

Conceptual Replication of “Compatibility between Physical Stimulus Size and Left-right
Responses: Small is Left and Large is Right”

Alexander Prochnow, Paul Jänsch, Pia Stermann and Ludger Masch

University of Osnabrück

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Alexander Prochnow, Paul Jänsch, Pia Stermann and Ludger Masch: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Alexander Prochnow, Paul Jänsch, Pia Stermann and Ludger Masch, Postal address. E-mail: my@email.com

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: ATOM; compatibility; stimulus size; response location

Word count: X

Conceptual Replication of “Compatibility between Physical Stimulus Size and Left-right Responses: Small is Left and Large is Right”

ATOM (“A theory of magnitude”) is a theory proposed by Walsh(2003), which basically assumes a generalized magnitude-processing system in the brain where relations between the cortical representations of “time, space and quantity” exist. The theory predicts that if there were such relations then there must be interference and congruency effects in experimental research investigating stimulus-response tendencies comprising different numerical, temporal or spatial properties. Many studies have investigated different aspects and proposals of this theory in terms of Stimulus-response (S-R) compatibility, where S-R compatibility refers to how “natural” it is for a certain required response (or action) to follow a given stimulus type. Typically, such compatibility effects are investigated by comparing reaction times for a required action or response in experimental tasks with a specifically defined Stimulus-Response mapping that is varied within the experiment. For example Dehaene et al.(1990) found compatibility effects between numerical size and horizontal response location, which Dehaene et al.(1993) later coined as “Spatial-Numerical Association of Response Codes” (SNARC) effect. The SNARC effect is interpreted as a spatial mapping of numbers to horizontal locations. In 2013, Ren et al. have furthermore analogously investigated the relationship between physical stimulus size and horizontal response location. Their results show similar compatibility effects of stimulus size and horizontal response location, but only for right hand responses. In 2018, Wühr and Seegelke attempted a conceptual replication and extension of this study. They claim that their results indicate the existence of a general magnitude code, as proposed in ATOM in that we might intrinsically associate small objects with the left side and large objects with the right. In the experiment, subjects are presented with “small” or “large” stimulus objects (squares) and are required to judge their size as “small” or “large”. The keys for subjects to communicate these judgements are mapped to either left- or right-hand responses and these mappings are switched once within the experiment. Wühr and

Seegelke(2018) found, that participants were faster to respond to a large stimulus object with a key that is on the right of the keyboard (e.g. “Backspace”, pressed with the right hand), than with a key that is on the left of the keyboard (e.g. “Tabulator”, pressed with the left hand). For small stimulus objects, they found that reaction times were numerically faster for left-hand responses than for right-hand responses. For participants to judge whether the stimulus is a small or large object, ATOM would propose that subjects need to mentally activate an internal representation of size and the corresponding spatial response mapping in order to generate an action as response according to their judgement about stimulus size, and that these representations are linked. Ren et al. (2011), and Wühr and Seegelke (2018) showed that stimulus size and horizontal response location indeed exhibit compatibility effects, which can be regarded as evidence for this sub-aspect of ATOM. They found that the horizontal response mapping (key press with either left or right hand) to judgements about stimuli of different size influences response behavior and matters for success and swiftness of judgements of “small” or “large” stimulus objects. We here aimed to assess the findings of Wühr and Seegelke(2018) via a conceptual replication of the first experiment and extended it by not only taking right-handed participants into account but also seperately testing left-handed participants. Our hypotheses were based on their results that small stimuli are associated with left-hand responses, whereas large stimuli are associated with right-hand responses. We sought to replicate similar stimulus size - response location compatibility effects as those reported by Wühr and Seegelke(2018), that means we wanted to find out whether different stimulus-response mapping conditions have an effect on participant’s reaction time to a specific type of stimulus (in this case large or small square). Additionally we aimed to investigate whether we can find compatibility effects in opposite direction for left-handed participants. However we changed In particular we tested for right-handed participants if response times for right-hand responses are faster to the larger stimulus than to the smaller stimulus and similarly if response times for left-hand responses are faster to the smaller stimulus than to the larger stimulus.

Furthermore we investigated if the stimulus size – response location compatibility effect is larger for right-hand responses than for left-hand responses. This would mean that the difference in response times for right-hand responses to be faster to the larger stimulus than to the smaller stimulus is larger than the difference in response times for left-hand responses to be faster to the smaller stimulus than to the larger stimulus. Correspondingly we tested the former hypotheses in the opposite direction for left-handed participants. In other words we assumed that the response times for left-hand responses are faster to the larger stimulus than to the smaller stimulus as well as that the response times for right-hand responses are faster to the smaller stimulus than to the larger stimulus. Analogously we hypothesized that here the stimulus size - response location compatibility effect will be larger for the left-hand responses instead of right-hand responses. In the end we compared the size of possible compatibility effects regarding people’s handedness and assumed that the stimulus size- response location compatibility effect is larger for right-handed people than for left-handed people. I.e., the absolute difference between response times in the compatible condition and response times in the incompatible condition is larger for right-handed than for left-handed people.

Methods

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study.

Participants

We recruited $N = 72$ participants (43 woman, 27 men, 2 others, $22.60 = \text{mean}$, age range: 14-55 years) by distributing the link to the experiment via social media, mainly WhatsApp, Facebook and Reddit (see Table 1 for demographics). The link was hosted on Netlify. 59 of the participants were right-handed and 12 were left-handed, as well as 1 participant indicating mixed-handedness. They were advised to take part only once,

however, we did not control for this. On average, participants needed 12.89 minutes to complete the experiment, with a minimum time of 9.84 minutes and a maximum time of 21.07 minutes after filtering (explained in the data preparation section). The link was available for four days until we closed collection of data (from the 31st of July 2020 until the 3rd of August 2020). Before taking part, the participants declared to have normal or corrected-to-normal visual acuity and should be naive with respect to the purpose of the study. Participation was voluntary and not rewarded with money or material things but with our deep gratitude.

Material

In each trial, a visual stimulus object was presented in the center of the screen with white background. A stimulus object in the experimental block was a simple black square, which was either small (about 3 x 3 cm, depending on the size of the screen) or large (edges twice as long as those of the small stimulus, i.e. about 6 x 6 cm, depending on the size of the screen). We used information about the stimulus objects (i.e. information about how they are to be displayed concerning their form, size and position on the display) as provided by Wühr and Seegelke (2018), however we could not guarantee an exact same formatting of stimuli for each participant in terms of equal edge lengths, but only in terms of relations between edge lengths, since the size of the displayed stimulus was relative to the screen on which it was presented. The 2 different types of stimulus objects were used for training trials and main trials in both experimental blocks. Additionally to the original study we included an intermediate block which was intended as a distraction. A stimulus object in this block was either a simple black circle or a simple black triangle, both of small to intermediate size - regarding the small and large square stimulus objects' edge lengths - in terms of the circle's diameter or the triangle's height.

Procedure

The experiment consisted of 10 parts. These parts comprised the two main experimental blocks, one for each of the two S-R mapping conditions, and one intermediate block with the distraction tasks. We varied the stimulus size - response (key) location mapping (S-R mapping) for the experimental blocks once over the course of the experiment. The order of S-R mapping conditions for the first and second experimental blocks (first = compatible – second = incompatible or first = incompatible – second = compatible) was determined experiment initially, uniformly at random once for each participant. In the compatible mapping condition, the small stimulus required a response with the left (“q”) key, whereas the large stimulus required a response with the right (“p”) key. The mapping was reversed in the incompatible mapping condition. At the beginning of the experiment, participants were welcomed and given (not further specified) information about the general procedure, namely, that the experiment will consist of several parts, which again consist of several trials of tasks of similar structure. It was recommended to the participants to complete the whole experiment (to prevent dropouts from the beginning on), and they were given information about the approximate duration of the experiment (about 15 minutes). More precise descriptions of the participant’s task were presented at the beginning of each block. These specific instructions included information about the sequence of events in a trial (fixation - stimulus presentation - key press or button click given the assigned S-R mapping), i.e. which keys to press for “small” or “large” stimuli (corresponding to the respective S-R mapping condition) in the experimental task blocks, or which buttons to click on for “circle” or “triangle” in the distracting task block, as well as information on whether the next part belongs to training (for the experimental blocks) or experimental phase (as in the intermediate block). Additionally, all instructions emphasized that participants should strive to optimize speed and accuracy (this was not mentioned in the original paper but seemed appropriate in this context, since a time-out occurred and an error message was displayed after the key press

when participants' response time(RT) exceeded a period of 2000 ms). After having read the initial instructions of a new experimental block, participants firstly performed a training trial phase (10 trials) before they proceeded with the main experimental trial phase for the respective S-R mapping condition($2 \text{ stimuli} \times 30 \text{ repetitions} = 60 \text{ trials}$). The first experimental block was followed by written instructions about the intermediate task, in which participants were shown either a circle or a triangle and had to classify the presented stimulus as "circle" or "triangle" by clicking on the respective button at the bottom of the screen. The intermediate task had ($2 \text{ stimuli} \times 10 \text{ repetitions} =$) 20 main trials. The specific number of the intermediate task's trials was not necessarily optimal, but we decided on this as a -hopefully well - balanced compromise between effective distraction and no unnecessary prolonging of the experiment. The order in which participants were shown either of the two stimuli in the intermediate task was determined uniformly at random. After completion of the intermediate block's trials, participants were shown written instructions of the following experimental block (similar to the initial experimental block's instructions, only different in the description of the S-R mapping, since it was reversed here). They then performed a training block (10 trials) and an experimental block with the respective second S-R mapping ($2 \text{ stimuli} \times 30 \text{ repetitions}$). During the training phases participants accustomed to the task, specifically to the respective S-R mapping, by completing the given amount of practice trials. Practice trials differed from main trials in that participants received a short feedback message, telling them whether their answer was correct (namely: "Correct!" or "Incorrect!"). The order in which individual trials are presented in the training and experimental phases is determined uniformly at random and on the fly for each participant. All experimental trials were structured as follows. Each trial started with the presentation of a fixation point (a small black "x" sign) for 1000 ms in the center of the screen. Afterwards one of the two stimulus objects were displayed until a keypress (experimental blocks) or button click (intermediate block) occurred. In the experimental blocks, participants responded by pressing the left

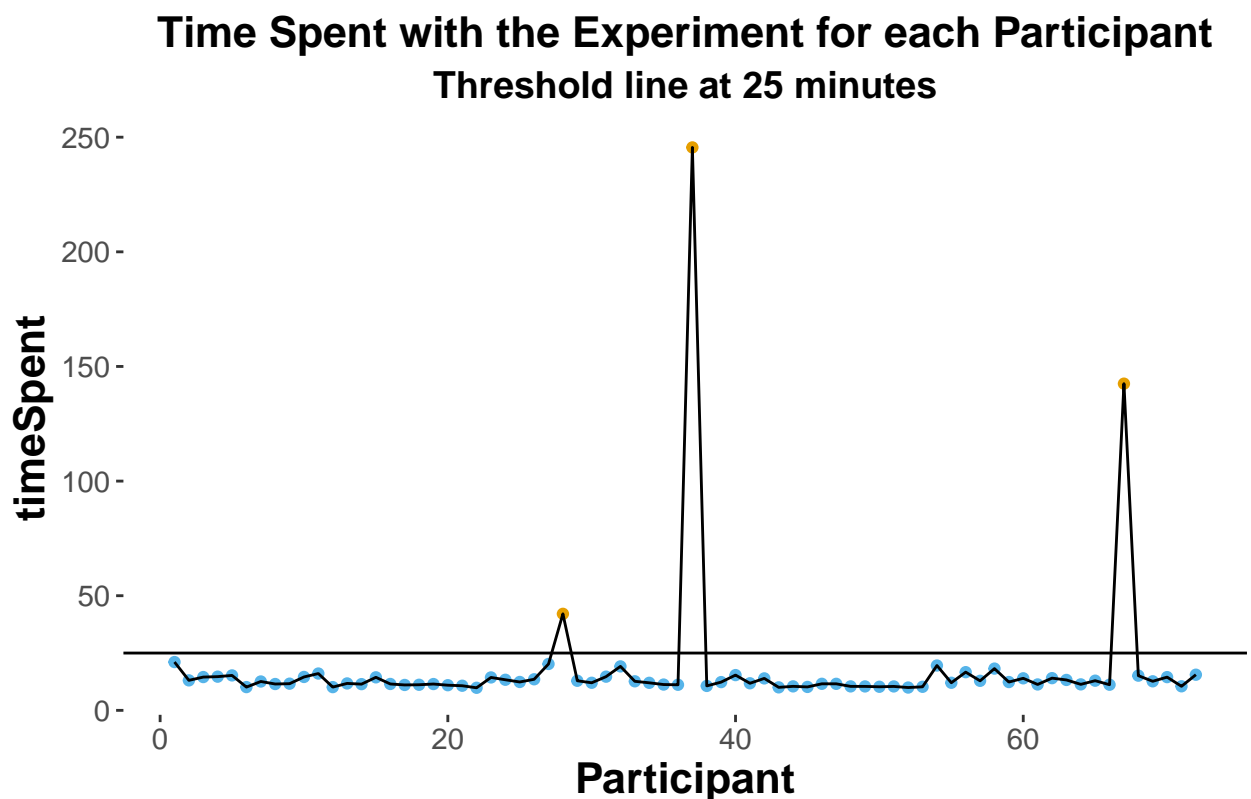
“q” key or the right “p” key on their keyboard. Participants should operate the left key (“q”) with the index finger of their left hand and the right key (“p”) with the index finger of their right hand. In the intermediate task, participants responded by clicking on the “circle” or “triangle” button at the bottom the screen. A correct response with a response time below 2000 ms after stimulus presentation was followed by a blank screen for 1500 ms. If the “q” or the “p” key was pressed after the maximum response period of 2000 ms, a corresponding error message (precisely: “Try to respond more quickly!”) was shown for 1500 ms in black color. In the intermediate block however each of the trials had to be initiated by clicking on a “start” button in the middle of the screen. In the main experimental blocks the next trial started automatically after the “p” or “q” key was pressed. Finally, the experiment terminated with a post-experiment survey asking participants for their preferred handedness (participants could choose between “right-handed” (default option), “left-handed” and “no preference”) and to optionally supply socio-demographic information (age, gender, level of education, native languages, whether there were any technical issues experienced during the experiment). Participants were able to leave comments on the experiment by entering them in a text field.

Data Preparation

First, we applied basic exclusions, namely only using experimental trials (not training or distractor trials) and only using trials where participants responded correctly to the stimulus by pressing the correct key. Second, we filtered out participants that responded faster than 100ms or slower than 1500ms. Third, we log transformed the reaction times for use in our Bayesian regression models. The following part will describe two data exclusions which we did not state in our preregistration report that filtered out 3 and 1 participant, respectively. We aim to be transparent with the following two decisions and hope our rationale for them is understandable despite them being made after having seen the data.

Originally, our data contained 3 participants who needed drastically longer for the

experiment than the average participant. We filtered out these 3 participants by choosing a threshold of 25 minutes, which these participants were well above. Figure X shows our rationale for choosing the threshold, as these three participants likely took a break during the experiment or performed another activity, which was not intended by our experimental design. This additional data exclusion criterion was not stated in our preregistration report, because we had not foreseen this unintended effect of conducting an online experiment, however the longer break in between trial blocks might cause those participants to perform differently in the following trials. To be on the safe side we therefore excluded these participant's data from our analyses.

**Figure X**

Additionally, we filtered out one mixed-handed participant, since their data does not fit into our hypotheses, however we wanted to include the option in our post-experiment survey to not classify data incorrectly. We made this decision after submitting the preregistration report, but think this exclusion follows naturally from our hypotheses, since

they only specify to use data from either left- or right-handed participants or both.

Data analysis

We used R (Version 3.6.1; R Core Team, 2020) and the R-packages *aida* (Version 0.1.0; person, 2020), *brms* (Version 2.12.0; Bürkner, 2017, 2018), *broom* (Version 0.5.6; Robinson & Hayes, 2020), *dplyr* (Version 0.8.5; Wickham et al., 2020), *forcats* (Version 0.5.0; Wickham, 2020), *ggplot2* (Version 3.3.0; Wickham, 2016), *knitr* (Version 1.28; Xie, 2015), *papaja* (Version 0.1.0.9997; Aust & Barth, 2020), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *Rcpp* (Version 1.0.4; Eddelbuettel & François, 2011; Eddelbuettel & Balamuta, 2017), *readr* (Version 1.3.1; Wickham, Hester, & François, 2018), *stringr* (Version 1.4.0; Wickham, 2019), *tibble* (Version 3.0.0; Müller & Wickham, 2020), *tidyr* (Version 1.0.2; Wickham & Henry, 2020), and *tidyverse* (Version 1.3.0; Wickham, Averick, et al., 2019) for all our analyses.

We analyzed our data with regard to the seven hypotheses we defined previously. To assess hypotheses 1-6 we defined two identical Bayesian regression models. They analyzed the influence of the factors S-R Mapping and Response separately on the (log-)Reaction Time as well as the interaction effect of S-R Mapping and Response on (log-)Reaction Time. We additionally accounted for group level variation between participants by adding random effects, namely random intercepts and slopes for the factor Response. We used the R package “brms”. Our first two models (model_right and model_left) were defined by the formula:

$$\log RT \sim \text{Mapping} * \text{Response} + (\text{Response} || \text{Participant})$$

For both models, the data is assumed to be Student-t-distributed and we did not replace the flat priors. Every other parameter of the models remained on the default setting. The models ran on the preprocessed data as described in section 22. data exclusion, however model_right was only fitted with data from right-handed participants,

while `model_left` only used data from left-handed participants for fitting. The exact analysis can be found in the attached script. To assess hypothesis 7 we defined an additional model (`model_both`) which only slightly differed from the two previous models and was defined by the formula:

$$\text{logRT} \sim \text{Mapping} * \text{Handedness} + (1 \parallel \text{Participant})$$

The handedness of the participants replaced the response location as a factor influencing (log-)Reaction time. Random effects now only comprised a random intercept. This model used as dataset containing the exact same number of left- and righthanded participants. The family again was set to Student-t and the priors were left on the default setting (flat priors). The independent, manipulated variable was the S-R Mapping with levels compatible and incompatible, the further independent variable Response with levels left and right. These factors were encoded using the default dummy/treatment coding with the group level combination “compatible” and “left” as the reference level. For Model 3, a further independent variable was the Handedness, dummy coded with reference level “righthanded”. The dependent variable was the (log-)Reaction Time.

Results

Discussion

References

- Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, 10(1), 395–411. <https://doi.org/10.32614/RJ-2018-017>
- Eddelbuettel, D., & Balamuta, J. J. (2017). Extending extitR with extitC++: A Brief Introduction to extitRcpp. *PeerJ Preprints*, 5, e3188v1. <https://doi.org/10.7287/peerj.preprints.3188v1>
- Eddelbuettel, D., & François, R. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8), 1–18. <https://doi.org/10.18637/jss.v040.i08>
- Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from <https://CRAN.R-project.org/package=purrr>
- Müller, K., & Wickham, H. (2020). *Tibble: Simple data frames*. Retrieved from <https://CRAN.R-project.org/package=tibble>
- person. (2020).
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Robinson, D., & Hayes, A. (2020). *Broom: Convert statistical analysis objects into tidy tibbles*. Retrieved from <https://CRAN.R-project.org/package=broom>
- Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from <https://ggplot2.tidyverse.org>

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*.

Retrieved from <https://CRAN.R-project.org/package=stringr>

Wickham, H. (2020). *Forcats: Tools for working with categorical variables (factors)*.

Retrieved from <https://CRAN.R-project.org/package=forcats>

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . .

Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>

Wickham, H., François, R., Henry, L., & Müller, K. (2020). *Dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>

Wickham, H., & Henry, L. (2020). *Tidyr: Tidy messy data*. Retrieved from

<https://CRAN.R-project.org/package=tidyr>

Wickham, H., Hester, J., & Francois, R. (2018). *Readr: Read rectangular text data*.

Retrieved from <https://CRAN.R-project.org/package=readr>

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, Florida:

Chapman; Hall/CRC. Retrieved from <https://yihui.org/knitr/>