Cognitive Psychology 91 (2016) 63–81

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

How numbers mean: Comparing random walk models of numerical cognition varying both encoding processes and underlying quantity representations

Dale J. Cohen [a,*], Philip T. Quinlan [b]

[a] University of North Carolina Wilmington, United States
[b] The University of York, United Kingdom

ARTICLE INFO

Article history:
Accepted 11 October 2016
Available online 7 November 2016

Keywords:
Numerical cognition
Random walk
Numerical distance
Physical similarity
Numerical architecture
Simulation

ABSTRACT

How do people derive meaning from numbers? Here, we instantiate the primary theories of numerical representation in computational models and compare simulated performance to human data. Specifically, we fit simulated data to the distributions for correct and incorrect responses, as well as the pattern of errors made, in a traditional "relative quantity" task. The results reveal that no current theory of numerical representation can adequately account for the data without additional assumptions. However, when we introduce repeated, error-prone sampling of the stimulus (e.g., Cohen, 2009) superior fits are achieved when the underlying representation of integers reflects linear spacing with constant variance. These results provide new insights into (i) the detailed nature of mental numerical representation, and, (ii) general perceptual processes implemented by the human visual system.

1. Introduction

Understanding how numerical symbols and their associated quantities are represented and used is a primary aspiration of those working in numerical cognition. Once we understand how symbols are represented and assigned meaning, we will understand how best to frame questions about how organisms become numerate (Cantlon, Platt, & Brannon, 2008; Feigenson, Dehaene, & Spelke, 2004), how numeracy changes over time (Geary, 1994), how and why numerical errors arise (Macaruso, McCloskey, & Aliminosa, 1993), how language and numeracy are related (Hauser, Chomsky, & Fitch, 2002), whether or not numeracy is a cultural universal (Dehaene, Isard, Spelke, & Pica, 2008), and, so on. Despite almost half a century of experimental study, the underlying representation of numerical symbols remains hotly debated.

Here, we present a model of number symbol encoding, representation, and retrieval. We instantiate this model in a random walk simulation and predict performance in the core paradigm in the field of experimental numerical cognition: the relative quantity task. Although we specifically simulate performance in the relative quantity task, the actual model informs more generally about how a number conveys the quantity it denotes. Indeed, the processes that we use to simulate performance in the task may reflect on more fundamental properties of the human perceptual system. Such general implications of the work are examined in detail later.

* Corresponding author at: Department of Psychology, University of North Carolina Wilmington, 601 South College Road, Wilmington, NC 28403-5612, United States.
E-mail address: cohend@uncw.edu (D.J. Cohen).

http://dx.doi.org/10.1016/j.cogpsych.2016.10.002
0010-0285/

1.1. The relative quantity task

The relative quantity task is a deceptively simple paradigm: On a typical trial, the participant must assess, as quickly and accurately as possible, which of two digits is the larger. The variations on the specific form of this task are numerous, ranging from the straightforward, simultaneous presentation of two Arabic digits (Dehaene, Dupoux, & Mehler, 1990; Hinrichs, Yurko, & Hu, 1981; Moyer & Landauer, 1967), to paradigms that include priming components (Dehaene et al., 1998; Ratinckx, Brysbaert, & Fias, 2005; Van Opstal, Gevers, De Moor, & Verguts, 2008), Flanker components (Notebaert & Verguts, 2006), multiple dimensions (e.g., in numerical Stroop tasks; Pavese & Umiltà, 1998; Ratinckx & Brysbaert, 2002; Tzelgov, Yehene, Kotler, & Alon, 2000; Waldron & Ashby, 2001), and, so on. Nevertheless, the reaction time (RT) data from the relative quantity task, and its variations, have revealed two effects that have proved foundational in all subsequent theory development; these are the *numerical distance effect* and the *size effect* (see Fig. 1A). The numerical distance effect is characterized by RTs (for correct responses) that monotonically decrease as the numerical distance between the two digits increases. The size effect is that, for a fixed difference between two digits, correct RTs increase monotonically as the size of the digits increase.

Moyer and Landauer (1967) conducted the classic experiment that identified the numerical distance effect and size effects. The authors presented two Arabic digits, side-by-side, and asked participants to identify as quickly and accurately as possible the numeral denoting the larger quantity. Moyer and Landauer (1967) explained the numerical distance effect and size effect by proposing that numbers are represented as magnitudes that are similar to those in the physical world and the discriminability of two perceived magnitudes is determined by the ratio of the actual magnitudes (i.e., these representations obey Weber's law).

Since the original description of the numerical distance and size effects, various accounts have been put forward to explain the psychological representations of numbers hypothesized by Moyer and Landauer (1967). All such accounts adopt a general Signal Detection Theory framework. That is, these accounts generally assume that the quantity associated with a given digit is represented on an internal continuum (e.g., an internal "number line," Dehaene, 2003) and there exists some perceptual variability (i.e., noise) associated with the placement of the digit on this continuum. Take, for example, the number "5." All theories assume that each time an observer experiences this symbol, the observer will have a slightly different "sense" of the quantity associated with "5." So, sometimes the observer's sense is more than 5 and sometimes less. Accordingly, each digit's quantity representation is captured by a distribution of values on the continuum that we term a *psychological distribution of quantity* (PDQ) (see Fig. 1B). The PDQ captures the perceptual noise associated with one's understanding of the quantity associated with a digit. The PDQs of successive digits are rank ordered and overlap (see Fig. 2).

In Signal Detection Theory, the degree of difficulty in distinguishing between two stimuli is determined by the amount of overlap between their corresponding perceptual distributions: the greater the overlap, the more difficult it is to distinguish between the two stimuli. This premise translates directly when discussing number representation. Most accounts explaining the psychological representations of numbers, assume the difficulty in distinguishing between the quantities of two numbers is determined, primarily, by the amount of overlap between their PDQs (see Fig. 2). 'Difficulty,' in this context, is defined by greater RTs and errors in the relative quantity task.

The fundamental differences between the key theories lie with the assumptions they make about the nature and spacing of the PDQs on the mental number line (see Fig. 1). According to the linear account, successive quantity representations are rank ordered at equal intervals and the different PDQs have the same variance (we call this account the *Linear Theory*, see e.g., Cantlon, Cordes, Libertus, & Brannon, 2009). The two other competing accounts of quantity representation are the *Logarithmic Theory* (Dehaene, 1992, 2003) and the *Scalar Variance Theory* (see Church, Meck, & Gibbon, 1983; Gallistel & Gelman, 1992; Meck & Church, 1983; Meck, Church, & Gibbon, 1985). The Logarithmic Theory also posits that the different PDQs have the same variance, but claims that the means of the ordered PDQs are spaced on a logarithmic scale. As such, the means of successive PDQs get closer together as the numbers increase (see Fig. 1B). In contrast, for the Scalar Variance Theory, the means of the ordered PDQs are spaced linearly but their variances scale linearly (i.e., increase) with quantity (see Fig. 1B).

In the absence of providing a detailed model, intuitions about general patterns of performance suggest that the numerical distance effect can be accommodated by all three theories because the PDQs of numerals denoting adjacent quantities (e.g., "5" and "6") overlap more than the PDQs of numerals denoting distant quantities (e.g., "5" and "1"). It is similarly apparent that the Logarithmic Theory and Scalar Variance Theory accounts can accommodate the size effect. The size effect is hypothesized to result because, for a given quantity distance (e.g., "1") the PDQs of numerals denoting large quantities (e.g., "7" and "8") overlap more than the PDQs of numerals denoting small quantities (e.g., "2" and "3"). For the Logarithmic Theory this is true because the means of the PDQs for successive smaller quantities are farther apart than those for successive larger quantities. For the Scalar Variance Theory this is true because the SDs of the PDQs for successive smaller quantities are smaller than those for successive larger quantities. On first glance, though, the size effect sits less well with the Linear Theory.

This, in brief, is the current state of affairs with respect to explaining performance in the relative quantity task and the evidence relevant to discussion of the underlying representation of integers. Here, we assess the validity of the stated models and a new model; by (i) specifying the details of models of this task from encoding to response, (ii) simulating data based on the specified details, and, (iii) and assessing the model fits against human data.