Available online at www.sciencedirect.com
ScienceDirect
Journal of Memory and Language 59 (2008) 495–523
Journal of Memory and Language
www.elsevier.com/locate/jml

Beyond mean response latency: Response time distributional analyses of semantic priming

David A. Balota [a,*], Melvin J. Yap [b], Michael J. Cortese [c], Jason M. Watson [d]

[a] Department of Psychology, Washington University, St. Louis, MO 63130, United States
[b] Department of Psychology, National University of Singapore, Block AS 4, #02-07, 9 Arts Link, Singapore 117570, Republic of Singapore
[c] Department of Psychology, University of Nebraska at Omaha, 6001 Dodge Street, Omaha, NE 68182, United States
[d] Department of Psychology, University of Utah, 380 South, 1530 East, Room 502, Salt Lake City, UT 84112, United States

Received 1 March 2007; revision received 3 October 2007
Available online 4 March 2008

Abstract

Chronometric studies of language and memory processing typically emphasize changes in mean response time (RT) performance across conditions. However, changes in mean performance (or the lack thereof) may reflect distinct patterns at the level of underlying RT distributions. In seven experiments, RT distributional analyses were used to better understand how distributions change across related and unrelated conditions in standard semantic priming paradigms. In contrast to most other lexical variables, semantic priming in standard conditions simply shifts the RT distribution, implicating a headstart mechanism. However, when targets are degraded, the priming effect increases across the RT distribution, a pattern more consistent with current computational models of semantic priming. Interestingly, priming effects also increase across the RT distribution when targets are degraded and primes are highly masked, supporting a memory retrieval account of priming under degraded conditions. Finally, strengths and limitations of alternative approaches for modeling RT distributions are discussed.

Keywords: Semantic priming; Masked priming; Distributional analysis; Lexical decision; Pronunciation

Breakthroughs in science often reflect improvements in the measurement tool investigators use to study a phenomenon. This can be most obviously seen in fields such as astronomy and biology, wherein the developments of higher magnification systems opened up new worlds for exploration. The recent advances in neuroimaging methods are another prime example of the power of measurement development. The present paper describes a step in this direction by increasing the magnification of the chronometric tools used to study psycholinguistic, and other response time (RT) dependent, phenomena.

Chronometric studies of language, memory, and attention have accumulated a vast amount of knowledge regarding the nature of representations, the processes engaged to tap such representations, and the time-course of the interactions between representations and processes. In order to better understand how one might increase the magnification of the standard chronometric approach, let us briefly consider the implicit assumptions researchers make.

In standard paradigms, researchers often manipulate a variable by including multiple observations (typically

* Corresponding author.
E-mail address: dbalota@artsci.wustl.edu (D.A. Balota).

0749-596X/$ - see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.jml.2007.10.004

10–20) at each level of an independent variable (IV). A mean is then typically calculated for each level of an IV, and these means are submitted to inferential tests (most often analyses of variance) to estimate how reliable effects are across participants (and/or across items). Consider the classic semantic priming effect, which we will target in the present study. Here, the finding is that participants produce faster response latencies to a target, when the target word is related to a prime word (e.g., DOCTOR–NURSE), compared to when it is unrelated (e.g., FOREST–NURSE)[1]. The implicit assumption that researchers make is that the related and unrelated conditions produce symmetric RT distributions, and hence, the mean is a reasonably good estimate of the central tendency of these distributions. So, if one observes a 50 ms semantic priming effect, this indicates that the distribution of the related condition is shifted 50 ms away from the unrelated condition.

However, we all know that this implicit assumption is wrong. That is, RT distributions are rarely symmetrical around a mean, but are almost always positively skewed (see Luce, 1986, for a comprehensive review). Fig. 1 reflects an RT distribution from a single participant across approximately 2400 observations in lexical decision performance. Notice the strong positive skewing of the distribution. Hence, returning to the 50 ms semantic priming effect in the means, we are confronted with a number of first-order reasons why one might obtain such a difference: (a) The modal portion of the distribution may shift, without changing the tail; (b) The tail of the distribution may increase without changing the modal portion of the distribution; (c) Both the modal portion and tail may increase.

If researchers know that RT distributions are skewed, and that there are multiple ways in which an effect in means may be observed, then why does the field continue to use estimates of the mean to gain insights into the cognitive architecture? Clearly, there are many advantages in support of the mean. First, and probably most importantly, the mean is relatively easy to calculate and understand. Means are a fundamental summary statistic and dominate much of our common knowledge of the world (e.g., mean income, average miles per gallon, batting average, etc.). Second, the estimates are relatively stable. Why should one worry about the underlying distributions if the effects with means are replicable across studies? Third, and related to this, higher-order estimates of the RT distribution, such as skewness and kurtosis, are considerably less reliable (see Ratcliff, 1979). Why spend the additional

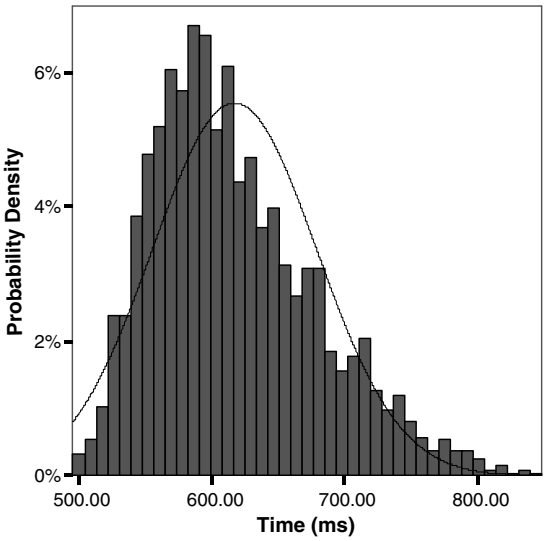

Fig. 1. Response time distribution for lexical decision performance across 2428 words taken from Balota et al. (2004).

effort to capture more subtle aspects of RT distributions if there is indeed a lack of stability in these estimates? In order to obtain stable estimates of higher order moments, one needs considerably more observations then the standard 10–20 observations per participant/cell. Does the added benefit justify the cost?

Although there are advantages to the mean, we, along with many others (e.g., Heathcote, Popiel, & Mewhort, 1991; Luce, 1986; Ratcliff, 1979; Rouder, Lu, Speckman, Sun, & Jiang, 2005; Van Zandt, 2002), believe that the *zeitgeist* is appropriate for researchers to move beyond the mean. The goal of the present paper is to provide a review of recent developments and extensions of RT distributional analyses to visual word recognition research. We should emphasize here that these arguments are not restricted to psycholinguistic variables, but indeed are relevant to all chronometric explorations of performance. However, in order to exemplify the power of this approach, we will focus on one of the most frequently studied effects in language and memory processing, i.e., the semantic priming effect.

Measuring aspects of the RT distribution: Beyond the mean

If it is time to move beyond the mean in estimating the influence of a variable or variables on RT distributions, how might one measure such influences? There are typically three major approaches that are used in the literature. First, one may have an explicit model that predicts how an underlying RT distribution may change as a function of a manipulation. Hence, one can simply fit the empirical data to the model's specific predictions

[1] Here we use the term "semantic" priming effect for simplicity; however, it should be noted that some, if not most, of the priming effects observed in these tasks may reflect associative relations, instead of semantic (see Hutchison, 2003, for a review).