – incompatible or incompatible – compatible) was not counterbalanced across participants. Because it was easier to implement in _magpie, we instead assigned the order of mapping conditions in an entirely random, ad hoc defined order at the beginning of the experiment. 19 participants (or 37%) had the compatible condition as their first S-R mapping, and 32 participants (or 63%) had the incompatible condition as their first S-R mapping.

All participants saw both stimuli (small and large square) throughout the experiment. Each square was shown in a random, ad hoc defined order.

Trial sequence. Each trial started with the presentation of a black fixation cross at the center of the screen (_magpie's default font and size). The fixation cross appeared for 1000 ms. Next, a square was presented at the screen center (either 2 x 2 cm or 4 x 4 cm). The stimulus was presented on a light grey background (hex #f8f8f8). The participants responded by pressing either the left “q” key or the right “p” key, with a maximum period of 2000 ms. If responded correctly within the time limit, the stimulus disappeared, and a blank screen was shown for 1500 ms. If the participant gave an incorrect answer, an error message “Incorrect answer!” appeared for 1500 ms, and afterward, the experiment automatically proceeded to the next trial. If the participant did not respond within the time limit, an error message “Please answer more quickly! (Press 'q' or 'p' to continue)” appeared and remained until either the “q” key or the “p” key was pressed. Both error messages were written in black and of _magpie's default font and size. The trial sequence was the same for both practice and main

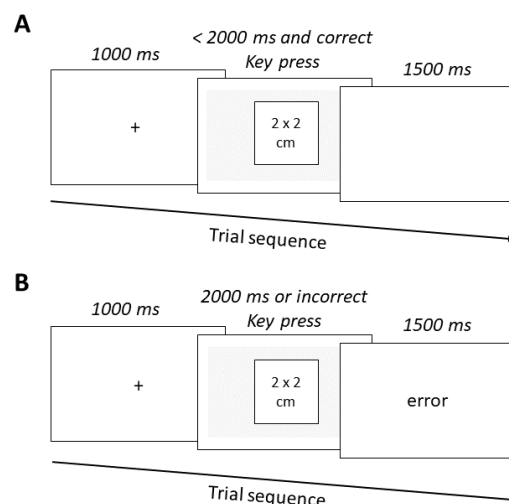


Figure 2: Trial sequence for both practice and main trials. Panel A depicts a correct trial (blank screen), panel B depicts an incorrect or 'too slow' trial (error message). Here, only the small stimulus is shown.

trials. A visualization of the two possible trial sequences is shown in figure 2. Note that for the large stimulus, the square of size 4 x 4 cm was shown.

Measured variables. The single dependent variable we measured for both confirmatory and exploratory analysis was the reaction time between stimulus onset and button press. Concretely, RT is a metric variable capturing the reaction times. Furthermore, we stored the mapping condition in the binary variable `CONDITION` (values: compatible, incompatible). This variable was manipulated by us. Furthermore, we measured whether each trial was correct or not. `CORRECTNESS` is a binary variable with values “correct” or “incorrect”. We used this variable to filter out incorrect trials. After filtering out incorrect trials, we created a binary variable `CORRECT_RESPONSE`, which had the values “left” or “right”. We created this variable based on the values of binary variables `CONDITION` (compatible, incompatible) and `EXPECTED` (small, big). The variable RT was used as the variate, and `CONDITION` and `CORRECT_RESPONSE` served as covariates (see Data analysis below).

For our exploratory analysis, we encouraged participants to indicate their dominant hand (right, left, I am able to use both hands equally well) in the post-experiment questionnaire. We stored the answer in the nominal variable `HANDEDNESS` (see Further exploration below).

Data exclusion. Following the original study by Wühr and Seegelke (2018), we excluded every individual trial faster than 100 ms and slower than 1500 ms (199 trials, $\approx 3\%$). As we were only dealing with reaction times and not error rates, we used the binary variable `CORRECTNESS` (correct, incorrect) to filter out incorrect trials. No data from practice trials entered the analysis.

Data analysis. For data analysis, we relied on the Stan modeling language (Carpenter et al., 2016) through the ``brm`` function of the ``brms`` package (Buerkner, 2016). For testing all three hypotheses, we made use of the ``compare_groups`` function of the ``faintr`` package introduced in Franke and Röttger (2019).

We fitted a Bayesian hierarchical model to reaction times (RT) as a function of dummy-coded factors `CONDITION` (reference level: compatible), `CORRECT_RESPONSE` (reference level: right), and their two-way interaction. We added random intercepts and slopes, allowing the two predictor variables of interest and their interaction to vary by participants (variable: `SUBMISSION_ID`). We excluded correlation coefficients. We used the default, flat priors of the ``brms`` package, namely a Student’s t -distribution($\nu = 3$, $\mu = 512$, $\sigma = 114.2$) for the mean of the reference cell (right-responses in the compatible condition), and a Student’s t -distribution($\nu = 3$, $\mu = 0$, $\sigma = 114.2$) for the standard deviation for the likelihood function. For the regression coefficients, as well as for the standard deviations for random effects, we also used the default priors of the ``brms`` package, which all were unbiased. We used this model for addressing all three hypotheses. We chose this model after comparing various models with the Leave-One-Out Cross-Validation (LOO-CV) method. Please refer to the [Model Comparison](#) script for the justification of our model choice.

Four sampling chains ran for 3000 iterations each, with a warm-up period of 1500, thereby yielding 6000 samples for each parameter tuple. To prevent divergent transitions, we

decreased the sampler's step-size by increasing `adapt_delta` to 0.99. Furthermore, we increased the number of tree depth to 15. We set a seed to 525.

For all relevant cell means and differences between them, we report the expected values under the posterior distribution and their 95% credible intervals (CrIs). For differences between cells, we also report the posterior probability that a difference δ is bigger than zero. If a hypothesis states that $\delta > 0$, we judge there to be compelling evidence for this hypothesis if zero is not included in the 95% CrI of δ .

Results

Reaction times were faster in the compatible mapping condition ($\mathbb{E}(\mu_{\text{compatible}}) = 532$, CrI = [505, 558]) than in the incompatible condition ($\mathbb{E}(\mu_{\text{incompatible}}) = 564$, CrI = [533, 595]). There is compelling evidence for this difference ($\mathbb{E}(\mu_{\text{incompatible}} - \mu_{\text{compatible}}) = 32$, CrI = [14, 48], $P(\delta > 0) = 0.99$). Given the data and the model, we conclude that there is evidence in favor of hypothesis 1.

In the compatible mapping condition, reaction times were faster if participants gave a right-response ($\mathbb{E}(\mu_{\text{compatible, right}}) = 530$, CrI = [504, 556]) than left-response ($\mathbb{E}(\mu_{\text{compatible, left}}) = 534$, CrI = [507, 561]). However, there is no sufficient evidence that this difference is larger than zero ($\mathbb{E}(\mu_{\text{compatible, left}} - \mu_{\text{compatible, right}}) = 4$, CrI = [-5, 15], $P(\delta > 0) = 0.81$). We conclude that the data and the model do not support hypothesis 2.

In the incompatible mapping condition, there was no substantial discrepancy between right-responses ($\mathbb{E}(\mu_{\text{incompatible, right}}) = 564$, CrI [533, 593]) and left-responses ($\mathbb{E}(\mu_{\text{incompatible, left}}) = 565$, CrI = [534, 596]) for the same condition. The model does not support this relationship either ($\mathbb{E}(\mu_{\text{incompatible, right}} - \mu_{\text{incompatible, left}}) = -1$, CrI = [-14, 11], $P(\delta > 0) = 0.43$). We conclude that the data and model do not support hypothesis 3.

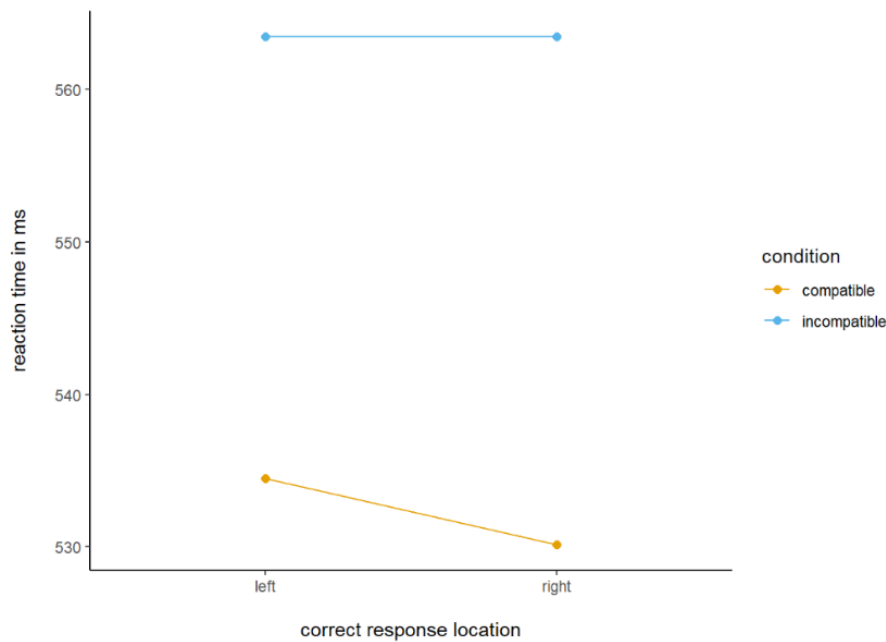


Figure 3: Mean reaction times (dots) as a function of condition (compatible, incompatible) and correct response location (left, right).

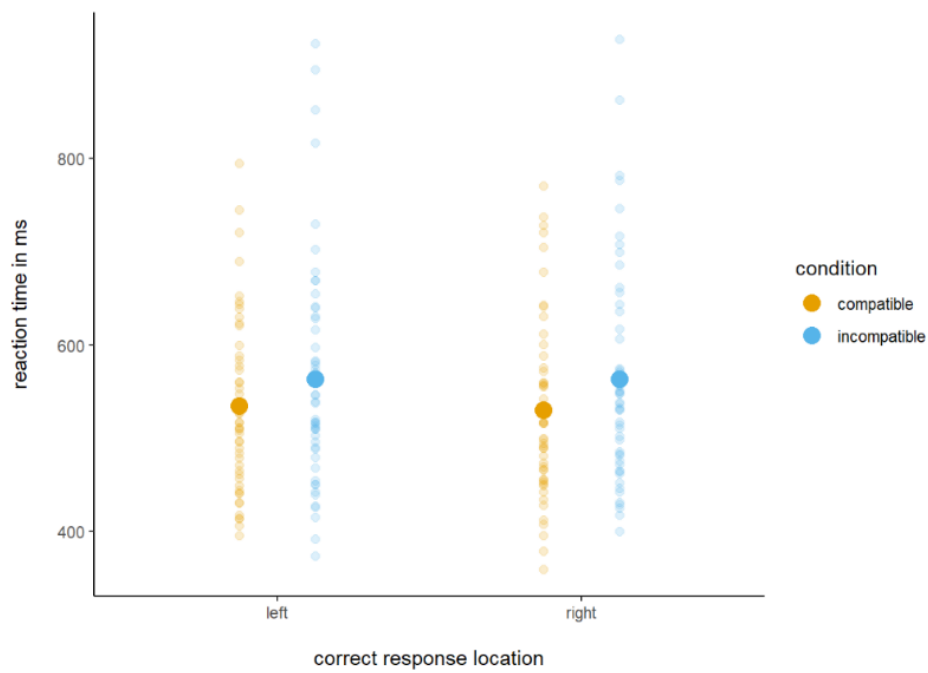


Figure 4: Mean reaction times (thick dots) for each combination of condition (compatible, incompatible) and correct response location (left, right). The transparent dots represent the mean reaction times for each participant. This plot is inspired by a tutorial of Franke and Röttger (2019).

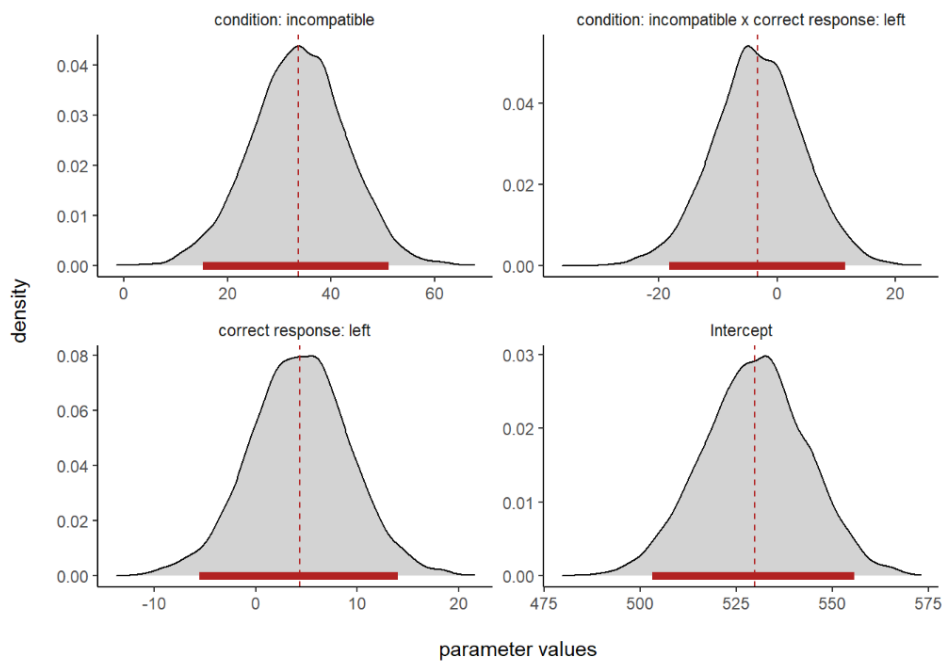


Figure 5: Posterior estimates of population-level effects. The dashed lines indicate the mean posterior estimates; the thick, horizontal lines indicate the 95% credible intervals. Note that the 95% credible intervals for the “correct_response: left” coefficient (bottom left) and the interaction term “condition: incompatible x correct_response: left” (top right) includes zero as credible value. This plot is inspired by a tutorial of Franke and Röttger (2019).

Further Exploration

Motivation. Wühr and Seegelke (2018) noted in the “Conclusion and directions for future research” section that the “participant’s handedness may modulate the effect as well” (p. 9). As their study results were only based on data from right-handed participants (N=24), the suggested internal representation of stimulus size and horizontal location may differ for left-handers. We therefore additionally investigated whether handedness modulates the S-R compatibility effect in a sense, that the internal horizontal representation might be switched. Concretely, we were interested in whether right-handers associate large objects with the right side, and small objects with the left side, and vice versa for left-handers. Questions 4 and 5 addressed this theory of ours. We thought that a positive outcome for questions 4 and 5 may inspire follow-up confirmatory research that investigates this relationship and ultimately may shed more light on the internal representation of magnitude.

It is further reasonable to think that – regardless of the stimulus size – left-handers respond faster with the left hand, and vice versa for right-handers. Therefore, in question 6 below, we tested whether, in both conditions, right-handers were faster in right-responses. Conversely, in question 7, we tested whether, in both conditions, left-handers were faster for left responses.

We addressed the following questions:

4. Do right-handers respond faster in the compatible mapping condition than left-handers?
5. Do left-handers respond faster in the incompatible mapping condition than right-handers?
6. Do left-handers respond faster to left-responses than to right-responses?
7. Do right-handers respond faster to right-responses than to left-responses?

Exploratory Analysis. We ran a Bayesian hierarchical model with variate RT and covariates CONDITION, CORRECT_RESPONSE, and HANDEDNESS (with three-way interaction). Furthermore, we specified by-participant random intercepts and slopes for the interaction of covariates CONDITION and CORRECT_RESPONSE. For testing the abovementioned questions, we again used the `compare_groups` function of the `faintr` package.

Four sampling chains ran for 3000 iterations each, with a warm-up period of 1500, thereby yielding 6000 samples for each parameter tuple. To prevent divergent transitions, we decreased the sampler’s step-size by increasing `adapt_delta` to 0.95. Furthermore, we increased the number of tree depth to 15. We set a seed to 525. Please refer to statistical analysis [made available here](#) for further information.

Results. In total, there were 44 right-handed, 1 both-handed, and 6 left-handed participants taking part in the experiment. We did not find any indication of evidence in favor of any of the former questions. A group comparison of the reaction times between left- and right-handers in the compatible condition did not yield reason to believe that right-handers might react faster than left-handers ($\mathbb{E}(\mu_{\text{compatible, left}} - \mu_{\text{compatible, right}}) = 6$, CrI = [-80, 88], $P(\delta > 0) = 0.56$). We conclude that the data and model do not support question 4.

A group comparison of the reaction times between left- and right-handers in the incompatible condition did not yield reason to believe that left-handers might react faster than right-handers ($\mathbb{E}(\mu_{\text{incompatible, right}} - \mu_{\text{incompatible, left}}) = -0.7$, $\text{CrI} = [-100, 97]$, $P(\delta > 0) = 0.5$). We conclude that the data and model do not support question 5.

For question 6, there was no substantial difference between groups ($\mathbb{E}(\mu_{\text{response:right, handedness:left}} - \mu_{\text{response:left, handedness:left}}) = 3$, $\text{CrI} = [-20, 25]$, $P(\delta > 0) = 0.61$), leading us to reject the idea that left-handers respond faster to left-responses than to right-responses.

Also for question 7, there was no substantial difference between groups ($\mathbb{E}(\mu_{\text{response:left, handedness:right}} - \mu_{\text{response:right, handedness:right}}) = 4$, $\text{CrI} = [-5, 12]$, $P(\delta > 0) = 0.81$), leading us to reject the idea that right-handers respond faster to right-responses than to left-responses.

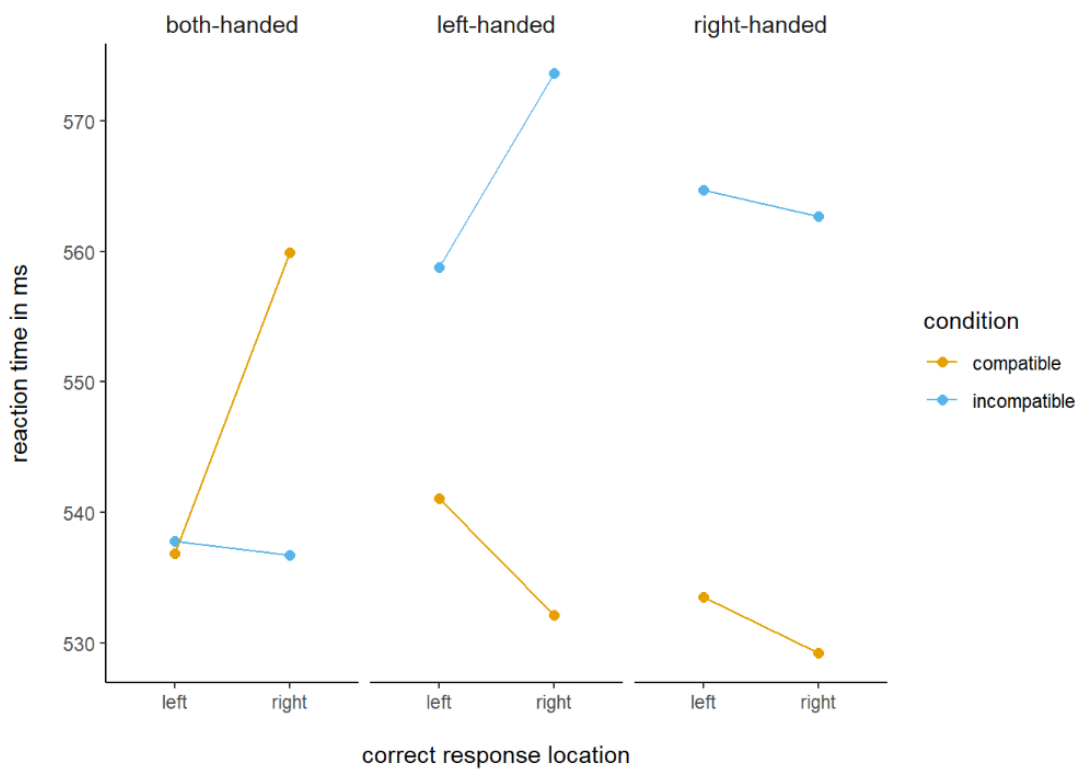


Figure 6: Overall mean reaction times as a function of handedness, condition, and correct response location. Contrary to what we have initially speculated, the data suggest that also left-handed participants reacted faster in the compatible mapping condition than in the incompatible mapping condition.

Conclusion

In this experiment, we directly replicated the compatibility effect between physical stimulus size and horizontal response location, which was previously shown by Wühr & Seegelke (2018), who themselves conceptually replicated this effect shown by a study from Ren et al. (2011). As well as the original study by Wühr and Seegelke (with a sample size of $N=24$ participants), our experiment ($N=51$) provides evidence for an association between smaller stimulus objects with left-hand responses and larger stimulus objects with right-hand responses (Hypothesis 1). Concretely, our results show that the participants responded faster in the compatible mapping condition than in the incompatible mapping

condition. This result further manifests the idea of intrinsic reciprocal interactions across magnitude dimensions, as suggested by ATOM.

However, our results do not confirm the second hypothesis of Wühr and Seegelke (2018), namely that the compatibility effect is more pronounced in the right responses than in left responses. Specifically, we did not find that right-hand responses were substantially faster in the compatible mapping condition than left-hand responses (Hypothesis 2). Likewise, we did not find that right-hand responses were slower in the incompatible mapping condition compared to left-hand responses (Hypothesis 3).

An exploratory analysis did not yield any indication to believe that the participant's dominant hand may play a role in explaining the observed S-R compatibility effect. As only 6 out of 51 (or 12%) participants indicated to be left-handed compared to 44 out of 51 (or 86%) right-handers, it would be insightful to investigate the relationship between reaction times, mapping conditions and handedness with data from an approximately even number of left-vs. right-handed participants.

Links

GitHub repository: <https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility>

Preregistration: <https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/blob/master/writing/02-preregistration-SR-Compatibility-Task.pdf>

Materials (stimuli and instructions image): https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/tree/master/experiment/02_main/materials

Experimental design: <https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/blob/master/writing/01-experimental-design-SR-Compatibility-Task.pdf>

Raw data (N=24) from Wühr and Seegelke (2018): <https://doi.org/10.5334/joc.19.s1>

Raw data of pilot study (N=4) by us: https://raw.githubusercontent.com/ooezenoglu/XPlab-2020-SR-Compatibility/master/data/01_pilot/01-raw-data-pilot.csv

Raw data of main study (N=51) by us: https://raw.githubusercontent.com/ooezenoglu/XPlab-2020-SR-Compatibility/master/data/02_main/01-raw-data-main.csv

Link to images provided in the final report: <https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/tree/master/writing/03-images>

Model comparison: https://htmlpreview.github.io/?https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/blob/master/analysis/01_pilot/01-SR-compatibility-model-comparison.html

Analysis script (confirmatory): https://htmlpreview.github.io/?https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/blob/master/analysis/02_main/01-SR-compatibility-analysis-main.html

Analysis script (exploratory): https://htmlpreview.github.io/?https://github.com/ooezenoglu/XPlab-2020-SR-Compatibility/blob/master/analysis/02_main/02-SR-compatibility-exploratory-analysis.html#data-wrangling

References

- Buerkner, Paul-Christian** (2016). “brms: An R package for Bayesian multilevel models using Stan”. In: *Journal of Statistical Software* 80.1, pp. 1–28.
- Carpenter, Bob et al.** (2016). “Stan: A probabilistic programming language”. In: *Journal of Statistical Software* 20, pp. 1–37.
- Dehaene, S., Bossini, S., & Giraux, P.** (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General*, 122, pp. 371–396. DOI: <https://doi.org/10.1037/0096-3445.122.3.371>
- Franke, M., & Roettger, T. B.** (2019, July 13). Bayesian regression modeling (for factorial designs): A tutorial. DOI: <https://doi.org/10.31234/osf.io/cdxv3>
- Henik, A., & Tzelgov, J.** (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, 10, pp. 389–395. DOI: <https://doi.org/10.3758/BF03202431>
- Fischer, M. H.** (2008). Finger counting habits modulate spatial-numerical associations. *Cortex*, 44, 386–392. DOI: <https://doi.org/10.1016/j.cortex.2007.08.004>
- Besner, D., & Coltheart, M.** (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, 17, 467–472. DOI: [https://doi.org/10.1016/0028-3932\(79\)90053-8](https://doi.org/10.1016/0028-3932(79)90053-8)
- Tzelgov, J., Meyer, J., & Henik, A.** (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18, 166–179. DOI: <https://doi.org/10.1037/0278-7393.18.1.166>
- Ren, P., Nicholls, M. R., Ma, Y., & Chen, L.** (2011). Size matters: Non-numerical magnitude affects the spatial coding of response. *Plos ONE*, 6. DOI: <https://doi.org/10.1371/journal.pone.0023553>
- Cohen Kadosh, R., Cohen Kadosh, K., Linden, D. J., Gevers, W., Berger, A., & Henik, A.** (2007). The brain locus of interaction between number and size: A combined functional magnetic resonance imaging and event-related potential study. *Journal of Cognitive Neuroscience*, 19, 957–970. DOI: <https://doi.org/10.1162/jocn.2007.19.6.957>
- Bueti, D., & Walsh, V.** (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society B*, 364, 1831–1840. DOI: <https://doi.org/10.1098/rstb.2009.0028>
- Kaufmann, L., Vogel, S. E., Wood, G., Kremser, C., Schocke, M., Zimmerhackl, L., & Koten, J. W.** (2008). A developmental fMRI study of nonsymbolic numerical and spatial processing. *Cortex*, 44, 376–385. DOI: <https://doi.org/10.1016/j.cortex.2007.08.003>
- Walsh, V.** (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends In Cognitive Sciences*, 7, pp. 483–488. DOI: <https://doi.org/10.1016/j.tics.2003.09.002>
- Wühr, P., & Seegelke, C.** (2018). Compatibility between Physical Stimulus Size and Left-right Responses: Small is Left and Large is Right. *Journal of Cognition*, 1(1): 17, pp. 1–11. DOI: <http://doi.org/10.5334/joc.19>