Categorizing digits and the mental number line

Dennis Reike¹ · Wolf Schwarz¹

Published online: 13 February 2019
© The Psychonomic Society, Inc. 2019

Abstract

Following the classical work of Moyer and Landauer (1967), experimental studies investigating the way in which humans process and compare symbolic numerical information regularly used one of two experimental designs. In *selection tasks*, two numbers are presented, and the task of the participant is to select (for example) the larger one. In *classification tasks*, a single number is presented, and the participant decides if it is smaller or larger than a predefined standard. Many findings obtained with these paradigms fit in well with the notion of a mental analog representation, or an Approximate Number System (ANS; e.g., Piazza 2010). The ANS is often conceptualized metaphorically as a mental number line, and data from both paradigms are well accounted for by diffusion models based on the stochastic accumulation of noisy partial numerical information over time. The present study investigated a categorization paradigm in which participants decided if a number presented falls into a numerically defined central category. We show that number categorization yields a highly regular, yet considerably more complex pattern of decision times and error rates as compared to the simple monotone relations obtained in traditional selection and classification tasks. We also show that (and how) standard diffusion models of number comparison can be adapted so as to account for mean and standard deviations of all RTs and for error rates in considerable quantitative detail. We conclude that just as traditional number comparison, the more complex process of categorizing numbers conforms well with basic notions of the ANS.

Keywords Categorization · Numerical distance effect · Mental number line · Diffusion models

Suppose you want to purchase a shirt, and have set a price frame between 20 (as a lower quality limit) and 50 USD (the maximum you can afford). On one occasion, an attractive shirt was offered for a given price and you had to decide whether to buy it or not. Or, a police officer looks for a wanted person described to be between 175 and 180 cm tall; s/he spots a resembling suspect of clearly recognizable height. What these, and many other, situations have in common is that a person must decide whether or not a given number (indicating, e.g., a price, or a height) falls into a given category that is numerically defined. Categorizing numbers is one elementary instance of mental operations that are studied in the area of numerical cognition. The multidisciplinary and many-faceted interest in this field of

research reflects that numbers and quantitative information assume crucial significance in many contexts, and that humans process with great efficiency numerical magnitude information as conveyed in symbolic formats, such as digits. Correspondingly, the way in which we learn, retrieve, process, and compare numerical magnitudes as symbolized by digits is among the best-investigated processes in cognitive psychology.

According to a prominent model of number processing (e.g., Dehaene & Brannon, 2011; Nieder, 2005; Reike & Schwarz, 2016, 2017; Rugani & de Hevia, 2017; Schwarz & Reike, 2019; Walsh, 2003), humans automatically convert digits into analog representations much like sensory representations of physical attributes such as brightness or length. These representations are often thought to be approximate, preverbal, and highly over-learned, and to have a quasi-spatial format, often referred to metaphorically as a “mental number line”. Convergent evidence supporting the notion of an Approximate Number System (ANS) comes from experimental, comparative, physiological, and developmental psychology and has been extensively reviewed (e.g., Lambrechts, Walsh, & van Wassenhove, 2013; Piazza, 2010; Whalen, Gallistel, & Gelman, 1999).

✉ Dennis Reike
reike@uni-potsdam.de

Wolf Schwarz
wschwarz@uni-potsdam.de

¹ Department of Psychology, University of Potsdam,
PO Box 60 15 53, 14415 Potsdam, Germany

Diffusion models (for general background, see Luce, 1986; Ratcliff & Smith, 2004) are often seen as a natural quantitative framework to incorporate the concept of analog magnitude representations. They have considerable biological validity because they successfully account, at the single-cell level, for known neural correlates of two-choice decisions (e.g., Gold & Shadlen, 2007).

The basic idea behind stochastic accumulation models is that the information on which participants base their decisions is not conveyed in an all-or-none fashion. Instead, each digit activates internal analog representations retrieved from memory. These representations provide noisy partial information about the digit, which is then accumulated over time. In two-choice designs, an overt decision is reached when the cumulated decision-related evidence reaches for the first time one of two critical levels, the response barriers. These barriers represent the critical amount of differential evidence required by the participant for the particular response with which they are associated. As the accumulated partial information is assumed to be noisy, this decision stage forms a stochastic process, which allows diffusion models to account for latency distributions and error rates. For a decision involving a given digit i , response latencies and error rates depend on the response barriers and the retrieval rate (the drift μ_i) which represents the mean evidence increment per time unit contributed by that digit. Converging evidence suggests that diffusion models provide a coherent conceptual framework that accounts in considerable quantitative detail for the way in which humans compare numbers (e.g., Kamienkowski, Pashler, Dehaene, & Sigman, 2011; Reike & Schwarz, 2016; Schwarz & Ischebeck, 2003, 1998; Sigman & Dehaene, 2005; Smith & Mewhort, 1998). More recently, the same model framework has been used to account, for example, for age effects in numerical cognition, both in young children (e.g., Thompson, Ratcliff, & McKoon, 2016) and in older adults (e.g., Reike & Schwarz, 2019).

A noticeable limitation of this line of research is that, following the seminal work of Moyer and Landauer (1967), studies investigating the way in which humans process and compare symbolic numerical information regularly are limited to one of two major experimental designs (cf., Dehaene, 1989). In *selection tasks* two different numbers are presented, and the task of the participant is to select, for example, the larger one. In *classification tasks*, a single number is presented, and the participant decides if it is smaller or larger than a predefined numerical standard.

Although a small number of studies explored number-related tasks not based on these standard designs (e.g., Dehaene & Akhavein, 1995; DeRosa & Morin, 1970; Schwarz & Stein, 1998) it remains true that the vast majority of studies on symbolic numerical comparison is based on selection and classification tasks. Therefore, it is at

present an open question if the main conceptual conclusions regarding the concept of mental analog representations (e.g., linear vs. logarithmic number representation, for discussion see, e.g., Dehaene, 2003; Nieder, 2005; Ratcliff & McKoon, 2018) generalize to paradigms which differ from these standard designs. More specifically, one relevant example, investigated in detail in the present study, concerns the type of cognitive process illustrated by the two introductory examples mentioned above: the categorization of numbers by numerical magnitude. Categorization tasks have long been studied in classical psychophysical contexts (e.g., Cartwright, 1941) but they have not been used systematically to investigate how we process symbolic numerical information.

A diffusion model of number categorization How can extant diffusion models of numerical comparison be generalized so as to account for number categorization, without changing their well-supported fundamental representational assumptions?

The central notion of our generalization is that a participant in a categorization task has to judge the digit's typicality, that is, the degree to which a presented digit corresponds to the ideal, or prototype, that defines the given category. For example, for a category formed by the successive integers [3, 4, 5, 6, 7], the property defining the prototype may be called the "fiveness" of the digit shown. In the present experiments, participants decided whether or not a single visually presented number falls into the central category [3, 4, 5, 6, 7]. As in our design both responses were equally likely, we assume that our participants have no systematic bias for or against category membership decisions, and place their response barriers about equally distant from the starting point. Therefore, the core of our model represents a set of assumptions about how the drift rate varies as a function of the digit shown.

The basic assumptions underlying this notion are depicted in Fig. 1 in which positive decisions about the category membership of a digit ("inside") are associated with the upper response barrier. Thus, we assume that with each digit i is associated a generic drift rate μ_i that is positive for numbers falling into the category and negative otherwise. The (positive) drift rate for evidence supporting category membership will then be maximal for the prototype digit, 5. Conceptually, the drift rate will change its sign at the boundaries of 2.5 and 7.5 delimiting the central category. Below 2.5 and above 7.5, the drift rates will be increasingly negative, in line with the lower response barrier indicating non-membership of the digit to the "fiveness" category ("outside"). As the core notion here is the diffusion-based judgment of how typical a given digit is relative to the category's prototype (e.g., Rosch, 1975), we will for short refer to this set of assumptions as the "typicality diffusion model" (TDM).

The conceptual logic underlying the TDM is well explained by analogy to standard categorization tasks (e.g., Macé, Joubert, Nespoulous, & Fabre-Thorpe, 2009), for example, when people judge whether the picture of an animal shows a bird (in our task, this target category consists of the central digits [3, 4, 5, 6, 7]). In this situation, according to the prototype theory subjects judge the “birdness” of the animal shown, much like the “fiveness” in our numerical task. A prototypical bird like a sparrow would correspond to the central digit 5, whereas an ostrich is one of the least typical birds that is still a category member, corresponding to the digits 3 or the 7 in our task. A bat is one of the elements closest to the “bird” category without actually being a member, corresponding to the digits 2 or the 8 in our task. In contrast, the numbers 1 and 9 correspond to, say, an ant or an elephant, that is, to an animal that is quite far away from the central “bird” category.

Figure 1 (left panels) illustrates these assumptions with two examples representing high vs. low sensitivity for judging category membership. In the first case (filled circles), the steep gradient of the linear digit-to-drift function corresponds to a sharp representational contrast between members and non-members of the category. Under these conditions, we expect relatively fast and error-free decisions to all digits, leading to a relatively small performance amplitude across different digits. In the second

case (open circles), the flat gradient of the digit-to-drift function corresponds to a more diffuse representational contrast between members and non-members of the category. Under these conditions, we expect relatively slow and error-prone decisions, corresponding to relatively large differences in speed and accuracy in responding to digits close to vs. far from the category boundaries. As shown in the right panels of Fig. 1, small vs. large barriers essentially govern the speed vs. accuracy trade-off, and modulate the performance amplitude across different digits. Barriers closer to the starting point ($\pm a'$ in Fig. 1) will generally reduce decision time and thus also digit-related differences in response speed; at the same time they will, however, increase differences in response accuracy. Conversely, barriers far from the starting point ($\pm a$ in Fig. 1) will increase digit-related differences in response speed; at the same time they will, however, reduce the error rate, and thus also differences in response accuracy.

The present study The present study had three major aims. First, we sought to broaden the empirical basis of current theorizing about numerical comparison processes, by asking if and how the process of number categorization can be understood in terms of basic principles gleaned from selection and classification paradigms. Second, we asked if current stochastic accumulation models of number comparison can be adapted in a consistent, seamless manner so that they are able to account parsimoniously for number categorization as well. Third, based on a unified model framework with a minimal number of free parameters, we aimed to account comprehensively not just for mean RT but also for error rates and RT standard deviations (SDs) of categorization performance.

Method

Participants

Thirty paid (8 euros) University of Potsdam students (mean age, 27) participated in one session lasting approximately 60 min.

Stimuli and apparatus

Each stimulus display consisted of one of the digits from 1 to 9, presented at the center of a 75-Hz, 1028 × 768 px (1 deg = 65 px) monitor; the display timing was synchronized with the video refresh cycle. The digits were presented in white against a black background in Verdana font using 36-px font height. A voice key was used to measure the onset of the vocal response.

Fig. 1 The typicality diffusion model (TDM) and its predictions for different drift rates and barriers. The *left panels* show the drift rate (*top*), predicted mean response time (RT, *middle*), and predicted error rate (*bottom*) for high (filled circles) and low (open circles) sensitivity for judging category membership. The *right panels* show the diffusion model (*top*), predicted mean RT (*middle*), and predicted error rate (*bottom*) for barriers emphasizing accuracy ($\pm a$, filled circles) and speed ($\pm a'$, open circles) barriers.

Procedure

In each of 11 blocks, each of the digits 3, 4, 5, 6, and 7 were presented eight times, and each of the other digits 1, 2, 8, and 9 ten times in random order. Each trial started with the presentation of a fixation point. After a random delay of 500–750 ms the digit was presented. The task was to indicate, as quickly as possible, whether the presented digit was within the range 3–7 (“inside”) or not (“outside”), by uttering the syllable “To” or “Ti” (the response mapping was counterbalanced between participants). The experimenter recorded the answer given by the participant. Finally, a blank screen was presented for 1250 ms.

Preliminary data reduction

The first block was considered training and excluded from all analyses. In addition, RTs more than ± 3 SDs from the participant’s mean RT were excluded (e.g., Miller, 1991), this led to the exclusion of 313 (1.3%) trials.

Results

For each participant and each digit, we computed mean RT, error rate, and the SD of RT from all valid, accepted trials. Mean RT, square root arcsin transformed error rates (Sokal & Rohlf, 1995), and variance of RT were each subjected to an ANOVA with the within-subjects factor numerical magnitude (1–9). As illustrated in Fig. 2, the numerical magnitude of the digit influences mean RT ($F(8, 232) = 14.86$, $MSE = 1000.62$, $p < .001$), error rate ($F(8, 232) = 10.59$, $MSE = .01$, $p < .001$), and the SD of RT ($F(8, 232) = 8.94$, $MSE = 8063036.34$, $p < .001$) in a highly systematic manner. To investigate the dependence of mean RT on the digit presented more systematically, we computed Bonferroni-corrected polynomial trends up to the order of four (e.g., Maxwell & Delaney, 2004). Consistent with the pattern shown in Fig. 2, the quartic (fourth-order) component was highly significant ($F(1, 232) = 102.44$, $MSE = 1290.34$, $p < .001$), representing a strong quartic trend effect ($\eta = 0.55$). In contrast, the contribution of the linear, quadratic, and cubic trend components were all negligible (all F 's < 2).

Diffusion model of number categorization We first fitted the TDM (Fig. 1) to the subject-averaged data (i.e., we first computed the digit-specific mean RTs, error rates, and RT variances for each participant separately, and only then averaged these quantities across participants), assuming the drift rates to decline linearly (for a relaxation of this assumption, see Footnote 1) as a function of the distance to the category prototype of five (as shown in the top left panel in Fig. 1). As indicated by the broken vertical lines

in Fig. 1, under the present design the TDM predicts the poorest performance at the hypothetical category borders (i.e., 2.5 and 7.5),

$$\mu_i = \begin{cases} k \cdot (i - 2.5) & i \leq 5 \\ k \cdot (7.5 - i) & i \geq 5 \end{cases}$$

The parameter k modulates how sensitively the accumulation process discriminates between digits falling within vs. outside of the central category. The barrier parameter a reflects how much evidence the average participant accumulates in order to decide whether the presented digit is within the target category or not. In addition, the non-decision component M includes all processes not involved in the actual decision process. The variance of the decision time (D) is defined by the basic diffusion process. In the TDM, overall RT consists of a decision component D , plus a non-decision component M . As a first approximation, we assumed D and M to be independent so that the variance of RT is the sum of the variances of D and M , $\sigma_{RT}^2 = \sigma_D^2 + \sigma_M^2$. Thus, predicting the SDs of RT requires one additional model parameter, σ_M^2 .

Mean correct RT, error rate, and SD of RT per digit were averaged across all 30 participants. We then divided the difference between predicted and observed mean RT by

Fig. 2 Mean correct response time (RT), mean error rate, and mean standard deviation (SD) for all nine digits (abscissa). The ordinate for each panel is from top to bottom: RT (ms), error rate (percent), SD (ms). Circles show the data and lines show the predictions of the typicality diffusion model (TDM). Error bars represent ± 2 standard error

its standard error, and squared this quantity to obtain an approximate $\chi^2(1)$. The difference between each observed and predicted error rate as well as the difference between observed and predicted SD of RT were evaluated by $\chi^2(1)$ components in the same manner. These three approximate $\chi^2(1)$ values per digit were then added, and aggregated across all nine digits. Finally, this goodness-of-fit statistic was numerically minimized in the four model parameters, a , k , μ_M , and σ_M^2 .

The best fit of the TDM is shown in Fig. 2. The response barrier estimate was $a = 11.66$. From the best-fitting sensitivity parameter, $k = 0.31$, we computed the net drift rate μ_i for each digit i . The mean of the non-decision component μ_M indicates that, on average, 620 ms of the RT were not involved in the decision process. The variance of the M stage was $\sigma_M^2 = 82.91^2$. As indicated by error bars in Fig. 2, the mean RT, error rate, and SD predictions are quite close to the observed data, with most predictions falling within the 95% confidence interval of observed mean RT, error rate, and SD¹. Note that the mean RTs for digits 1 and 9 are well in line with the model (top panel of Fig. 2), and show no additional extra speed-up (e.g., “end effect”), over and above the predictions of the TDM.

An additional advantage of the TDM is that it can be fitted to the data set of each participant individually. We used the same fitting procedure as described above to fit the TDM to the 30 individual data sets separately. As Table 1 shows, the parameters estimated by fitting the subject-averaged data are all well within the interquartile ranges of the parameters obtained from fitting the individual-subject data sets (see Table 1), suggesting that our TDM parameter estimates show a high degree of internal consistency.

General discussion

The present study sought to broaden the empirical basis of current theorizing about numerical comparison processes by asking our participants to decide if a given number falls into a predefined category (interval). Our main intention was to bridge, within a single coherent model frame, the wide conceptual gap between seemingly unrelated paradigms in numerical cognition, such as selection and classification tasks on the one hand, and categorization on the other.

Our results suggest that the process of categorizing numbers can be well understood by extending basic

concepts of current diffusion models of number comparison originally derived from more traditional selection and classification designs, although our model is certainly not the only way to conceptualize the task of categorizing numbers (e.g., Banks, Fujii, & Kayra-Stuart, 1976). More specifically, even though the TDM shown in Fig. 1 is highly constrained and uses only a minimal number of free parameters, it accounts comprehensively, in considerable quantitative detail, for the speed and accuracy with which humans categorize numbers (Fig. 2). Furthermore, the model is readily fitted to the RTs and error rates of each participant individually, and we find the median parameter estimates of these single-subject model fits to correspond consistently to the estimates derived from subject-averaged data. These results are remarkable as the course of the RT means in Fig. 2 shows a striking fourth-order polynomial pattern that is considerably more complex than the standard results from selection and classification designs which invariably show a simple monotone distance-dependent decline.

From the perspective of numerical cognition, our results complement and extend previous studies suggesting a fine-grained or analogical internal representation of numerical magnitude, the ANS. Thus, categorizing numbers as falling within or outside of a given numerical interval can be understood as a cognitive process that is perfectly consistent both with the fundamental representational assumptions inherent to the ANS, and also with basic assumptions (Fig. 1) about how these covert representations produce systematic features (means, SDs, error rates) of overtly observable decisions. At the same time, the excellent fit of the TDM implies critical and informative testable predictions of how numerical categorization processes could be modulated by manipulating further relevant experimental factors.

For example, the effects of stressing speed versus accuracy are predicted by varying the barrier parameter a while keeping all drift parameters constant (Fig. 1, right panels). Furthermore, the rates (μ_i) with which the numerical values of digits are retrieved have been shown to be modifiable, for example, by the digit's physical font size (the number-size congruency effect; Besner & Coltheart, 1979; Henik & Tzelgov, 1982; Reike & Schwarz, 2017). Diffusion models of this effect (e.g., Schwarz & Ischebeck, 2003) assume that both relevant numerical magnitude and irrelevant physical size information combine to influence the effective overall drift rate. In a categorization context, this mechanism generates predictions that are easily derived on the basis of the TDM in Fig. 1. Similarly, several recent studies (e.g., Thompson et al., 2016) suggest that during the early cognitive development of children the retrieval of information about numbers becomes considerably more efficient, which in the TDM should be reflected in

¹We also fitted in the same manner as described above a more general model in which the number-to-drift mapping was a power function with exponent β as additional free parameter. The best estimate $\hat{\beta}$ was 1.01, which corresponds essentially to the linear TDM. In addition, we also fitted logarithmic and exponential versions of the number-to-drift mapping. As compared to the linear TDM, both mappings resulted in considerably poorer model fits.

Table 1 Model parameter estimates

Parameter	Estimate from averages	Median of individual estimates	Interquartile range
a	11.66	11.12	[9.16, 17.48]
k	0.31	0.45	[0.29, 0.73]
μ_M	620	578	[542, 715]
σ_M^2	82.91 ²	80 ²	[62 ² , 99 ²]

Note. Parameter estimates from subject-averaged data, and medians and interquartile ranges of the 30 parameter estimates of the TDM model fitted to the mean RTs, error rates, and SD of RT of individual participants

corresponding increases of the values of the sensitivity parameter, k . More work is required to test these detailed predictions in the area of numerical categorization but the present model certainly provides a coherent theoretical basis on which related experiments can potentially improve our understanding of how we represent and process numerical information.

Acknowledgements We would like to thank Constantin G. Meyer-Grant for collecting the data.

The present work was supported by a research grant (SCHW 611/5-1) from the Deutsche Forschungsgemeinschaft (DFG).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 435–447.
- Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, 17, 467–472.
- Cartwright, D. (1941). Relation of decision-time to the categories of response. *American Journal of Psychology*, 54, 174–196.
- Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception & Psychophysics*, 45, 557–566.
- Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 314–326.
- Dehaene, S. (2003). The neural basis of the Weber–Fechner law: A logarithmic mental number line. *Trends in Cognitive Sciences*, 7, 145–147.
- Dehaene, S., & Brannon, E. M. (Eds.) (2011). *Space, time and number in the brain*. London: Academic Press.
- DeRosa, D. V., & Morin, R. E. (1970). Recognition reaction time for digits in consecutive and nonconsecutive memorized sets. *Journal of Experimental Psychology*, 83, 472–479.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574.
- Henik, A., & Tzelgov, J. (1982). Is 3 greater than 5: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, 10, 389–395.
- Kamienkowski, J. E., Pashler, H., Dehaene, S., & Sigman, M. (2011). Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making. *Cognition*, 119, 81–95.
- Lambrechts, A., Walsh, V., & van Wassenhove, V. (2013). Evidence accumulation in the magnitude system. *PLoS ONE*, 8, e82122.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York: Oxford University Press.
- Macé, M. J.-M., Joubert, O. R., Nespoulous, J.-L., & Fabre-Thorpe, M. (2009). The time-course of visual categorizations: You spot the animal faster than the bird. *PLoS ONE*, 4, e5927.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data*, 2nd Edn. New York: Taylor & Francis.
- Miller, J. O. (1991). Reaction time analysis with outlier exclusion: Bias varies with sample size. *The Quarterly Journal of Experimental Psychology*, 13A, 907–912.
- 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.
- Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, 14, 542–551.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Ratcliff, R., & McKoon, G. (2018). Modeling numerosity representation with an integrated diffusion model. *Psychological Review*, 125, 183–217.
- 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.
- Reike, D., & Schwarz, W. (2017). Exploring the origin of the number size congruency effect: Sensitivity or response bias? *Attention, Perception, & Psychophysics*, 79, 383–388.
- Reike, D., & Schwarz, W. (2019). Aging effects on symbolic number comparison: No deceleration of numerical information retrieval but more conservative decision-making. *Psychology and Aging*, 34, 4–16.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192–233.
- Rugani, R., & de Hevia, M.-D. (2017). Number-space associations without language: Evidence from preverbal human infants and non-human animal species. *Psychonomic Bulletin & Review*, 24, 352–369.
- Schwarz, W., & Stein, F. (1998). On the temporal dynamics of digit comparison processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1275–1293.
- 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.

- Schwarz, W., & Reike, D. (2019). The number–weight illusion. *Accepted for publication in Psychonomic Bulletin and Review*.
- Sigman, M., & Dehaene, S. (2005). Parsing a cognitive task: A characterization of the mind's bottleneck. *PLoS Biology*, 3, 334–349.
- Smith, D. G., & Mewhort, D. J. K. (1998). The distribution of latencies constrains theories of decision time: A test of the random-walk model using numeric comparison. *Australian Journal of Psychology*, 50, 149–156.
- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The principles and practices of statistics in biological research*, 3rd Edn. New York: W. H. Freeman.
- Thompson, C. A., Ratcliff, R., & McKoon, G. (2016). Individual differences in the components of children's and adults' information processing for simple symbolic and non-symbolic numeric decisions. *Journal of Experimental Child Psychology*, 150, 48–71.
- Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, 7, 483–488.
- Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science*, 10, 130–137.