

Exploring the origin of the number-size congruency effect: Sensitivity or response bias?

Dennis Reike¹ · Wolf Schwarz¹

Published online: 27 January 2017
© The Psychonomic Society, Inc. 2017

Abstract Physical size modulates the efficiency of digit comparison, depending on whether the relation of numerical magnitude and physical size is congruent or incongruent (Besner & Coltheart, *Neuropsychologia*, 17, 467–472, 1979), the number-size congruency effect (NSCE). In addition, Henik and Tzelgov (*Memory & Cognition*, 10, 389–395, 1982) first reported an NSCE for the reverse task of comparing the physical size of digits such that the numerical magnitude of digits modulated the time required to compare their physical sizes. Does the NSCE in physical comparisons simply reflect a number-mediated bias mechanism related to making decisions and selecting responses about the digit's sizes? Alternatively, or in addition, the NSCE might indicate a true increase in the ability to discriminate small and large font sizes when these sizes are congruent with the digit's symbolic numerical meaning, over and above response bias effects. We present a new research design that permits us to apply signal detection theory to a task that required observers to judge the physical size of digits. Our results clearly demonstrate that the NSCE cannot be reduced to mere response bias effects, and that genuine sensitivity gains for congruent number-size pairings contribute to the NSCE.

Keywords Numerical cognition · Number-size congruity effect · Signal detection theory

✉ Dennis Reike
reike@uni-potsdam.de

Wolf Schwarz
wschwarz@uni-potsdam.de

¹ Department of Psychology, University of Potsdam, P.O. Box 60 15 53, 14415 Potsdam, Germany

Numbers and numerical information play a central role in how humans represent, communicate, and respond to quantity-related aspects of their environments. Correspondingly, a considerable amount of research has addressed many cognitive, neuronal, developmental, and clinical aspects of how we process quantitative information (for general review, see Dehaene, 2011; Nieder, 2005).

According to one influential model of number comparison (Moyer & Landauer, 1967; for recent formulations and review, see Ditz & Nieder, 2016; Ganor-Stern & Goldman, 2015; Gilmore, Attridge, & Inglis, 2011; Inglis & Gilmore, 2014; Maloney, Risko, Presto, Ansari, & Fugelsang, 2010; Nuerk, Moeller, Klein, Willmes, & Fischer, 2011; Reike & Schwarz, 2016; Sigman & Dehaene, 2005), digits are automatically converted into percept-like analog representations that are then in turn compared with each other, much like sensory representations of physical attributes such as brightness or orientation. One line of evidence consistent with this model derives from conflict paradigms in which the to-be-compared digits are presented with varying physical (i.e., font) sizes. According to the analog representation model, the digit's (task-irrelevant) physical size should modulate the efficiency of the comparison process, depending on whether the relation of numerical magnitude and physical size is congruent (as in 8 – 2) or incongruent (e.g., 8 – 2). This number-size congruency effect (NSCE) has indeed first been reported by Besner and Coltheart (1979), and has since then often been confirmed and extended (e.g., Algom, Dekel & Pansky, 1996; Cohen Kadosh et al., 2007; Fitoussi, 2014; Girelli, Lucangeli, & Butterworth, 2000; Henik & Tzelgov, 1982; Kaufmann et al., 2005; Pansky & Algom, 1999; Schwarz & Heinze, 1998; Schwarz & Ischebeck, 2003; Szűcs & Soltész, 2007; Takahashi & Green, 1983; Tzelgov, Meyer, & Henik, 1992). Henik and Tzelgov (1982) first studied the reverse task of comparing the physical size

of digits, independent of their numerical meaning. Contrary to strictly serial models in which the physical features of the digit (including its size) are fully identified before the digit's numerical magnitude is accessed, they observed reliable NSCEs such that the (logically task-irrelevant) numerical value of the digits modulated the time required to compare their physical sizes (for recent related studies, see Arend & Henik, 2015; Cantlon, Platt, & Branno, 2009; Faulkenberry, Cruise, Lavro, & Shaki, 2016; Gabay, Leibovich, Henik, & Gronau, 2013; Goldfarb & Treisman, 2010; Pina, Castillo, Cohen Kadosh, & Fuentes, 2015; Reber, Christensen, & Meier, 2014; Risko, Maloney, & Fugelsang, 2013; Santens & Verguts, 2011; Sobel, Puri, & Faulkenberry, 2016).

An as-yet-open question, addressed in the present experiment, is that the NSCE in physical comparisons might simply reflect a number-mediated bias mechanism related to making decisions and selecting responses about the digit's sizes (e.g., Risko et al., 2013; Schwarz & Heinze, 1998; Sobel et al., 2016). Expressed in the terminology of signal detection theory (SDT; Macmillan & Creelman, 2005), one interpretation of the NSCE is that the digit's numerical magnitude may induce a number-related response bias because observers implicitly tend to associate numerical magnitude with physical size. In this view, congruent pairings of physical size and numerical magnitude do not actually enhance the sensitivity of the size discrimination process but more simply benefit from a tendency to respond more readily "larger" ("smaller") with numerically large (small) digits. Alternatively, the NSCE might more adequately be interpreted as reflecting a genuine sensitivity benefit tied to the numerical meaning of the digits presented (for example, Sobel, Puri, Faulkenberry, & Dague, 2016, Exp. 3 report that numerical information can influence but probably not guide visual search). In this interpretation, the NSCE reflects a true increase in the ability to discriminate small and large font sizes when these sizes are congruent with the digit's symbolic numerical meaning—over and above mere response tendencies.

With respect to these alternatives, it is unfortunate that most extant research designs exploring the NSCE are based on the analysis of response time. Chronometric measures often help to investigate the onset and time course of the NSCE but they also tend to make it difficult to distinguish between bias and sensitivity accounts (e.g., Luce, 1986, chapters 3–4; Matt & Dzharfarov, 2014). Therefore, in the present study we developed a new research design to apply SDT to a task in which observers judged the physical size of digits. Specifically, single digits were presented in either of two global sizes (sets S (mall) and L (arge)); *between* these sets digits differed massively in size. In addition, *within* each size set digit exemplars could have a slightly smaller ($S-$, $L-$) or larger ($S+$, $L+$) font size (for details, see Methods and Fig. 1), and the task of the observer was

to classify the digit shown as being a smaller or a larger exemplar within its size set. In this design, the difference between "global" font size (S vs. L) is large (and was made even more obvious by using different colors to indicate the different-sized sets) but not task-relevant, whereas the task-relevant feature (the "local" size within each set) requires a more difficult discrimination of minute size differences.

The aim of the present study was to analyze results from this size identification design within the framework of SDT. In particular, we aimed to extract, and to compare, separate bias and sensitivity measures for congruent and incongruent pairings of numerical magnitude and (global) physical size to further clarify the origin and nature of the NSCE.

Methods

Participants Thirty-six University of Potsdam students, aged 18–39, with normal or corrected-to-normal vision participated in one session of approximately 1 h. They received a payment of 8 euros or course credit for their participation.

Stimuli and apparatus The stimuli consisted of the digits 1–9, excluding 5. Each trial display consisted of a single digit presented against a gray background in Verdana font on a 75-Hz, 1028 × 768 px VGA color monitor. As shown in Fig. 1, we divided the stimuli into two sets: a small size set (S), and a large size set (L). The mean font size for the small size set was 30 pixels (px; 1 deg = 65 px at the viewing distance of approximately 114 cm); for the large size set it was 52.5 px. Therefore, the mean size distance of the fonts of these sets was 22.5 px. Within both size sets the digits were presented in two slightly different physical sizes: smaller (–) and larger (+) exemplars. Within the small size set the digits were presented in a font size of 28 px ($S-$) or 32 px ($S+$), whereas within the large size set they were

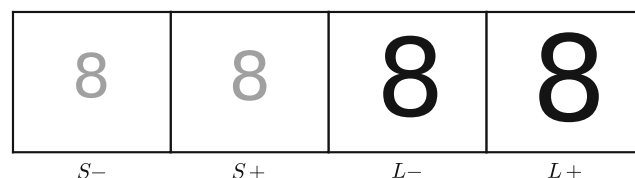


Fig. 1 Sizes of digit stimuli as used in the present experiment. For example, the digit 8 was shown in four different physical sizes. The *two left panels* illustrate the small size set (S); the *two right panels* illustrate the large size set (L). Both size sets consist of two slightly different sizes: smaller (–) and larger (+) exemplars. Digits from one size set are presented in *dark blue* and from the other size set in *brown*. The *two left digits* ($S-$, $S+$) represent incongruent conditions because the global physical size (small, S) does not correspond to the numerical magnitude (large, >5). Similarly, the *two right digits* ($L-$, $L+$) represent congruent conditions because the global physical size (large, L) corresponds to the numerical magnitude (large, >5)

presented with font sizes of 50 px ($L-$) or 55 px ($L+$), making the within-set size differences (S : 4 px; L : 5 px) much smaller than the mean between-set difference (22.5 px). In addition, the digits from one size set were presented in dark blue and from the other size set in brown (counterbalanced between subjects) in order to avoid any uncertainty about the size set (S or L) to which the digit presented belonged (as variations of the color-to-size mapping had no significant main or interaction effect, this factor is ignored in the sequel).

Procedure In each of ten blocks, all eight digits were presented twice in each of the four sizes in random order. Each trial started with the presentation of a fixation cross; after a delay of 500 ms the digit was presented until a response was given. The task was to indicate by button press whether the presented digit was a physically smaller ($-$) or a larger ($+$) exemplar within its size set. Following the response, the participants received visual feedback about the correct answer and the response they had given. A blank screen was presented after 1000 ms for 500 ms, before the next trial started. After each block, the overall error rate was computed for this block separately for each size set (S and L). In the first block, the error rate had to be between 15 and 30, in block two until block four between 10 per cent and 35 per cent, and after the fourth block between 5 per cent and 40 per cent. If the error rate of one size set did not meet these criteria, then the font size of the larger ($+$) exemplars of this set was changed one px in the required direction. For example, if the error rate for the digits in the small size set (S) in the first block was below 15 per cent, then the size of the larger exemplars of the small digits ($S+$) was reduced from 32 px to 31 px. Required size changes were applied after each block.

Preliminary data reduction Only blocks in which the error rates for both size sets were within the targeted accuracy range were included in the analyses. The first block was considered practice and was always excluded. According to the block inclusion rules described, on average 5.5 (SD = 1.4) blocks for each participant were included in the analysis; including all blocks, except block 1, produced the same pattern of in-/significant results. In the present design, a digit presented in any trial is characterized by three independent attributes: its membership in the physical size set L or S , its numerical magnitude (<5 vs. >5), and by being a smaller ($S-$, $L-$) or larger ($L+$, $S+$) exemplar within its size set. Only the last attribute is response relevant; the first two attributes are logically and statistically independent of each other, and of the correct response. Congruency was defined by the correspondence of the two response-irrelevant attributes physical size and numerical magnitude, so that digits <5 in the set S and digits >5 in the set L

were congruent, whereas digits >5 in the set S and digits <5 in the set L were incongruent. Thus, for each participant the sensitivity index d' and the response bias index $\ln \beta$ (Macmillan & Creelman, 2005) were computed separately for two \times two conditions, physically small or large (defined by set S vs. L) \times numerically small or large (defined by magnitude <5 vs. >5).

Results

Sensitivity index d' Values of d' for each participant were subjected to a 2-factorial ANOVA with the two within-subject factors physical size set (2: S/L) and numerical magnitude (2: $<5/>5$).

As intended, the required discrimination of sizes within each set was not an easy task, yielding an overall d' of 1.42 (SE 0.09) (corresponding to 74.4 % correct responses, relative to a chance level of 50 %).

Neither physical size set nor numerical magnitude had a main effect on d' . Central to our study, the interaction between physical size set and numerical magnitude was significant (Fig. 2, left panel), $F(1, 35) = 9.82$, $MSE = 0.15$, $p < .01$, $\eta^2 = .22$. Differences in d' were found within each size set separately: for physically small digits the difference between numerically small ($d' = 1.54$) and numerically large digits ($d' = 1.40$) was significant, $t(35) = -1.78$, $p < .05$, $\eta^2 = .18$. Similarly, for physically large digits, the d' difference between numerically small digits ($d' = 1.24$) and numerically large digits ($d' = 1.50$) was significant, $t(35) = 3.06$, $p < .01$, $\eta^2 = .43$. Thus, the congruency effect found did not depend on just one specific physical size set, as shown in Fig. 2.

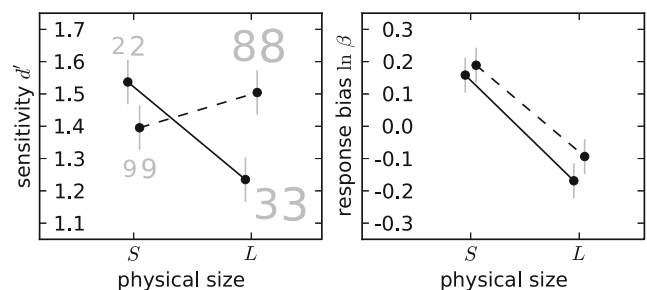


Fig. 2 Sensitivity index d' and response bias index $\ln \beta$ for each combination of physical size (set S vs. L) and numerical magnitude (<5 vs. >5). *Left panel:* Circles show the mean sensitivity index d' (ordinate) for numerically small (<5 ; solid line) and large (>5 ; dashed line) digits as a function of size set (S small / L large; abscissa). Each condition is illustrated by one smaller ($-$) and one larger ($+$) digit exemplar. *Right panel:* Circles show the mean response bias index $\ln \beta$ (ordinate) for numerically small (solid line) and large (dashed line) digits as a function of size set (abscissa). Error bars represent ± 1 SE (Loftus & Masson, 1994)

In order to validate and further investigate these findings, we also considered the congruency effect (defined as the d' difference between congruent and incongruent conditions) for each participant separately. On average, d' was 0.21 (SE 0.07) larger for congruent than for incongruent conditions; the participants answered correctly in 76.1 % of all congruent, and in 72.8 % of all incongruent trials. Overall, 25 of 36 participants (69 %) showed a positive difference between the d' s of congruent vs. incongruent conditions, with the mean positive differences (0.40) being larger than the mean negative differences (-0.23), $W+ = 515$, $W- = 151$, $p < .01$, $\eta^2 = .11$, Wilcoxon signed-rank test.

Response bias index $\ln \beta$ Values of $\ln \beta$ (positive values indicating a tendency to respond “smaller”) for each participant were subjected to a repeated-measures ANOVA with the same factors as used for d' .

As expected under our balanced design, the grand mean of $\ln \beta$ was close to zero (0.02; SE 0.09), indicating that our participants had no overall bias towards using one of the two responses. Only the factor size set exerted a systematic main effect on $\ln \beta$. Participants tended to classify the digits as being smaller ($-$) if they belonged to the small size set (S), and as being larger ($+$) if they belonged to the large size set (L), $F(1, 35) = 7.83$, $MSE = .10$, $p < .01$, $\eta^2 = .18$. As shown in Fig. 2 (right panel), the mean response bias index was $+0.17$ for digits belonging to size set S , and -0.13 for digits belonging to size set L . Numerical magnitude had no main effect on $\ln \beta$. Finally, differences in $\ln \beta$ between congruent and incongruent trials were quite small in both size sets, as reflected in the nonsignificant interaction of physical size and numerical magnitude, $F(1, 35) = .18$, $MSE = .10$, $p = .67$, $\eta^2 = .01$.

Discussion

As reviewed in the Introduction, judgments about the physical size of digits are faster and less error-prone when physical size and numerical magnitude are congruent rather than incongruent (Henik & Tzelgov, 1982; for recent review, see Arend & Henik, 2015; Fitousi, 2014). Does this well-established performance benefit reflect a genuine enhancement of sensitivity, that is, an increase in the ability to discriminate small and large font sizes when numerical magnitude and physical size are congruent? Or is it more adequately attributed to a number-mediated response bias mechanism, for example, a tendency to respond more readily “larger” with numerically large, and more readily “smaller” with numerically small digits?

The present results strongly suggest that, in a general sense, response biases indeed systematically influence the

way in which observers judge the physical size of digits. Specifically, as shown in Fig. 2 (right panel), digits in the size set S are more readily classified as “smaller”, and digits in the size set L more readily as “larger”, even though smaller ($S-$, $L-$) and larger ($S+$, $L+$) exemplars were equally frequent in both size sets. If such biases selectively favor congruent number-size pairings and if “more readily” in the SDT sense translates into faster responses (e.g., Luce, 1986, ch. 7), then the standard finding of shorter response times for congruent number-size pairings could be attributed to similar response bias mechanisms.

However, our results clearly demonstrate that the NSCE cannot simply be reduced to bias effects, and that genuine sensitivity gains for congruent number-size pairings contribute to the NSCE over and above mere response tendencies.¹ Specifically, as shown in Fig. 2 (left panel), $S-$ vs. $S+$ digits in the size set S are discriminated with higher sensitivity when they are numerically small than when they are numerically large. Conversely, $L-$ vs. $L+$ digits in the size set L are discriminated with higher sensitivity when they are numerically large than when they are numerically small. Note that in our design within either size set S and L separately, each numerical magnitude was presented as a smaller and larger exemplar equally often; therefore, the differential sensitivity effects obtained cannot be attributed to overall response biases. Consider, for example, the simple biased response strategy of classifying numerically small digits more readily as “smaller”, and numerically large digits more readily as “larger”. In our design, this biased strategy would not produce the differential sensitivity effect shown in the left panel of Fig. 2 because numerically small and large digits were presented equally often as small and large exemplars in both size sets. Also, our results provide no evidence for this specific form of number-related response bias (Fig. 2, right panel). As discussed above, a related type of simple biased response strategy is to classify digits in size set S more readily as “smaller”, and digits in size set L more readily as “larger”. Again, this biased strategy (which is more prominent in our data; see Fig. 2, right panel) would not produce the differential sensitivity effect observed because smaller and larger exemplars were presented equally often within both size sets. The same conclusion applies to any biased response strategy based on some form of combination of numerical magnitude (<5 vs. >5) and physical size (S vs. L).

The present results suggest that the to-be-judged physical sizes of the digits ($S-$ vs. $S+$, and $L-$ vs. $L+$)

¹Note that in the context of signal detection theory partitioning the total congruency effect into an inhibitory and a facilitatory component (as in RT studies involving a neutral condition; Henik & Tzelgov, 1982) requires a more elaborate SDT design (Macmillan & Creelman, 2005, ch. 5).

are internally represented in a noisy format that is modifiable by numerical magnitude. This claim is consistent with the general theory of magnitude (ATOM) positing that there is a common cortical metric underlying several quantity-related attributes, such as space, time, and number (e.g., Bonato, Zorzi, & Umiltà, 2012; Cohen Kadosh, Lammertyn, & Izard, 2008; Eiselt & Nieder, 2013; Henik, Leibovich, Naparstek, Diesendruck, & Rubinsten, 2012; Leon & Shadlen, 2003; Schwarz & Eiselt, 2009; Walsh, 2003; Winter, Marghetis, & Matlock, 2015). For example, in their coalescence diffusion model Schwarz and Ischebeck (2003) proposed that information from task-irrelevant attributes (e.g., numerical magnitude, and physical font size S vs. L) often cannot be completely ignored. In their model, task-irrelevant information effectively combines with (and thereby modifies) information from task-relevant stimulus attributes to form an amalgam representation reflecting both, relevant and irrelevant stimulus aspects. On the face of it, our findings could be seen as being more compatible with an interaction at an early representational rather than at a late decision stage (e.g., Risko et al., 2013; Schwarz & Heinze, 1998; Sobel et al., 2016; Szűcs & Soltész, 2007); however, it should be stressed that SDT models per se are mute with respect to chronometric aspects. More generally, our findings clearly fit in with, and further extend, previous results from a variety of perceptual tasks (e.g., Casarotti, Michielin, Zorzi, & Umiltà, 2007; Corbett, Oriet, & Rensink, 2006; Fischer, Castel, Dodd, & Pratt, 2003; Godwin, Hout, & Menneer, 2014; Nieder, 2005; Schwarz & Eiselt, 2009, 2012; Sobel et al., 2016) suggesting that symbolic numerical meaning is extracted at an early processing stage from visual displays containing digits, and under favorable (i.e., congruent) conditions, may enhance perceptual sensitivity, even in basic psychophysical tasks involving judgments about physical size.

Author Note The present work was supported by a research grant (SCHW 611/3-1) from the Deutsche Forschungsgemeinschaft (DFG).

References

- Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory & Cognition*, *24*, 557–572.
- Arend, I., & Henik, A. (2015). Choosing the larger versus choosing the smaller: Asymmetries in the size congruity effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*, 1821–1830.
- Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, *17*, 467–472.
- Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience & Biobehavioral Reviews*, *36*, 2257–2273.
- Cantlon, J.F., Platt, M.L., & Branno, E.M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*, 83–91.
- Casarotti, M., Michielin, M., Zorzi, M., & Umiltà, C. (2007). Temporal order judgment reveals how number magnitude affects visuospatial attention. *Cognition*, *102*, 101–117.
- Cohen Kadosh, R., Cohen Kadosh, K., Linden, D.E.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.
- Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Progress in Neurobiology*, *84*, 132–147.
- Corbett, J.E., Oriet, C., & Rensink, R.A. (2006). The rapid extraction of numeric meaning. *Vision Research*, *46*, 1559–1573.
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics, revised and updated edition*. New York, NY: Oxford University Press.
- Ditz, H.M., & Nieder, A. (2016). Numerosity representations in crows obey the Weber–Fechner law. *Proceedings of Royal Society B*, *283*, 20160083.
- Eiselt, A.-K., & Nieder, A. (2013). Representation of abstract quantitative rules applied to spatial and numerical magnitudes in primate prefrontal cortex. *The Journal of Neuroscience*, *33*, 7526–7534.
- Faulkenberry, T.J., Cruise, A., Lavro, D., & Shaki, S. (2016). Response trajectories capture the continuous dynamics of the size congruity effect. *Acta Psychologica*, *163*, 114–123.
- Fischer, M.H., Castel, A.D., Dodd, M.D., & Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nature Neuroscience*, *6*, 555–556.
- Fitousi, D. (2014). On the internal representation of numerical magnitude and physical size. *Experimental Psychology*, *61*, 149–163.
- Gabay, S., Leibovich, T., Henik, A., & Gronau, N. (2013). Size before numbers: Conceptual size primes numerical value. *Cognition*, *129*, 18–23.
- Ganor-Stern, D., & Goldman, R. (2015). Reaching towards an end: Numerical end and distance effects in motor movements. *Journal of Cognitive Psychology*, *27*, 490–498.
- Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *The Quarterly Journal of Experimental Psychology*, *64*, 2099–2109.
- Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology*, *76*, 104–122.
- Godwin, H.J., Hout, M.C., & Menneer, T. (2014). Visual similarity is stronger than semantic similarity in guiding visual search for numbers. *Psychonomic Bulletin & Review*, *21*, 689–695.
- Goldfarb, L., & Treisman, A. (2010). Are some features easier to bind than others? The Congruency Effect. *Psychological Science*, *21*, 676–681.
- Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*, 389–395.
- Henik, A., Leibovich, T., Naparstek, S., Diesendruck, L., & Rubinsten, O. (2012). Quantities, amounts, and the numerical core system. *Frontiers in Human Neuroscience*, *5*, 186.
- Inglis, M., & Gilmore, J. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*, 147–155.
- Kaufmann, L., Koppelstaetter, F., Delazer, M., Siedentopf, C., Rhomberg, P., Golaszewski, S., ... Ischebeck, A. (2005). Neural correlates of distance and congruity effects in a numerical Stroop task: An event-related fMRI study. *NeuroImage*, *25*, 888–898.

- Leon, M.I., & Shadlen, M.N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron*, *38*, 317–327.
- Loftus, G.R., & Masson, M.E.J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin and Review*, *1*, 476–490.
- Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.
- Macmillan, N.A., & Creelman, C.D. (2005). *Detection theory: a user's guide*, 2nd edn. Mahwah, NJ: Erlbaum.
- Maloney, E.A., Risko, E.F., Preston, F., Ansari, D., & Fugelsang, J. (2010). Challenging the reliability and validity of cognitive measures: The case of the numerical distance effect. *Acta Psychologica*, *134*, 154–161.
- Matt, J., & Dzhafarov, E.N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review*, *121*, 1–32.
- Moyer, R.S., & Landauer, T.K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*, 1519–1520.
- Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience*, *6*, 177–190.
- Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K., & Fischer, M.H. (2011). Extending the mental number line a review of multi-digit number processing. *Journal of Psychology*, *219*, 3–22.
- Pansky, A., & Algom, D. (1999). Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 39–58.
- Pina, V., Castillo, A., Cohen Kadosh, R., & Fuentes, L.J. (2015). Intentional and automatic numerical processing as predictors of mathematical abilities in primary school children. *Frontiers in Psychology*, *6*, 375.
- Reber, R., Christensen, B.T., & Meier, B. (2014). Effects of meaning and symmetry on judgments of sizes. *Frontiers in Psychology*, *5*, 1270.
- Reike, D., & Schwarz, W. (2016). One model fits all: Explaining many aspects of number comparison within a single coherent model—A random walk account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 1957–1971.
- Risko, E.F., Maloney, E.A., & Fugelsang, J.A. (2013). Paying attention to attention: evidence for an attentional contribution to the size congruity effect. *Attention, Perception, & Psychophysics*, *75*, 1137–1147.
- Santens, S., & Verguts, T. (2011). The size congruity effect: is bigger always more? *Journal of Experimental Psychology: Human Perception and Cognition*, *118*, 94–110.
- Schwarz, W., & Eiselt, A.-K. (2009). The perception of temporal order along the mental number line. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 989–1004.
- Schwarz, W., & Eiselt, A.-K. (2012). Numerical distance effects in visual search. *Attention, Perception & Psychophysics*, *74*, 1098–1103.
- Schwarz, W., & Heinze, H.-J. (1998). On the interaction of numerical and size information in digit comparison: A behavioral and event-related potential (ERP) study. *Neuropsychologia*, *36*, 1167–1179.
- Schwarz, W., & Ischebeck, A. (2003). On the relative speed account of number-size interference effects in comparative judgments of numerals. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 507–522.
- Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology*, *3*, 334–349.
- Sobel, K.V., Puri, A.M., & Faulkenberry, T.J. (2016). Bottom-up and top-down attentional contributions to the size congruity effect. *Attention, Perception, & Psychophysics*, *78*, 1324–1336.
- Sobel, K.V., Puri, A.M., Faulkenberry, T.J., & Dague, T.D. (2016). Visual search for conjunctions of physical and numerical size shows that they are processed independently. *Journal of Experimental Psychology: Human Perception & Performance*. doi:10.1037/xhp0000323.
- Szűcs, D., & Soltész, F. (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia*, *45*, 3190–3202.
- Takahashi, A., & Green, D. (1983). Numerical judgments with Kanji and Kana. *Neuropsychologia*, *21*, 259–263.
- 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.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*, 483–488.
- Winter, B., Marghetis, T., & Matlock, T. (2015). Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex*, *64*, 209–224.