

# An Integrated Model of Choices and Response Times in Absolute Identification

Scott D. Brown  
University of Newcastle

A. A. J. Marley  
University of Victoria

Christopher Donkin and Andrew Heathcote  
University of Newcastle

Recent theoretical developments in the field of absolute identification have stressed differences between relative and absolute processes, that is, whether stimulus magnitudes are judged relative to a shorter term context provided by recently presented stimuli or a longer term context provided by the entire set of stimuli. The authors developed a model (SAMBA: selective attention, mapping, and ballistic accumulation) that integrates shorter and longer term memory processes and accounts for both the choices made and the associated response time distributions, including sequential effects in each. The model's predictions arise as a consequence of its architecture and require estimation of only a few parameters with values that are consistent across numerous data sets. The authors show that SAMBA provides a quantitative account of benchmark choice phenomena in classical absolute identification experiments and in contemporary data involving both choice and response time.

*Keywords:* absolute identification, absolute identification models, response time, response time distributions, sequential effects

Performance in absolute identification tasks has fascinated researchers for over 50 years (e.g., Garner, 1953; Miller, 1956; Pollack, 1952, 1953). Research in the past 35 years has emphasized both data and formal theories (e.g., Braida & Durlach, 1972; Durlach & Braida, 1969; Laming, 1984; Lockhead, 2004; Luce, Nosofsky, Green, & Smith, 1982; Marley & Cook, 1984; Petrov & Anderson, 2005; Stewart, Brown, & Chater, 2005; Treisman & Williams, 1984) and, most recently, has been concerned with both the choices made and the time it takes to make them (Kent & Lamberts, 2005; Lacouture & Marley, 1991, 1995, 2004). As Shiffrin and Nosofsky (1994, p. 358) stated in an article reassessing the significance of Miller's (1956) classic paper, "absolute identification has captured the imagination . . . not only because the empirical results are so startling . . . but also because [they] provide perplexing prob-

lems for classic psychophysical models." Luce (1986, chapter 10) gave an excellent summary of data and theory to that date, and Lockhead (2004) summarized data and theory most relevant to relative interpretations of absolute identification, where the relativity is with respect to stimuli and responses from previous trials. Stewart et al. (2005) and Petrov and Anderson (2005) provided comprehensive reviews of choice data and the related theory, with emphasis on theoretical approaches over the past 20 years.

A typical absolute identification task requires a participant to identify, on each trial, which stimulus has been presented from a relatively small prespecified set. In general, people are unable to accurately identify more than about 8–10 stimuli that vary on a single physical dimension. For example, the stimuli might be a set of 10 lines varying only in length, with the shortest line labeled 1 and the longest 10. A participant previews the entire labeled set and is then shown the lines one at a time, over numerous trials, and asked to identify the presented line with the appropriate response label. Typically, a participant in this task is unable to achieve an overall accuracy above about 80%, which is surprising given that the stimuli are chosen such that comparative judgments of any pair of them are completely accurate (i.e., judging whether one stimulus is smaller or greater than another stimulus presented in rapid succession).

With such an extensive history, the study of absolute identification is a mature field with many well-established benchmark behavioral phenomena that describe how choices and response times (RTs) are affected by stimulus manipulations and by the history of stimuli and responses. We broadly separate these phenomena into global and local effects.

---

Scott D. Brown, Christopher Donkin, and Andrew Heathcote, School of Psychology, University of Newcastle, Newcastle, Australia; A. A. J. Marley, Department of Psychology, University of Victoria, Victoria, British Columbia, Canada.

This research has been supported by Australian Research Council Linkage International Project LX0348125 to Doug Vickers, A. A. J. Marley, Andrew Heathcote, Philip Smith, and Michael Lee; Natural Science and Engineering Research Council Discovery Grant 8124-98 to A. A. J. Marley; and Australian Research Council Discovery Project DP0881244 to Scott D. Brown and Andrew Heathcote. We are indebted to Doug Vickers, who inspired us to begin this line of research but passed away before it came to fruition.

Correspondence concerning this article should be addressed to Scott D. Brown, School of Psychology, University of Newcastle, Callaghan NSW 2308, Australia. E-mail: scott.brown@newcastle.edu.au

1. *Global effects: Stimulus range and set size.* For a fixed set size ( $N$ ) of stimulus–response pairs, performance—measured by the amount of information transmitted—increases quickly to an asymptotic level of 2–3 bits as the range of stimuli on the physical dimension increases (Braidia & Durlach, 1972). Similarly, as set size increases, the amount of information transmitted increases rapidly at first but then asymptotes at 2–3 bits (e.g., Garner, 1953; Pollack, 1952, 1953). Responses to the largest and smallest stimuli in a set are faster and more accurate than responses to the middle stimuli—the bow effect. Both accuracy and RTs worsen for any given stimulus as other stimuli are introduced to the set (Kent & Lamberts, 2005; Lacouture & Marley, 1995). As shown by Luce et al.’s (1982)  $d'$  analysis,<sup>1</sup> bow effects in accuracy are due partly to bow effects in sensitivity and partly to response bias produced by the constraints on available responses for stimuli near the ends of the range. Ward (1987) showed that scaling methods that require either relative or absolute judgments reveal profound effects of stimulus–response mappings over days.
2. *Local effects: Sequential effect on accuracy and errors.* Previous stimuli and responses can affect the response to the current stimulus (e.g., Lacouture, 1997; Ward & Lockhead, 1970, 1971). When an incorrect response is given, it tends to be toward, rather than away from, the stimulus from the previous trial (assimilation). The opposite pattern occurs for longer lags (contrast): Errors tend to be away from, rather than toward, stimuli from two or more trials previously.<sup>2</sup> Accuracy is improved when stimuli are constrained to be similar on successive trials (e.g., Luce et al., 1982). In particular, the difference in magnitudes between the stimuli presented on the current and previous trials influences response accuracy (e.g., Petrov & Anderson, 2005; Rouder, Morey, Cowan, & Pfaltz, 2004; Stewart et al., 2005).

Previous theoretical accounts of absolute identification have attempted to show that some or all of these phenomena could be accounted for using only relative processes or only absolute processes. A model that uses only absolute processes is one where decisions are made about stimulus magnitudes on the basis of comparisons with some longer term referents (e.g., the context-coding component of Braidia et al., 1984; Lacouture & Marley, 1995, 2004) or a longer-term frame of reference (e.g., Marley & Cook, 1984). On the other hand, a model that uses only relative processes (e.g., Laming, 1984; Lockhead, 2004; Stewart et al., 2005) posits that decisions are made using only comparisons with recent stimuli and responses. Range and set size effects have most often been attributed to absolute processes (e.g., Braidia et al., 1984; Marley & Cook, 1984). Sequential effects, particularly assimilation and contrast, have been frequently explained by shorter term relative judgment processes. Other than the model developed in this article, there is no model, either relative or absolute, that accounts for all of the global and local benchmark phenomena described above.

Stewart et al. (2005) presented a league table (their Table 2) comparing absolute identification models on their ability to account for nine benchmark phenomena, along with associated clas-

sical data sets, under three broad headings: limited information transmission, bow (set size) effects, and sequential effects. The most comprehensive relative theory, Stewart et al.’s relative judgment model (RJM<sup>3</sup>), performs well on choice-related phenomena but does not address RTs. The most comprehensive absolute theory, Lacouture and Marley’s (1995, 2004) mapping model, performs well on global phenomena for both choice and RT but does not address sequential effects. Stewart et al. counted RT as one benchmark phenomenon, even though a wide range of benchmark RT phenomena have been identified (Lacouture, 1997; Lacouture & Marley, 1995, 2004), some paralleling those found in choices, including set size effects, stimulus magnitude effects, and sequential effects, and some specific to RT, such as distribution shapes for correct and error responses for each stimulus.

We propose a model of both choice and RT in absolute identification in which sequential effects, including assimilation and contrast, result from short-term memory effects in an absolute judgment process. Our model further develops key concepts from several previous models. Reflecting this cumulative history, the acronym for the model, SAMBA, highlights the three core elements: selective attention (Marley & Cook, 1984), mapping (Lacouture & Marley, 1995, 2004), and ballistic accumulation (Brown & Heathcote, 2005). Stewart et al. (2005), in their league table, showed that the separate components of SAMBA are, by themselves, inadequate. This gives motivation for SAMBA, which integrates and extends these components, to account for assimilation, contrast, asymmetries in bow effects, and local judgment effects. SAMBA provides a unified account of choice phenomena as well as the associated RT phenomena, as we demonstrate by fitting SAMBA to Lacouture’s (1997) full range of choice and RT data, which is averaged over participants, and to Kent and Lamberts’s (2005) and Lacouture and Marley’s (2004) individual subject data. While SAMBA is the first model of absolute identification to provide a comprehensive account of RTs and response choices that includes sequential effects, of course we do not claim that other such models cannot be developed. For example, both Lacouture and Marley’s (2004) and Kent and Lamberts’s models provide a good account of RTs and most choice phenomena but do not explain sequential effects, but these models may be further developed to cover such effects. Similarly, models that cover response choices but not RTs (such as Stewart et al.’s, 2005, RJM or Petrov & Anderson’s, 2005, ANCHOR) may be developed further to include an RT mechanism. We note, however, that an account of RT added to a choice model is not guaranteed success—see, for example, Karpiuk, Lacouture, and Marley (1997).

Recently, Brown, Marley, and Lacouture (2007) highlighted the theoretical importance of sequential effects in accuracy (see also Petrov & Anderson, 2005; Rouder et al., 2004; Stewart et al., 2005), thus going beyond just assimilation and contrast, which describe sequential effects in errors. Brown et al. focused on Rouder et al.’s (2004) analysis of accuracy as a function of the

<sup>1</sup> Throughout this article, we use Luce et al.’s (1982) method of calculating  $d'$  to quantify sensitivity.

<sup>2</sup> Since stimuli and correct responses are correlated, care must be taken in interpreting these as solely stimulus (or response) effects.

<sup>3</sup> The term *relative judgment model* was earlier used by DeCarlo and Cross (1990) for their model of magnitude scaling—their Equation 16.

difference between the current and previous stimuli. Analyzing data from Lacouture's (1997) line length task, they observed improved accuracy for a stimulus similar to the one before and also for a stimulus very different from the one before. Subsequently, Stewart (2007) noted the same pattern in three other data sets (Kent & Lamberts, 2005; Neath & Brown, 2006; Stewart et al., 2005). Brown et al. attributed higher accuracy when successive stimuli are similar to a comparison of the present stimulus with the prior stimulus and higher accuracy when successive stimuli are very different to comparison with an end stimulus. Stewart explained the pattern by modifying the RJM to include a memory for the stimulus two trials back that on some trials is used instead of the memory for the stimulus one trial back as the basis for relative judgment. Other experiments by DeCarlo and Cross (1990) and DeCarlo (1994) demonstrated that instructions significantly impact whether magnitude judgments are made relative to short- or long-term referent stimuli and responses. These findings present problems for models that rely solely on absolute or solely on relative mechanisms, requiring suitable extensions of the relative approach, as shown by Stewart, and a suitable extension of the absolute approach, as we show for SAMBA.

SAMBA is an integrative model not only because it accounts for choice and RT data but also because it includes relative as well as absolute processes. We show that SAMBA is able to model the complex effects shown in Brown et al.'s (2007) analysis by replacing one of the end anchors used by Marley and Cook's (1984) selective attention process with the estimated magnitude of the previous stimulus. However, this relative judgment process is not required for SAMBA's account of classic absolute identification phenomena, including assimilation and contrast, so we omit it in our fits to benchmark data sets exemplifying these phenomena. Only Lacouture's (1997) and Stewart et al.'s (2005) data show effects that are strong enough to require explicit modeling by the relative judgment process in SAMBA.

In the following sections, we first provide details of SAMBA, then describe a set of benchmark empirical phenomena and associated classic data sets. Along with these descriptions, we show that SAMBA accurately fits each of these classic benchmark data sets. The benchmark sets presented here were chosen to provide insight into the workings of SAMBA and to illustrate its account for phenomena beyond the scope of any of the models from which it was derived, with a particular emphasis on sequential effects. Note that SAMBA also accounts for many other benchmark phenomena that we do not have space to illustrate here, such as the effects of set size on RT. Finally, we present comprehensive fits of SAMBA to two data sets: Lacouture (1997) and Stewart et al.'s (2005) Experiment 1. We used Lacouture's data to test SAMBA's ability to simultaneously account for all of the choice and RT phenomena in a complex data set using a single set of parameter values. Although Stewart et al. did not collect RTs, their data are important because of experimental manipulations that allow for a strong test of SAMBA and for comparison of its fits with those of the RJM.

### The Theoretical Challenge

The paradox of absolute identification is that the task is superficially very simple yet the performance of participants is both inaccurate and surprisingly complex. Our approach to this chal-

lenge is similar to that taken by Ratcliff's (1981) theory of perceptual matching: We explicitly model an integrated architecture for perceptual, memory, and decision processes in sufficient detail to obtain predictions for the broadest possible range of observed behavior. However, our approach differs from Ratcliff's and many other theories of absolute identification, which assume that repeated presentations of any given stimulus result in a distribution of internal magnitude estimates and that certain parameters of the distribution, often the variance, must be estimated for each experimental context in which the stimulus appears. Such approaches are analogous to signal detection theory, providing a successful description of the data without addressing the deeper question of how these magnitude estimates arise. As well as being less intellectually satisfying, models that begin with parameterized distributions also fail to provide constrained accounts of some of the most fundamental benchmark phenomena. For example, there are powerful effects caused by very simple stimulus factors such as set size ( $N$ ), stimulus spacing, and stimulus rank order. In a framework similar to Ratcliff's, these effects are modeled by changes in parameter estimates, but this approach fails to provide an explanation of how the changes arise.

With SAMBA, we adopt a less flexible approach that provides strict constraints and reduces the number of parameters required. We directly model the process by which an observer produces a magnitude estimate when confronted with a stimulus (see also Kent & Lamberts, 2005). We model this process using an extended version of Marley and Cook's (1984) selective attention theory. This theory explains how an observer can attach a numerical magnitude estimate to a stimulus and how these magnitude estimates are distributed across repeated trials. This establishes a mechanism for producing the distributions and for describing how those distributions change under experimental manipulations without requiring arbitrary parameter changes. The remaining parts of the SAMBA architecture provide a similarly constrained explanation of the process by which a response is chosen in light of the magnitude estimate.

An even deeper philosophical question arises when considering the theoretical completeness of the simple parametric approach to stimulus representations. The participant's task in absolute identification is to attach a response label, such as 1, 2 and so on, to a stimulus, with the physical stimulus measured in, for example, decibels, hertz, or meters. A theoretical account is incomplete if it begins by assuming that numbers are attached to stimuli and that these are simply transformed into response labels. Lacouture and Marley's (1995, 2004) and Stewart et al.'s (2005) models suffer from this weakness. Put another way, the central task of absolute identification is to associate a numeral with a stimulus magnitude, so it is dissatisfying to consider a theory that begins by assuming that stimulus magnitude estimates are available with no explanation of how they arise.

SAMBA differs from the simple distributional approach in a second way that greatly reduces the number of parameters that must be estimated in fits to data. Like signal detection theory, simple distributional models typically make decisions using a set of cutpoint or referent parameters. The disadvantage of this approach is that the number of referent parameters grows linearly with the number of response alternatives, and usually, no explanation is provided of how participants choose appropriate values. SAMBA adopts a framework similar to that of Petrov and Ander-

son's (2005) ANCHOR model: We assume that participants learn average magnitude estimates corresponding to each response and that these magnitudes act as referents. As this learning is assumed to be accurate, at least when feedback is correct, the referent values are entirely determined by the experimental design, and hence, in model fitting, they do not have to be estimated. Consequently, SAMBA's flexibility in fitting data is greatly curtailed. Lacouture and Marley's (1995) parameter-free mapping model is used to transform referent estimates to what are, effectively, a set of tuning curves that provide input to SAMBA's decision stage. Hence, we meet the theoretical and practical challenge posed by the potentially large number of responses in absolute identification, as the number of parameters required to fit SAMBA does not change with the number of responses.

### Model Overview

SAMBA integrates three successful elements from previous models. It uses the time-dependent (i.e., not the asymptotic) version of Marley and Cook's (1984) selective attention model of stimulus representation and Brown and Heathcote's (2005) ballistic accumulator model of response selection.<sup>4</sup> These model elements do not constitute a complete account because the selective attention component produces a single magnitude estimate while the ballistic accumulator component requires  $N$  numerical inputs, one for each possible response. We link the two components with Lacouture and Marley's (1995) mapping process. The vertical integration of model components is an important feature of SAMBA. Often, in cognitive psychology, different levels of processing are considered separately, and models are developed independently for each. Greater model constraint and explanatory power can be achieved by integrating models to span several levels of explanation, from stimulus representation to response selection (see Ratcliff, 1981, for a similar approach to perceptual matching). The three elements we borrow from prior models successfully explain some aspects of absolute identification behavior, but not others, making each incomplete. Some of these problems are naturally fixed by integration within SAMBA, but others are not. We address these remaining issues by modeling the selective attention and ballistic accumulator processes at the level of the duration of the trial; the model can be specified in real time as more temporally fine-grained data become available (see the General Discussion).

Figure 1A illustrates the three stages of the model. SAMBA's first stage is a modified version of Marley and Cook's (1984) selective attention theory. The selective attention (SAMBA) stage maintains a representation for the context of the experiment and uses this context to produce estimates of sensory magnitude. The context representation is maintained by activation of a range of units that are in one-to-one correspondence with stimulus magnitudes, such as line lengths or intensity of tones of the same frequency. Input to the selective attention stage comes from a relatively accurate psychophysical representation of each stimulus. We assume that the stimuli are represented topographically at the psychophysical level and that psychophysical input causes selection of a corresponding unit. Importantly, neither the psychophysical representation nor the selected unit directly provides a numerical estimate of the stimulus magnitude. In this, we agree with Krantz's (1972, p. 175) view that "I do not see how sensations

could be paired directly with numbers at all." Instead, stimulus magnitudes are estimated by the summation of activity between the unit selected by the psychophysical input and the ends of the active context.

This magnitude estimate is transformed into  $N$  response strengths<sup>5</sup>—one for each of the  $N$  possible responses—by Lacouture and Marley's (1995) mapping. The transformation works using a long-term memory for each stimulus, given by the average of its magnitude estimates on previous trials. The final stage of SAMBA is an elaborated version of Brown and Heathcote's (2005) ballistic accumulator. This stage takes the  $N$  response strengths produced by the mapping and assigns each one to a separate decision accumulator. The activations of these accumulators increase at rates determined in part by the response strengths' output from the mapping and in part by the dynamics of the accumulators, including mutual inhibition. A response is made as soon as any accumulator's activation exceeds a response threshold. RT is modeled as the time taken for that accumulator to reach threshold plus a constant amount of time taken for nondecision processes. We now describe each stage in more detail.

As input to SAMBA, we assume a simple spatial psychophysical stimulus representation corresponding to topological projections in the sensory areas of the brain, such as retinotopic or tonotopic maps (e.g., Romani, Williamson, & Kaufman, 1982; Wiemer & von Seelen, 2002). The physical magnitudes of stimuli are mapped quite accurately onto this psychophysical dimension, with only a small amount of variability. Most absolute identification experiments involve stimuli that are sufficiently separated so that psychophysical variability does not cause errors. Hence, variability in the psychophysical stage is neglected in the fits of SAMBA that we report, except in two cases where stimuli are closely spaced.

The selective attention stage produces a magnitude estimate from the ordinal psychophysical representation. This process, with the addition of decision cutpoints, has had considerable success in fitting classic choice data in absolute identification (Marley & Cook, 1984) and also magnitude estimation data (Marley & Cook, 1986). The selective attention model has been successful because it provides a stimulus representation that dynamically adapts to changes in the experimental design, such as stimulus spacing and choice set size. Although motivated differently, this stage performs similarly to the successful theory of sensory trace and context coding (see, e.g., Berliner, 1973; Berliner, Durlach, & Braida, 1977; Braida et al., 1984) and to the attention band proposals of Weber, Green, and Luce (1977), Luce et al. (1982), and Nosofsky (1983).

Selective attention gives SAMBA a mechanism by which to attach context- and time-dependent numerical representations to stimuli and stimulus differences, without making arbitrary assignments. For example, suppose an experiment uses pure tone stimuli, all of the same frequency but of different intensities. Over preex-

<sup>4</sup> Software to implement the model is available from <http://science-it.newcastle.edu.au/~sdb231/software/samba/>

<sup>5</sup> We use *response strengths* instead of *mapping outputs* or *drift rates* simply as a mnemonic convenience. We hope the reader will be reminded by this terminology that these (unobserved) quantities drive the ballistic accumulators toward making an overt response.

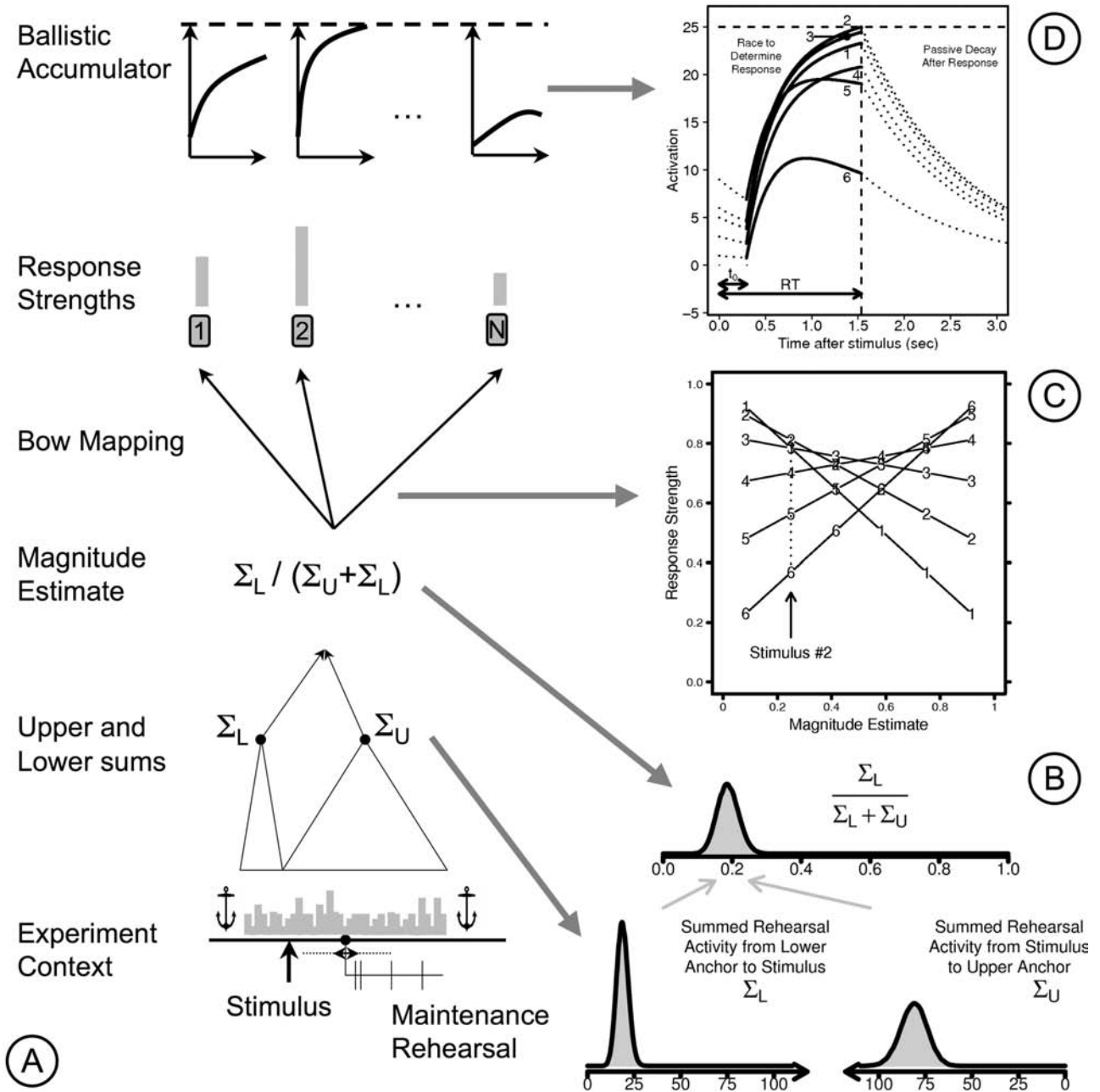


Figure 1. The SAMBA model. Panel A illustrates the model at a general level, to be read from bottom to top. Panel B shows how the magnitude estimate produced by the selective attention stage varies over trials. Summed activities from the stimulus location to the lower (L) and upper (U) anchors are combined to make a magnitude estimate in the interval [0, 1], and this varies with changes in the activity of the Poisson accumulators. Panel C shows how the magnitude estimate is transformed into  $N$  response strengths by the mapping process. In the example, a magnitude estimate of 0.25 (corresponding to Stimulus 2) is transformed into  $N = 6$  response strengths, shown by the heights of the six lines above  $x = .25$ . Panel D shows how the ballistic accumulator stage makes a response decision. The response strengths from the mapping stage drive ballistic accumulators, with the first one to reach a threshold determining the choice. Between trials, the activity in the ballistic accumulators decreases by passive leakage.

perimental training and initial trials, the participant might estimate that the stimuli range is between 50 dB and 70 dB, and so the context becomes the range between these values. Importantly, the selective attention stage does not assume that the psychophysical

representations are addressable—that is, there are no numbers associated with the psychophysical locations. Instead, we explicitly model the process by which participants estimate stimulus magnitudes without having access to numerical tags. A finite set of

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

leaky accumulators is put in a one-to-one ordered correspondence with the psychophysical dimension. Presentation of a test stimulus results in two effects. First, a corresponding location on the topographic psychophysical representation becomes active (i.e., is selected). Second, the accumulator that corresponds to the psychophysical representation of the stimulus is selected. The selection of this accumulator allows the participant to identify its ordered location within the array of units. That is, the observer can identify which accumulators are below, or above, the one selected by the presented stimulus.

To aid in understanding how the selective attention process helps to determine magnitude information from an ordered set of accumulators, consider an analogy. Imagine a very long row of lamps on the wall of a room. A contiguous range of the lamps are lit green but with flickering intensities. The lit lamps represent the range of stimuli in an experiment, and the intensity of each lamp corresponds to the current level of activation of each accumulator. Now, suppose that the only method an observer has of measuring the activity of the lamps is to gauge the total intensity of portions of the array but not the absolute position of any single lamp. Further suppose that, if any particular lamp changes to red, an observer can select it and use this location to partition the array into those lamps above and those below. With only these abilities, the observer is able to estimate (relative) magnitudes using the total intensity of the green lamps below the selected red lamp and the total intensity of the green lamps above the selected red lamp. Each of these sums will be large or small depending upon the position of the currently lit red lamp, so they carry relative magnitude information. Finally, as with a flashlight shining on a surface, the average total intensity of the lit lamps is fixed, independent of the range of lit lamps.

Suppose, for example, that presentation of a Number 2 stimulus (a 52-dB tone in a set of range 50–70 dB) corresponds to a lamp near the lower end of the green range turning red. The separation between the lower end of the row of green lamps and the red lamp is small, so the sum of the intensity of the lamps between the two will also be small. However, the estimate of magnitude is noisy because the lamps are flickering, and so, the sums of their intensity will also vary. Other measurements are also possible—for example, if two lamps were to change color, the sum of the intensity of the lamps between them could be estimated in a similar way.

We call the first and last active accumulators on the attention dimension *anchors*; they correspond to stimulus magnitudes that span the range used in the experiment. The anchor positions are assumed to be under the direct control of the participant. The entire range of accumulators between the anchors is kept active by continual attention during an experiment, analogous to the process that keeps the lamps flickering. Activation is modeled as a Poisson process of mean rate  $\lambda$  events per trial, and so we sometimes refer to the accumulators as *Poisson accumulators*. Without loss of generality, we assume each Poisson event increases the activity of a randomly selected accumulator by a unit amount. This activation is combined with a passive decay process: In the absence of attention, each accumulator's activity decreases by a factor of  $\alpha$  over the course of a trial.<sup>6</sup> The combination of Poisson activation and passive decay results in each accumulator having an activation value that varies from trial to trial (Marley & Cook, 1984). The average total activity in all accumulators is set by the balance between the attention and decay rates, namely,  $\eta = \lambda/(1 - \alpha)$ .

This average total activity is the major determinant of the overall accuracy of responses in SAMBA.

Stimulus magnitudes are estimated by summing activity in subranges of accumulators, in particular, the total activity from the upper anchor (U) and from the lower anchor (L) to the current stimulus. Taking the sums between the current stimulus and each of the anchors produces two estimates of stimulus magnitude, one relative to the lower anchor ( $\Sigma_L$ ) and one relative to the upper anchor ( $\Sigma_U$ ). These are combined into a single magnitude estimate by the ratio  $\Sigma_L/(\Sigma_L + \Sigma_U)$ , which is naturally constrained to be between zero and one. As shown in Figure 1B, this magnitude estimate varies from trial to trial, even if exactly the same stimulus is repeated, because the activities in the Poisson accumulators vary as a result of the attention process. Importantly for absolute identification, the variability of the magnitude estimate is largest in the center of the range, causing SAMBA to predict a bow effect in response accuracy and sensitivity. This projection of psychophysical stimulus representations for all stimulus modalities onto a common bounded interval is supported by research on magnitude estimation and cross-modality matching by Krantz (1972), Teghtsoonian (1973), and Teghtsoonian and Teghtsoonian (1978, 1997).

The observer accumulates magnitude estimates  $\Sigma_L/(\Sigma_L + \Sigma_U)$  from each trial and stores an average magnitude estimate for each stimulus in a long-term memory. We assume that this long-term memory becomes stable relatively quickly, particularly when accurate feedback is provided. Since we aggregate data over many trials, we approximate the learning and memory process by the assumption that a participant has available an accurate memory for the average magnitude estimate corresponding to each stimulus. These memories, combined with the end anchors, constitute the set of referents stored in long-term memory that underpin SAMBA's absolute process, similar to the anchor values in Petrov and Anderson's (2005) ANCHOR model or the cutpoints in Stewart et al.'s (2005) RJM. The learning of referents is an important issue (see, e.g., Petrov & Anderson, 2005) and is an area where SAMBA could be extended in future. When discussing a false-feedback experiment below, we explore a simple beginning to a referent-learning mechanism in SAMBA.

The magnitude estimate produced by the selective attention stage must be transformed into  $N$  response strengths to provide inputs for the ballistic accumulator stage. The transformation is made by Lacouture and Marley's (1995, 2004) mapping process, illustrated in Figure 1C and presented in more detail below. The mapping is error free relative to its referents: For example, if the estimate of stimulus magnitude is closest to the referent for Stimulus 2, then the mapping will provide a response strength that is largest for Response 2 (strengths for Responses 1 and 3 will be next, and so on).

On each trial, the outputs of the mapping stage are analogous to a tuning curve over the possible responses, with the response strengths contingent on how well the current input matches the long-term referent for each stimulus. In contrast to the multiple

<sup>6</sup> The parameters  $\lambda$ ,  $\alpha$ , and  $\eta$  are similar to Marley and Cook's (1984) parameters of the same names but differ because we use discrete time (trials) and they used continuous time. If  $T$  is the duration of a trial, then our  $\lambda$  corresponds to their  $\lambda T$ , our  $\alpha$  corresponds to their  $e^{-\alpha T}$ , and our  $\eta$  is the reciprocal of theirs.

parameters of other tuning curve models, the positions, widths, and shapes of the tuning curves produced by the mapping stage are determined entirely by the values of the referents held in long-term memory. The fixed form of Lacouture and Marley's (1995, 2004) mapping provides significant constraint, greatly limiting SAMBA's flexibility for fitting choice and RT data. For instance, an alternative approach, explored by Karpiuk et al. (1997), links the output of the selective attention stage with the input to the decision stage via parameterized tuning curves, given by Link's (1992) wave theory. This framework has great flexibility to adjust decision cutpoints, whereas the mapping solution has a limited ability to adjust such cutpoints. Our position is that, until clearly required to fit data, the more constrained model is preferable.

To illustrate how the mapping operates and also how it naturally handles unevenly spaced stimuli, consider an experiment from Lockhead and Hinson (1986). In one part of this experiment, the stimuli were tones of intensity 58 dB, 60 dB, and 66 dB. Suppose that a participant placed the lower anchor 2 dB below the lowest stimulus (i.e., at  $L = 56$  dB) and 4 dB above the highest (i.e., at  $U = 70$  dB). With this setup, regardless of any parameter settings, the selective attention stage would produce stimulus magnitude estimates for the three stimuli with average values of  $\{1/7, 2/7, 5/7\}$ . Note that these magnitude estimates naturally reflect the unequal spacing of the stimuli—3 times the spacing between the upper two stimuli than between the lower two stimuli—and this property holds regardless of the locations that the observer selects for the anchor values ( $L$  and  $U$ ), provided the stimulus locations lie between them.

Continuing this example, the mapping transforms a magnitude estimate into three response strengths, one for each of the three possible responses (1, 2, or 3). The computations of the mapping stage are specified entirely by the long-term average stimulus magnitude estimates. A magnitude estimate, say,  $z$ , is linearly transformed into a response strength for each and every response  $j = 1 \dots N$  according to  $(2Y_j - 1)z - Y_j^2 \pm 1$ , where  $Y_j$  is the average magnitude estimate for the  $j$ th stimulus.<sup>7</sup> For the Lockhead and Hinson (1986) example, suppose the 60-dB stimulus was presented. On this particular trial, the selective attention stage might produce a magnitude estimate of .3, which is quite close to the long-term average value of  $2/7$  for this stimulus. The mapping stage transforms this value into three response strengths: the strength for Response 1 is  $(2 \times 1/7 - 1) \times 0.3 - (1/7)^2 + 1 = 0.765$ ; for Response 2, it is  $(2 \times 2/7 - 1) \times 0.3 - (2/7)^2 + 1 = 0.790$ ; and for Response 3, it is  $(2 \times 5/7 - 1) \times 0.3 - (5/7)^2 + 1 = 0.618$ . Notice that the strength is greatest for the correct response (Response 2).

Figure 1C illustrates the mapping solution for a more standard experiment, with six evenly spaced stimuli and six lines corresponding to the response strengths for the six possible responses. In the example, Stimulus 2 is presented so the selective attention stage will—on average—produce a magnitude estimate of .25, shown by the vertical arrow in Figure 1C. From this magnitude estimate, the mapping produces six response strengths, shown by the heights of the six lines immediately above the arrow. As required for a correct response, the strength corresponding to Response 2 is greatest, with Responses 1 and 3 next, Response 4 after that, and so on.

The mapping stage brings many useful properties to SAMBA. First, it provides a powerful way to simplify the model and greatly

reduce the number of free parameters required to produce RT predictions. Models of RT distributions for  $N$  alternative tasks generally have some multiple of  $N^2$  parameters (e.g., see Busemeyer & Townsend, 1992, 1993). These parameters specify response strengths or drift rates for each response, contingent on each stimulus. Lacouture and Marley's (1995, 2004) mapping replaces this entire set of  $O(N^2)$  parameters with a process for producing drift rates given a stimulus magnitude estimate as input.

The outputs of the mapping component provide the input to the final response selection or decision stage of SAMBA, which is based on Brown and Heathcote's (2005) ballistic accumulation model. The ballistic accumulator model brings three important properties to the modeling absolute of identification data. First, the use of an accumulator process allows SAMBA to make very detailed predictions about RT. Second, the process of passive leakage in ballistic accumulator activities provides a natural explanation of very short-term sequential effects without the need to posit extra processes. Third, the ballistic accumulator includes competition—also called lateral inhibition—between accumulators, and this naturally means that responses become slower as the number of alternatives (set size,  $N$ ) increases.

The decision stage associates each of the  $N$  responses with a competitive ballistic accumulator (Brown & Heathcote, 2005; see also Usher & McClelland, 2001). The inputs to the decision accumulators are the outputs of the mapping stage, with a common (single) sample of zero-mean Gaussian noise added to every response strength, with the standard deviation ( $\sigma_M$ ) estimated from the data. The noise is analogous to the drift variance included in almost all models of choice RT, as first introduced by Ratcliff (1978; his parameter  $\eta$ ). In our applications, the estimated standard deviation of the noise ( $\sigma_M$ ) was sufficiently small so that at least one decision unit always received a positive input. As the same sample is added to all inputs,  $\sigma_M$  models stimulus-independent variability due to global factors such as fluctuations in arousal. The value of  $\sigma_M$  has little influence on accuracy because it does not alter differences between inputs. Hence,  $\sigma_M$  has a selective influence on variability in RT, which is also determined by sequential effects in the decision stage discussed in detail later.<sup>8</sup>

Figure 1D illustrates an example decision process. Each of the  $N$  response accumulators begins the decision stage with a starting activation level determined by previous inputs. The starting levels are different for each accumulator, and the activation levels increase deterministically at rates dictated by the response strengths from the mapping stage and competition among accumulators. Activation  $x_j$  in accumulator  $j = 1 \dots N$  changes according to a linear, first-order system of differential equations:  $x_j'(t) = I_j$

<sup>7</sup> Lacouture and Marley (1995) motivated this parameter-free solution by requiring a solution that (a) makes use of all of the bounded input and bounded output range in every experiment and (b) has cutpoints that are midway between the mean activities for adjacent stimuli.

<sup>8</sup> A reviewer questioned whether the  $\sigma_M$  parameter might be replaced by within-trial variability in the decision stage or perhaps another source of variability in the magnitude estimates. Other sources of variability affect error rates, which  $\sigma_M$  does not, and would thus increase model flexibility by providing another mechanism for modeling error responses. Our decision not to include this extra flexibility is based on parsimony, as the data we examined did not seem to require it. However, we have no grounds to rule out other types of variability.

–  $\beta \sum_{p \neq j} x_p(t)$ . Here,  $I_j$  represents the response strength from the  $j$ th unit of the mapping stage plus the Gaussian noise sample with standard deviation  $\sigma_M$ . The parameter  $\beta > 0$  represents lateral inhibition and causes the increase in the activation to be nonlinear. The system of coupled differential equations describing accumulation during the decision stage can be solved analytically by matrix algebra for any response set size  $N$  (see Brown & Heathcote, 2007).

A response is chosen corresponding to the first accumulator to reach a threshold ( $C$ , the same value for all accumulators), and the RT is given by the time taken to reach that criterion plus a constant time for nondecision processes,  $t_0$ . This system of accumulators ensures that a finite response threshold will always be reached in a finite time. This occurs because at least one input is positive and because we use a simplified version of the ballistic accumulator model with no passive leakage within a trial (see Brown & Heathcote, 2005, for further details). The example activation trajectories in Figure 1D correspond to the mapping example illustrated in Figure 1C. The trajectory corresponding to the correct response (Response 2) increases fastest and reaches the threshold first, so a correct response would be made in this example with an RT of just over 1.5 s. If the response threshold were set lower, say at  $C = 20$  instead of  $C = 25$ , the RT would be faster, around 0.9 s, but an error would be made: The model would give Response 3 instead of the correct response (Response 2). The error occurs because the ballistic accumulator for Response 3 begins the decision stage with an advantage over the accumulator for Response 2, and it takes some time for this advantage to be overcome.

The decision stage of SAMBA is more constrained than Brown and Heathcote's (2005) ballistic accumulator model and most other models of two-choice RT (see Ratcliff & Smith, 2004, for a recent summary) in its assumptions about the starting points of the evidence accumulation processes. These models assume that accumulation for each response unit starts from a random value. In SAMBA, the accumulator start points are completely determined by passive decay from their values at the end of the previous trial (also see Laming, 1968). This mechanism not only specifies previously unspecified details of the genesis of start point variability in theories of choice RT but also enables the model to explain short-term sequential phenomena in absolute identification, including assimilation and response repetition effects. For typical parameter values, about one quarter of incipient choices generated by the selective attention and mapping stages of SAMBA are changed by the ballistic accumulators, due to differences in the starting points of the evidence accumulation processes left over from the previous trial.

Brown and Heathcote's (2005) ballistic accumulator model, used in SAMBA's decision stage, has been simplified further by Brown and Heathcote (in press) by omitting the lateral inhibition term (i.e., by setting  $\beta = 0$ ), making a linear ballistic accumulator (LBA) model. The LBA model has computational and analytic advantages over the ballistic accumulator model, but nevertheless, we choose not to use it in SAMBA. The lateral inhibition element in the ballistic accumulator makes SAMBA predict increased RT with increased set size ( $N$ ), without the need for any parameter changes with set size. As discussed by Brown and Heathcote (in press), if LBA were used in SAMBA, then an extra assumption would be required to fit the increase in RT with set size. For example, the LBA would predict increased RT with set size if the

outputs of the bow mapping stage decreased with increased set size (e.g., by fixing the sum of the response strengths across set size).

### *An Absolute Account of Sequential Effects on Errors*

SAMBA makes strong predictions about sequential effects in errors and RT because it unambiguously attributes contrast effects to processes producing a context-dependent stimulus representation, as well as repetition and assimilation effects to the processes producing a response. As in other models of choice RT, Brown and Heathcote (2005) assumed that the start points for the ballistic accumulators ( $x_0$ ) were independent random samples from a common uniform distribution. SAMBA replaces this assumption with a deterministic mechanism based on passive leakage. This passive leakage is illustrated in Figure 1D, where dotted lines show how accumulated evidence decays during nondecision and between-trial times. Computationally, after each decision is completed, the lateral inhibition and stimulus input processes ( $I_j - \beta \sum_{p \neq j} x_p$ ) are replaced by a passive decay process. Decay returns each accumulator's activation exponentially back toward zero at a constant rate:  $x'_j(t) = -\gamma x_j(t)$ . Therefore, after a constant intertrial interval (ITI), each accumulator begins the next decision process with activation  $Dz$ , where  $z$  was the activation level for that accumulator when the previous response was made and  $D$  is given by  $\exp(-\gamma \cdot \text{ITI})$ . In our model fits, we make the approximation that ITI is constant throughout the experiment, and estimate only the value  $D$ . At the beginning of each decision process, the accumulators corresponding to the previous response and those for nearby responses have an advantage, as illustrated in Figure 1D. This results in response assimilation and in RT and accuracy advantages for repeated stimuli.

Contrast effects arise through reallocation of activity in the selective attention stage. In particular, we assume that the participant has some control over the Poisson process and that this control is used to preferentially attend to locations selected by recently presented stimuli. Reallocation is modeled by a probability ( $M$ ) that, each time an accumulator is to be incremented by the Poisson process, the increment is directed to the accumulator selected by the most recently presented stimulus. Consequently, the activity of accumulators selected by recently presented stimuli is increased, which in turn makes magnitude estimates near those values larger. These locally larger values cause a contrast effect that decreases with time as activity in the accumulators decays.

To keep the model as simple as possible, we made the default assumption that participants redirect activity only toward the accumulator selected by the most recently presented stimulus. This assumption causes contrast to have its greatest magnitude at lag = 2, after the (stronger) assimilation effect has passed. This pattern matches most absolute identification data, but for data from Lacouture (1997), we observed the peak contrast effect later, at lag = 3 or 4. We modeled those data by assuming that the observer persists in redirecting activity for  $K$  trials. As before, each Poisson event has a probability  $M$  of being redirected, but the redirection is to one of the locations selected by the  $K$  most recently viewed stimuli.

### *Integrating Absolute and Relative Judgment*

Research into absolute identification has become focused on the distinction between absolute versus relative interpretations, for



both empirical phenomena and theoretical accounts. We think this distinction is not as useful nor as clear cut as others believe. Even the most relative models (such as Stewart et al.'s, 2005, RJM) employ absolute knowledge about the global nature of the experiment: For example, when the RJM is used to model unequally spaced stimuli, such as in Lockhead and Hinson (1986), the (global, absolute) scaled magnitudes of the stimulus differences must be taken into account in setting the spacing of cutpoints, ensuring that the global stimulus setup is captured in the model.

Our description of SAMBA has, so far, been entirely in terms of absolute processes in the sense that the present stimulus magnitude is evaluated against long-term referents such as the anchors. However, Rouder et al. (2004) observed an accuracy bonus for repeated stimuli and for stimuli that are very similar to the preceding stimulus that cannot be accommodated within SAMBA without including a partially relative process. Below, we extend Rouder et al.'s analysis and demonstrate that the accuracy bonus also translates to an RT bonus. The absolute version of SAMBA we have outlined so far accommodates this phenomenon in a qualitative sense; it predicts that responses to repeated stimuli are faster and more accurate. However, it fails to quantitatively fit these data, as the predicted accuracy and RT bonuses are smaller than observed.

The RJM also matches the accuracy effects qualitatively but fails quantitatively by predicting too large an accuracy bonus for several data sets (see Brown et al., 2007). Stewart (2007) showed how the RJM could provide a better quantitative fit when modified to be slightly more absolute, by extending the model's memory so that the stimulus two trials back is, on certain theoretically specified trials, used as the basis for judgments in place of the memory for the stimulus one trial back. Similarly, we have found that SAMBA can provide a better quantitative fit when modified to be slightly more relative—confirming that Rouder et al.'s (2004) analysis and our extension suggest the need for both absolute and relative processes in models such as RJM and SAMBA.

Working from the assumption that absolute identification must include both relative and absolute processes, we have developed SAMBA to provide the first theoretical account to integrate these processes in a consistent manner. The key to the integration is the use of anchors in SAMBA. In the absolute version of SAMBA, the anchors bracket the smallest and largest stimuli that are important in the experimental context, and incoming stimuli are judged against these anchors via the ratio  $\Sigma_I/(\Sigma_L + \Sigma_U)$ . This ratio estimates the magnitude of the current stimulus within the range defined by the two anchors, [L, U]. On the other hand, in the relative version of SAMBA, a stimulus is judged within a smaller interval, with one of the anchors replaced by the Poisson accumulator selected by the previous stimulus. If the previous stimulus is larger than the current one, the current stimulus is judged relative to the interval between the lower anchor and the previous stimulus; if the previous stimulus is smaller, the current stimulus is judged relative to the interval from the previous stimulus to the upper anchor. These assumptions are similar to assumptions made by RJM's relative judgment process, although there are also important differences (e.g., the RJM uses a separate mechanism when successive stimuli are judged to be equal).

To illustrate SAMBA's relative process, consider an example. Suppose Stimulus 4 was presented on the previous trial. If Stimulus 7 is presented next, it would be judged in the interval [4', U], where 4' indicates the Poisson accumulator associated with the

previous presentation of Stimulus 4. After being restricted to these subintervals, SAMBA works as before.<sup>9</sup> That is, the current stimulus magnitude is estimated by summing up the activation between the current stimulus and each end of the subinterval, that is, 4' and U, and these two sums are combined into the ratio of the form  $\Sigma_{4'}/(\Sigma_{4'} + \Sigma_U)$ . Then, the mapping stage operates in the usual manner to transform the magnitude estimate into a set of response strengths, except that the mapping is restricted to the responses commensurate with the subinterval. In particular, zero strengths are given to responses that are smaller than the previous response if the current stimulus is larger than the previous stimulus, and vice versa. Finally, the ballistic accumulator stage proceeds as normal to select a response, with an associated RT.

This mechanism is partially relative because it uses a memory of the previous stimulus, and it is partially absolute because it still requires one of the anchors from long-term memory and employs the mapping solution, which is based on long-term memories for stimulus magnitude estimates. In SAMBA, the relative judgment process is under strategic control, so that a participant must choose to use relative judgment. In fitting data from the comprehensive data set of Lacouture (1997), we obtain the best fit by assuming partial use of the relative mapping process. We model this with a parameter *P* indicating the proportion of trials and/or participants that use the relative process—there were insufficient data at the individual subject level to separate these interpretations. Only the analysis developed by Rouder et al. (2004) provides clear evidence for the use of the relative process through accuracy and RT bonuses for repeat and near-repeat stimuli (where a near-repeat stimulus is one that is close, in the rank order of the stimuli, to the stimulus presented on the previous trial). We model the data from Stewart et al.'s (2005) Experiment 1 with the simpler assumption that all participants used a relative process on all trials. All other analyses, including our fits to assimilation and contrast phenomena, do not depend on the relative process.

### Parameters of the Model

Table 1 shows SAMBA's 13 parameters, along with the data characteristics that they affect. A key feature of SAMBA is that the parameters' effects are defined by the architecture of the model and can be interpreted in terms of psychological processes. This provides strong predictions and allows for tests of selective influence—certain parameters must only be altered by particular experimental manipulations, and those manipulations must affect only the parameters in question. In this section, we outline some of these parameter constraints. The first stage of SAMBA, our simple psychophysical stimulus representation, has just one parameter ( $\sigma_p$ ). Each of the remaining stages is affected by two or more parameters, with one extra parameter ( $t_0$ ) to account for the sum of the times taken to complete processing in nondecision stages and to make a response after the decision is made.

<sup>9</sup> We have not directly modeled the process by which the observer decides whether the current magnitude estimate is smaller or larger than the previous one, but it would be trivial to do so. Since the numerical sizes of these magnitude estimates are available to the observer, a simple accumulator mechanism could instantiate our assumption of very fast and accurate larger versus smaller decisions.

Table 1  
Model Elements

Stage affected	Symbol	Description	Principal effect
Psychophysical	$\sigma_P$	Psychophysical noise	Overall accuracy for small stimulus range
Selective attention	$\eta = \lambda/(1 - \alpha)$	Accumulator update equation: $x_{n+1} = \alpha x_n + \text{Poisson}(\lambda)$ Ratio of the mean number of pulses in the Poisson process to the accumulator decay rate	Overall accuracy and bow in accuracy
	$\alpha$	Rate of decay for accumulators	Overall accuracy and contrast time course
	$M, K$	Mean proportion ( $M$ ) and duration ( $K$ ) of activity directed to prior stimulus location	Contrast
	$L, U$	Position of the lower and upper anchors	Size and symmetry of bow effects
Selective attention and mapping	$P$	Probability of using relative process	Sequential effects on accuracy and RT
Mapping	$\sigma_M$	Output for response $j$ given magnitude estimate $z$ : $(2Y - 1)z - Y_j^2 + 1 + N(0, \sigma_M)$ Standard deviation of noise added to outputs of the mapping	Variability of RT distributions
Decision		Activation change $x_j'(t) = \begin{cases} I_j - \beta \sum_{p \neq j} x_p(t) & \text{predecision} \\ -\gamma x_j(t) & \text{postdecision} \end{cases}$	
	$\beta$	Rate of lateral inhibition	Size of bow and set size effects in RT
	$C$	Decision criterion	Overall RT and accuracy
	$D$	Rate of decay of decision unit activation during intertrial time $D = \exp(-\gamma \cdot \text{ITI})$	Assimilation and shape of bow effect
Nondecision	$t_0$	Nondecision component of reaction times	Overall RT

Note. ITI = intertrial interval; RT = response time.

The first parameter ( $\sigma_P$ ) determines the standard deviation of noise in the psychophysical stimulus representation. It only has an appreciable effect on SAMBA's predictions when stimuli are close enough in magnitude to cause errors in comparative judgment (i.e., judgments about stimuli presented simultaneously or in a rapid sequence). With few exceptions, the experiments we model use adjacent stimuli that are sufficiently widely spaced so that we can fix  $\sigma_P$  at zero. Four of the next six selective attention stage parameters determine the distribution of Poisson activity in that stage, namely, the mean number of pulses in the Poisson process over one trial ( $\lambda$ ), the rate of decay of each accumulator ( $\alpha$ ), the proportion ( $M$ ) of activity directed toward the unit selected by the previous stimulus (or stimuli), and the duration in trial units ( $K$ ) of that direction. Activity is otherwise assumed to be distributed with equal probability across locations. The remaining two selective attention stage parameters, the positions of each end anchor ( $L$  and  $U$ ), determine the range of the accumulators and allow SAMBA to accommodate asymmetries in data (i.e., more accurate and slower responses to larger than smaller stimuli, or vice versa).

One parameter—the probability ( $P$ ) of using the relative judgment process—affects both the selective attention and mapping stages because the relative process replaces one of the anchor values and scales the output of the mapping stage. Another parameter—the standard deviation of noise ( $\sigma_M$ ) added to mapping outputs—affects only the mapping stage. Finally, there are three decision stage parameters that have the same value for all decision units: the evidence accumulation threshold ( $C$ ), the strength of competition between the decision units ( $\beta$ ), and the decay rate ( $D$ ) of decision unit activation.

### Benchmark Phenomena

There are a plethora of benchmark phenomena in absolute identification concerned both with the choices made by participants and with the distribution of the times (RTs) to make such choices. The parameters we use for the benchmark phenomena are based on the parameter values estimated from Lacouture's comprehensive data set (see Table 2). Only one or two parameters needed to be adjusted from this baseline to fit each benchmark experiment. The  $\sigma_P$  (psychophysical noise) parameter was fixed at zero for all fits, except those for Braida and Durlach's (1972) study and Stewart et al.'s (2005) Experiment 1, since each of those designs contained some conditions with closely spaced stimuli. Where sequential effects were not at issue, the local judgment process was not used (i.e.,  $P = 0$ ), and only one parameter was estimated for the Poisson process in the selective attention stage: the ratio  $\eta$ , which largely determines overall accuracy. In those cases, effectively only eight parameters were required to fit the data, and where only choices (not RT) were considered, this number dropped to six (without  $t_0$  and  $\sigma_M$ ). Hence, the model fits are parsimonious and use parameter values that are consistent across several data sets from different paradigms.

Stewart et al. (2005) provided a comprehensive review of benchmark phenomena concerning choices, but not RT, and demonstrated that their RJM accommodates the former. SAMBA accommodates all of the benchmark phenomena listed by Stewart et al., and others related to RT, as illustrated by our fits to comprehensive data sets below. Absolute models, from which SAMBA derives, have often been unable to account for sequential phenom-

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

Table 2

Parameter Estimates for Lacouture (1997), Experiment 1 From Stewart, Brown, and Chater (2005), and Lacouture and Marley (2004)

Data set	Selective attention					Mapping				Decision			
	$\sigma_P$	$\alpha$	$\lambda^b$	$M$	$K$	L	U	$P$	$\sigma_M$	$\beta$	$D$	$C$	$t_0^f$
Lacouture (1997)	—	.75	6 ( $\eta = 24$ )	.14	4	91 <sup>c</sup>	520 <sup>c</sup>	.525	.22	.0307 <sup>e</sup>	.07	878	.265
Stewart, Brown, and Chater (2005)	1.9 <sup>a</sup>	.92	0.87 ( $\eta = 7$ )	.36	3	10 <sup>d</sup>		(1)	—	—	.11	—	—
Lacouture and Marley (2004)	—	.80	19 ( $\eta = 95$ )	.06	3	86 <sup>c</sup>	384 <sup>c</sup>	—	.10	.08 <sup>e</sup>	.02	535	.223

Note. Dashes indicate parameters that were not estimated.

<sup>a</sup> Percentage of stimulus frequency. <sup>b</sup> Events per trial. <sup>c</sup> Pixels. <sup>d</sup> Percentage of stimulus range. <sup>e</sup> Per second. <sup>f</sup> Seconds.

ena, such as assimilation and contrast, so SAMBA's accounts of those phenomena are covered in detail below. The following section on critical tests addresses data patterns that have previously been thought to rule out absolute models.

### Stimulus Spacing Effects

We fit SAMBA to two data sets showing stimulus spacing effects, Braida and Durlach (1972) and Lockhead and Hinson (1986). SAMBA accommodates the effects of unequal spacing as a natural consequence of its architecture, with the physical values of the unequal stimulus spacings providing direct, parameter-free constraints on the psychophysical front end. Braida and Durlach examined the effect of changing the physical range of the stimuli. They performed eight absolute identification experiments, each with 10 pure tones equally spaced in intensity (dB), with a different stimulus range and, hence, spacing, in each experiment. The first experiment used a very small range, with the smallest and largest tones differing by only 2.3 dB, so adjacent stimuli were sufficiently closely spaced to bring psychophysical noise into play. The remaining experiments steadily increased the range up to 54 dB (see Figure 2). Two important effects were observed in these data. First, the overall accuracy, as measured by the amount of information transmitted from stimulus to response, increased as the range increased but quickly reached an asymptote. Second, stimulus sensitivity, as measured by  $d'$  per bel ( $B^{-1}$ ), showed the standard bow effect for large stimulus ranges but almost no bow effect for small ranges.

The upper panel of Figure 2 shows how information transmitted increases from almost zero at the smallest range to an asymptote of about 1.9 bits at the largest ranges. The lower panels show how sensitivity per bel decreases with increasing range while simultaneously the depth of the bow effect increases. As illustrated by the dashed lines, SAMBA provides an accurate account of both phenomena. As for all the benchmark phenomena, the model's fits were parsimonious, adjusting only some of a fixed set of reference parameter values used in fitting Lacouture's (1997) data (see Table 2). For Braida and Durlach's (1972) data, we omitted the local process (i.e., set  $P = 0$ ), increased the selective attention ratio to  $\eta = 30$ , and included a psychophysical noise parameter,  $\sigma_P = 0.96$  dB. The ratio  $\eta$  was increased to match the overall accuracy level of Braida and Durlach's participants, and the psychophysical noise parameter was changed as it was fixed at zero for Lacouture's

(large-range) data. We are reassured about the interpretability of model parameters by comparison with the estimates given by Marley and Cook (1984). Using the same data set but using an asymptotic approximation of the selective attention model and decision cutpoints, Marley and Cook estimated  $\sigma_P = 0.9$  dB and  $\eta = 26$ .

Importantly, no parameters were adjusted between the various stimulus ranges to achieve these fits. Instead, SAMBA captures the effects of increased stimulus range solely through the action of psychophysical noise ( $\sigma_P$ ). When stimulus separations are small, psychophysical noise causes confusion between adjacent stimuli, decreasing performance in the small-range conditions. As the stimulus range increases, the separation between adjacent stimuli rapidly grows larger than the imprecision introduced by psychophysical noise. Once the stimulus separation is greater than  $\sim 3$  dB, psychophysical variability (with a standard deviation of 0.96 dB) becomes unimportant, and performance asymptotes. The reason that SAMBA predicts almost flat bow effects for small stimulus ranges is that the very high error rates for small stimulus separations is almost exclusively due to psychophysical variability, which is constant across the stimulus range.

A similar limit on absolute identification performance relates to increases in the number of stimuli ( $N$ ) rather than stimulus range. Pollack (1952) and Garner (1953) measured performance in terms of the amount of transmitted information (in bits) when the number of stimuli increased. Their most important observation was a limit—no more than about 2.8 bits of information were transmitted through responses no matter how large the stimulus set size became. This limit is fundamental to absolute identification and is incorporated at an appropriately fundamental level in SAMBA. The selective attention process has a limited capacity (the total average activity in all accumulators,  $\eta$ ). This capacity is independent of manipulations such as stimulus set size ( $N$ ) or stimulus range (as in Braida & Durlach, 1972), and it determines the variability of the stimulus magnitude estimates.

Lockhead and Hinson (1986) performed another benchmark experiment that manipulated stimulus spacing using three tones that differed in intensity. In the equally spaced condition, adjacent tones were separated by 2 dB (at 58 dB, 60 dB, and 62 dB). In this condition, confusion matrices (i.e., the probability of each response conditional on each stimulus) were typical of other absolute identification data sets (see Figure 3, middle panel). Lockhead and

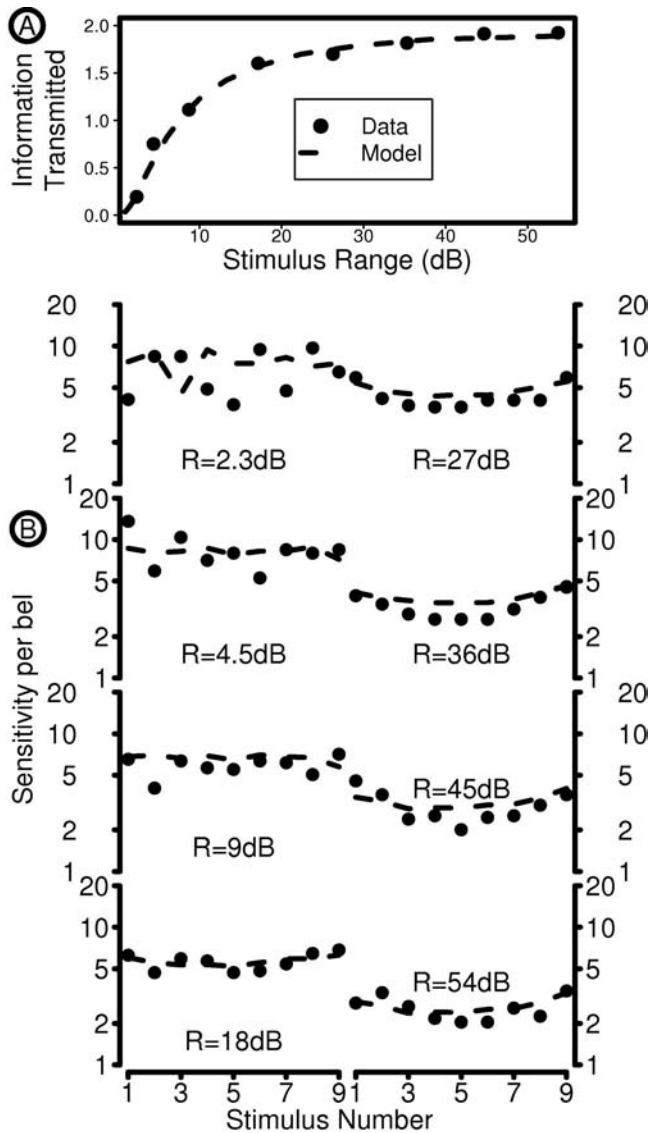


Figure 2. Data from Braida and Durlach's (1972) study (symbols) along with model fit (lines). Panel A shows how information transmitted increases to an asymptote as the stimulus range increases. Panel B shows how sensitivity per bel ( $B^{-1}$ ) decreases as range increases, with a corresponding increase in the depth of the bow effect.

Hinson created two other conditions by manipulating the spacing of the end stimuli. Compared with the equally spaced condition, in the low-spread condition, the lowest stimulus was made much lower (54 dB), and in the high-spread condition, the highest stimulus was made much higher (66 dB). The confusion matrices for these two conditions are shown in the left and right panels of Figure 3 (cf. Stewart et al., 2005, Figure 7). The important effect is that the unchanged stimuli (the upper two stimuli in the low-spread condition and the lower two stimuli in the high-spread condition) were more often confused in the low-spread and high-spread conditions than in the equally spaced condition. This poses a theoretical challenge since the relevant pair of stimuli do not differ physically between pairs of conditions, yet they are more

often confused when the third stimulus is far away than when it is near. The dashed lines in Figure 3 show that SAMBA can explain these data more parsimoniously than previous accounts—with no parameter changes between conditions. The different predictions in the three conditions arise solely from differences in the stimulus spacing, set directly by the experimental design.

Only one parameter was changed from our fits to Lacouture's (1997) data ( $\eta = 10.5$ ) to fit the overall accuracy of Lockhead and Hinson's (1986) participants—note that the local process was also omitted ( $P = 0$ ), although this makes little difference. SAMBA's differential predictions for the three conditions are a natural consequence of the geometry of the model's mapping solution. When the stimuli are equally spaced, the standard map applies, as described in Lacouture and Marley's (1995) original formulation. However, when the stimuli are unequally spaced, the section of the mapping around the relatively closely spaced stimuli becomes compressed as consequence of the mean activities in the selection attention stage being close, leading to poorer performance.

### Sequential Effects on Errors

The two most studied sequential effects in absolute identification are assimilation and contrast (see Holland & Lockhead, 1968; Lacouture, 1997; Mori & Ward, 1995; Ward & Lockhead, 1970). Assimilation and contrast are sequential effects concerned with the distribution of incorrect responses among the possible responses (as opposed to effects concerned with the overall number of incorrect responses, which we examine later). Figure 4A shows assimilation effects in detail for one benchmark data set, and Figures 4B and 4C show both assimilation and contrast for that data set and another benchmark data set.

Assimilation means that errors tend to be made toward rather than away from the previous stimulus. In Figure 4, this is shown using the average error, which is the average difference between the correct response and the actual response. For example, if the correct response is 3 and the subject responds 5, the error is 2. Assimilation is evident at Lag 1 (i.e.,  $X = 1$  in B and C)—that is, average errors are positive (respectively, negative) when the preceding stimulus is large (respectively, small). Contrast is evident in the opposite pattern for longer lags (i.e.,  $X = 2-8$  in B and C)—that is, the average errors are positive (respectively, negative) when the previous stimulus is small (respectively, large). In SAMBA assimilation is due to prior responses, through the starting point of the decision accumulators on the next trial, and contrast is due to prior stimuli, through attention allocated to the units selected by previous stimuli.

Creating a theory that simultaneously produces assimilation at short lags and contrast at longer lags is challenging: It must predict that responses are biased toward the previous response but away from responses prior to that one. SAMBA quantitatively describes assimilation and contrast quite well, as shown by the solid lines in Figure 4. Also, in agreement with data, SAMBA could never make the prediction that assimilation and contrast occur at the opposite time scales (i.e., lag = 1 for contrast, and lag > 1 for assimilation). In general, models that set parameters separately for assimilation and contrast could make such a prediction, which is problematic (Roberts & Pashler, 2000).

SAMBA cannot make this counterfactual prediction because it is constrained by the nature of its processing architecture. Assim-

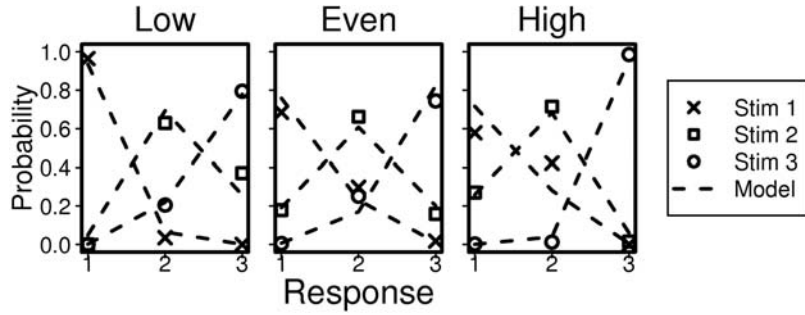


Figure 3. Data and fits for each of the three conditions in Lockhead and Hinson (1986). Each panel refers to a different experimental condition; from left to right, the lower stimulus is much lower (low spread), the stimuli are evenly spaced, and the upper stimulus is much higher (high spread). Each graph shows the probability of each response (abscissa) conditional on each stimulus (separate lines, see legend). Stim = stimulus.

ilation is naturally predicted by passive decay in the ballistic accumulators. Between each trial, the activation values of the accumulators decay smoothly back toward a baseline level. This means that the accumulator corresponding to the response made on the previous trial will begin the next trial with an advantage, and that advantage will also extend to nearby responses, as they typically have activations close to that of the winning response. The rate of decay in accumulator values during the ITI (Parameter  $D$ ) governs the size of assimilation effects, but the effects must always be assimilative, never contrastive. The model's competitive response selection stage also restricts these effects to the previous trial only, never to earlier trials.

Similarly, SAMBA predicts both the direction and the time course of contrast due to preferential treatment for recent stimuli in the selective attention process. The Poisson process that activates accumulators is biased toward incrementing the accumulator selected by previous stimuli, normally the most recent. The magnitude of the bias is set by Parameter  $M$ , representing the proportion of activity redirected this way, and the duration of the bias is set by Parameter  $K$ . The contrast mechanism directs extra activity to accumulators selected by recently presented stimuli, and this ac-

tivity causes the expansion of magnitude estimates that include those locations. For example, suppose the previous stimulus was number 2, and the current stimulus is larger, 3. The magnitude of the current stimulus is estimated by summing activity between Stimulus 3 and the upper and lower anchors (giving the values  $\Sigma_L$  and  $\Sigma_U$ ). The value  $\Sigma_L$  includes extra activity in the accumulator selected by the previous stimulus (Stimulus 2), and so, the estimated magnitude of the current stimulus will be larger than it otherwise might be. A parallel argument shows that if the current stimulus is smaller than the previous stimulus, it is judged to be smaller than it otherwise might be. The resulting effect is contrast: Stimuli are judged to be further away from recently seen stimuli, and this is the case whatever the relative locations of the current and previous stimuli. Such context effects make adaptive sense for a participant tracking stimulus distributions that vary over time (Petrov & Anderson, 2005; Ward & Lockhead, 1970).

For the fits in Figures 4A, 4B, and 4C, we began with parameter values estimated from Lacouture's (1997) data and removed the local process (i.e., set  $P = 0$ ). To fit Ward and Lockhead's (1970) data in Figures 4A and 4B, we changed one parameter that affects assimilation (decay in the decision stage,  $D = .2$ ), one parameter

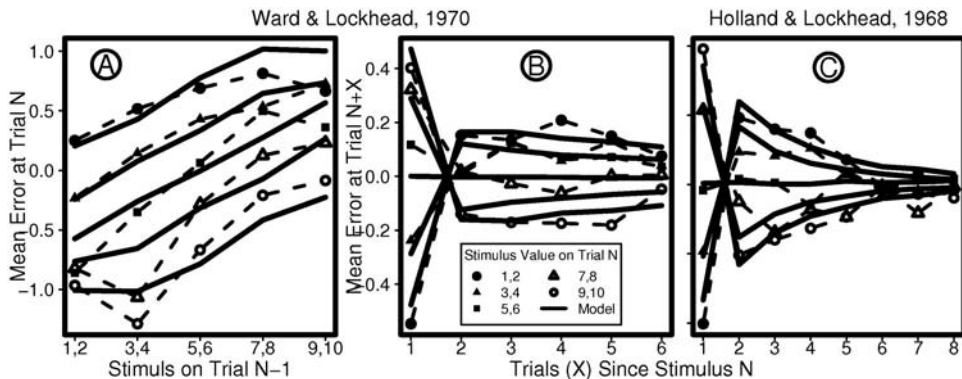


Figure 4. Panel A shows assimilation effects in data from Ward and Lockhead (1970): The average error on trial  $N$  was positive when the stimulus on the previous trial ( $N - 1$ ) was large, and vice versa. Panels B and C show both assimilation (at  $X = 1$ ) and contrast (at  $X > 1$ ) in Ward and Lockhead's and Holland and Lockhead's (1968) data. When  $X = 1$ , assimilation is shown by negative average errors when the previous stimulus was small (filled symbols) and positive average errors when the previous stimulus was large (unfilled symbols). The opposite pattern at longer lags ( $X > 1$ ) is the contrast effect. Solid lines are predictions from SAMBA.

that affects the magnitude of contrast effects ( $M = .75$ ), and two that govern its time course ( $\alpha = .9$ , and  $K = 2$ ). To fit Holland and Lockhead's (1968) data in Figure 4C, we changed just the one parameter that affects assimilation ( $D = .17$ ).

*Response Time Distributions*

Even when RTs are collected in absolute identification experiments, they are rarely subjected to the detailed analysis given to response choices. Previous research has identified several effects in mean RTs analogous to effects in choice. These include bow effects, in which responses to extreme stimuli are faster than those to middle stimuli (e.g., Lacouture, 1997; Lacouture & Marley, 1995, 2004), and set size effects, where RTs slow down as set sizes increase (e.g., Kent & Lamberts, 2005; Lacouture & Marley, 1995, 2004). Sequential effects on mean RTs have also been observed due to response repetition (Petrov & Anderson, 2005) and assimilation (Lacouture, 1997). We illustrate SAMBA's ability to accommodate these phenomena in mean RT later, in our fits to Lacouture's (1997) data.

An even more stringent model test is provided by fitting full RT distributions. This has rarely been attempted in absolute identification, with a few exceptions, notably Kent and Lamberts (2005) and Lacouture and Marley (2004). We have taken data from both of those studies and fit them with SAMBA. We present the fits to Kent and Lamberts's data and Lacouture and Marley's data here and those to Lacouture's (1997) data in our fits to comprehensive data sets below. Our analyses of the data from Kent and Lamberts's Experiment 1 and Lacouture and Marley are particularly important, as we fit full RT distributions for individual subjects, without averaging. For Lacouture and Marley's data, we go one step further and separately analyze data and model predictions for

the RT distributions of both correct and incorrect responses (data for incorrect responses were not available for Kent & Lamberts, 2005).

In their Experiment 1, Kent and Lamberts (2005) analyzed full RT distributions for 3 individual subjects, which we summarize using quantiles. For each of the 30 distributions (3 subjects  $\times$  10 stimuli), we calculated the 10%, 30%, 50%, 70%, and 90% quantiles, that is, the RT below which 10%, 30%, 50% (i.e., the median), 70%, and 90% of the data fall. These quantiles are shown along the bottom row of Figure 5, using three panels, one for each participant. The x-axis measures stimulus magnitude (from 1–10), and the five solid lines on each plot show quantiles calculated from the data. The upper and middle rows of Figure 5 show response accuracy and mean RT, respectively, also as functions of stimulus magnitude. The data show several standard effects. First, there are clear bow effects, where responses to middle stimuli are slower and less accurate than those to edge stimuli. For the RT distributions, these bow effects are greatest in the slow tails (the 90% quantile). The RT distributions are also positively skewed, with greater distances between the 70% and 90% quantiles than between the 10% and 30% quantiles. SAMBA captures all of these effects very well and provides a good quantitative fit to the data. The parameters used to fit the model are shown in Table 3. The RT quantile and accuracy data allowed us to estimate the decision stage parameters, the overall accuracy parameter ( $\eta$ ), and the anchor parameters (L and U). The published data contained no information on sequential effects, so we kept those parameters fixed at values estimated from Lacouture's (1997) data (see Table 2) and removed the local process ( $P = 0$ ).

Kent and Lamberts (2005) were unable to fit their model (the extended generalized context model for RTs, called ECGM-RT)

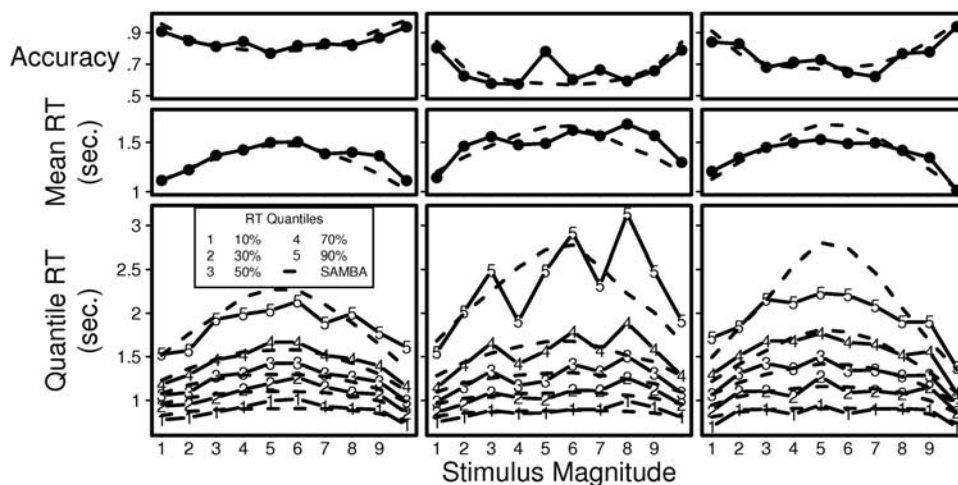


Figure 5. Top panels show accuracy, middle panels show mean response time (RT), and bottom panels show RT quantiles for correct responses for each of 3 participants (columns) from Kent and Lamberts (2005, Experiment 1). Solid lines correspond to data and dashed lines to predictions from SAMBA. The five numbered lines in the lower panels correspond to the 10%, 30%, 50% (median), 70%, and 90% quantiles of the RT distributions for the 10 stimulus magnitudes. The reader may note that our mean RT scales are more compressed than those in Kent and Lamberts's figures. This compression was required to show the very longest RT quantiles in the figure—Kent and Lamberts did not plot the tails of the distributions in their bow effect plots. sec. = seconds.

Table 3  
Parameter Estimates for Kent and Lamberts (2005, Experiment 1)

Subject	Selective attention ( $\eta$ )	Mapping			Decision			$t_0^c$
		L <sup>a</sup>	U <sup>a</sup>	$\sigma_M$	$\beta^b$	D	C	
1	75.0	80	330	.227	.053	.0044	806	.076
2	33.3	80	372	.298	.067	.021	730	.175
3	46.7	80	330	.231	.094	.010	772	.051

<sup>a</sup> Pixels. <sup>b</sup> Per second. <sup>c</sup> Seconds.

directly to RT distribution data because “although it might be possible in principle to estimate the properties of the residual-time distribution, the number of simulated trials needed to produce consistent estimates proved prohibitively large.” (p. 297). Instead they fit only mean RT and accuracy. Kent and Lamberts then used representative parameter values from these fits to generate illustrative RT distribution predictions. The predictions had qualitative trends that matched their RT data, although they were not intended to provide close quantitative fits of the sort we provide in Figure 5. Kent and Lamberts also did not report any data or model predictions for RT distributions associated with incorrect responses. Hence, a complete comparison of SAMBA and ECGM-RT will have to await the development of suitable estimation techniques for EGGM-RT. However, a comparison of the top two rows of Figure 5 with the corresponding data in Figure 1 of Kent and Lamberts shows that SAMBA’s fits just to mean RT and accuracy are comparable to those of the ECGM-RT, even though we did not directly optimize on those quantities. For comparison against possible future models for Kent and Lamberts’s data, we note that the chi-square values for SAMBA’s fits, summed across the 10 stimuli

separately for each participant, are 291, 235, and 397 for the left, middle, and right panels of Figure 5, respectively. These are similar to those obtained by fits of other RT models to data from binary choice tasks (e.g., Ratcliff, Gomez, & McKoon, 2004; Ratcliff & Rouder, 1998). These chi-square values were calculated in the usual manner for RT fits, on the basis of quantiles estimated from the data, making them inappropriate for comparison with theoretical chi-square distributions.

We were able to access the complete sequence of raw data for the single participant in Lacouture and Marley’s (2004) Experiment 2. We analyzed the data from the standard absolute identification section of that experiment, in which manual responses were used. We fit both the choice and RT data comprehensively with SAMBA, providing accurate fits of all benchmark phenomena, including bow effects and sequential effects. However, for brevity, we present only the analyses of RT distributions. There were 3,103 correct responses and 574 incorrect responses that were either 1 or -1 response away from correct. Figure 6 shows the same five quantile estimates as used for Kent and Lamberts’s (2005) data, graphed against stimulus magnitude: The left panel shows model predictions and data for the RTs of correct responses, the right panel for the RTs for the  $\pm 1$  errors. The parameters used to generate these fits are shown in Table 2.

SAMBA fits the shape of the RT distributions for both correct and incorrect responses, as shown by the relative spacing of the quantiles, and it accommodates changes in the shape and variance of the distribution with stimulus magnitude. The most serious misfit is to the slowest quantiles (90%), especially for incorrect responses to the largest (8, 9) and smallest (1, 2) stimuli; however, the data were quite noisy for the incorrect responses, especially in the slow tails of the distributions. The data reveal that the standard bow effect—longer RTs for middle than for end stimuli—is evi-

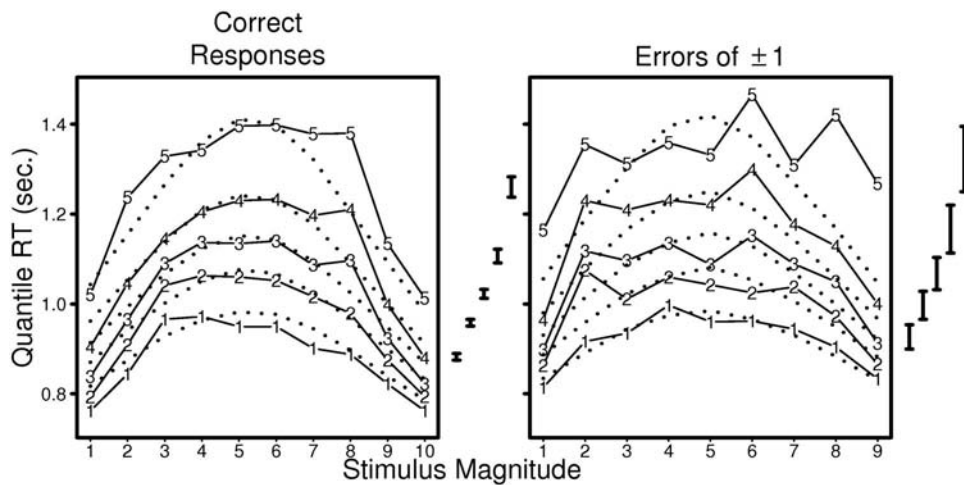


Figure 6. Response time (RT) quantiles for Lacouture and Marley (2004, Experiment 2). The five numbered lines correspond to the 10%, 30%, 50% (median), 70%, and 90% quantiles of the RT distributions for all 10 stimulus magnitudes. The solid lines, with numbers, show these quantile estimates for data, and the dotted lines show predictions from SAMBA. The left panel shows data and model predictions for correct responses, the right panel for errors where responses were either 1 or -1 away from correct (the undershoot errors have been flipped along the x-axis before averaging with the overshoot errors, to preserve direction). The error bars beside each plot show average standard errors for each quantile, calculated by bootstrap, as in Ratcliff, Gomez, and McKoon (2004). sec. = seconds.

dent in all quantiles, at least for correct responses. That is, both the fastest and the slowest parts of each RT distribution are slower for middle than for end stimuli. In binary choice RT modeling ( $N = 2$ ), researchers have sometimes observed quite small effects in the 10% quantile, on the order of tens of milliseconds (e.g., Ratcliff et al., 2004), and even these have proven theoretically challenging when attributed only to changes in the input to the decision process. The data from Lacouture and Marley (2004) and Kent and Lamberts (2005) show vastly larger bows in the 10% quantile, over 200 ms in magnitude.

SAMBA is able to account for large effects on the 10% quantile only because of changes in inputs (response strengths) to the ballistic accumulator stage. The ballistic accumulator stage is very tightly constrained—it has only three parameters that are free to vary when fitting RTs ( $t_0$ ,  $C$ , and  $\beta$ ) without constraint from other aspects of the data (e.g.,  $D$  is fixed by assimilation). Even with this constraint, SAMBA accounts for the large bow in the 10% quantile, and importantly none of these parameters varies between different stimulus conditions. Furthermore, the bow effect is predicted not by arbitrarily estimating response strengths for each stimulus but by the way the response strengths are produced by the earlier stages of SAMBA. The original ballistic accumulator model (Brown & Heathcote, 2005) is unable to accommodate such bows in leading edge RT without making post hoc assumptions about input strength parameters, assumptions that are superceded by SAMBA's architectural constraints.

Critical Tests of Absolute Versus Relative Models

In this section, we review some empirical results that have been construed as critical tests—that is, as providing qualitative evidence against absolute theories of absolute identification or in favor of relative theories. Sequential effects on response accuracy have been claimed to provide critical evidence supporting relative accounts of absolute identification and refuting absolute accounts. We examine several of these effects and illustrate SAMBA's predictions. Each of the effects has a common theme, comparing

responses made to stimuli that were preceded by very dissimilar stimuli with responses made to stimuli that were preceded by similar stimuli. The first effect we discuss was identified by Luce et al. (1982) and the second by Rouder et al. (2004). These findings suggested related effects, which we later examine further using data from Lacouture (1997) and Stewart et al.'s (2005) Experiment 1. First, however, we examine the effects of false feedback, which was claimed by Stewart et al. to provide a critical test between the class of relative judgment theories and the class of absolute judgment theories.

False Feedback

We first describe the fit of the revised SAMBA, then the nature of the revision—namely, the addition of a referent-learning mechanism. Stewart et al.'s (2005) Experiment 2 involved a standard absolute identification task with equal-loudness tones of different frequencies. On just a few trials in each block, Stimulus 3 was presented but—after making a response—the participant was told that it was Stimulus 4. This misleading feedback caused the participants to overestimate the magnitude of the stimulus that followed, as shown in the left panel of Figure 7 by the large positive error for trials following misleading feedback. Note that Stewart et al. reported their data broken down by whether the response on the previous trial was incorrect or correct. However, both RJM and SAMBA predict the same effect whether the previous response is correct or incorrect, and the effect of that variable in that data was small—so Figure 7 presents the results only for trials following a correct response.

Stewart et al. (2005) considered this experiment to be a critical test of absolute versus relative theories. As explained in their Table 8, relative theories (including RJM) predict an average error of 1 following the misleading feedback in Stewart et al.'s experiment and zero average error following standard feedback. In contrast, Stewart et al. concluded that existing absolute (mapping) theories predict a very different pattern: an error of +1 for correct feedback following an error response or for misleading feedback following

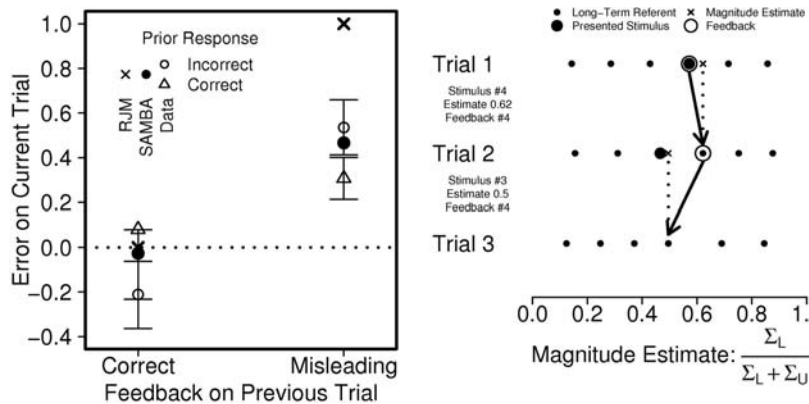


Figure 7. The left panel shows the effect of misleading feedback, with data taken from Stewart, Brown, and Chater's (2005) Experiment 2. Average error was close to zero when correct feedback was given on the previous trial—subjects performed the task properly. After misleading feedback, participants overestimated the magnitude of the current stimulus. Vertical bars show standard errors. The right panel illustrates the feedback mechanism in SAMBA. The three rows of dots show the positions of the long-term referents for each stimulus, for three trials. RJM = relative judgment model.



a correct response, and zero error otherwise. These predictions clearly do not accord with the data. However, they also do not accord with the predictions from the SAMBA model, which fit the data quite well. Figure 7 shows the predictions from the RJM as cross symbols and the predictions of the SAMBA model as solid circles. SAMBA's predictions were obtained using the same parameter values we used in fitting the data of Stewart et al.'s Experiment 1 (see Table 2). RJM's predictions match the data in a qualitative sense, with zero average error following correct feedback and large positive errors following the misleading feedback, but they substantially overestimate the magnitude of the misleading feedback effect. SAMBA's predictions also match the data in the qualitative sense. Moreover, their quantitative agreement with the data is much better than that of RJM. Importantly, the predictions from SAMBA are unaffected by the inclusion or exclusion of the relative mapping process. These results indicate that a purely absolute model can account for the effects of misleading feedback at least as well as a relative model.

We incorporated the effects of feedback into SAMBA in a simple and constrained manner, by partially developing a referent-learning algorithm of the sort discussed in detail by Petrov and Anderson (2005). On each trial, the SAMBA model produces an estimate of the stimulus magnitude—the ratio  $\Sigma_L/(\Sigma_L + \Sigma_U)$ . The mapping stage of the model operates using long-term memories for the mean values of these magnitude estimates for each stimulus, with a value of zero given to the lower anchor (L) and a value of one to the upper anchor (U). We assumed that feedback helps the observer maintain these long-term referents, as illustrated in the right-hand panel of Figure 7. We provide a simple physical analogy to a spring system as an intuitive description of the system; mathematically, the adjustment is made in a simple linear fashion.<sup>10</sup>

The top row of dots in Figure 7 shows the long-term stimulus referents for a hypothetical experiment with  $N = 6$  stimuli. Suppose Stimulus 4 is presented, shown by the larger filled circle in the top row, and this presentation produces a magnitude estimate of .62, shown by the small cross. The observer is then provided with correct feedback, shown by the ring around Stimulus 4. Feedback allows the referents to be updated for the following trial (second row) so that the referent for Stimulus 4 is moved to match its observed magnitude estimate (.62). The long-term referents for the all the stimuli move as if they are locations on a linear spring, whose ends are fixed at zero and one. The point on the spring corresponding to the last magnitude estimate (.62) is deflected so that it aligns with the expected magnitude estimate corresponding to the feedback (Stimulus 4). As in a spring system, this deflection causes compression of locations on one side and expansion on the other, with the ends remaining fixed. On the second trial (second row), Stimulus 3 is presented, producing a magnitude estimate of .5. False feedback now suggests that this was actually Stimulus 4, so the long-term referents are adjusted to make the referent for Stimulus 4 match the observed estimate (.5), shown by the arrow. In this case, this causes significant compression on the left side and significant expansion on the right side, again with the ends remaining fixed.

The example from the right panel of Figure 7 makes clear the reason for SAMBA's good fit to the data from the left panel of Figure 7. First, adjustments to the long-term referents are, on average, smaller when they follow correct feedback than when

they follow false feedback. This is because, when there are very few misleading feedback trials, the stimulus magnitude estimate is typically close to the correct value stored as the long-term referent, so only small adjustments are required when correct feedback is given. On average, when false feedback is given, much larger adjustments are required, causing greater subsequent errors (and corrections, once correct feedback is given again). A referee wondered whether the feedback mechanism makes SAMBA into a more relative (than absolute) model. It is true that the mechanism for updating long-term referents implies that information from recently presented stimuli is used in each judgment, but this does not make the model relative, once again illustrating the problems with separating absolute from relative models; all contemporary models include elements of both. Even with the referent-learning mechanism, judgments in SAMBA are fundamentally absolute, as each stimulus is judged against a set of long-term referents and anchors.

One assumption of the referent-learning mechanism that may require development concerns the magnitude of updates. We assumed that the long-term referent identified by feedback is moved all the way to the location of the stimulus magnitude estimate. A more realistic implementation may be softer, with the long-term referents moving only some fraction of the way. We did not adopt this approach here because it would have required estimation of a learning rate parameter that specifies the relative weights given to the value of the long-term referent and the magnitude estimate in the update rule. Such extra flexibility was not required by SAMBA to obtain a reasonable fit to the data.

Further experimental investigation of false-feedback effects would also be useful, as SAMBA and RJM make quite different detailed predictions. The RJM predicts the same effect of misleading feedback across the entire range—when false feedback of 4 is given to Stimulus 3, the RJM predicts an average error of 1 on the following trial, no matter whether, say, Stimulus 1 or 10, is presented (see, e.g., Stewart et al.'s, 2005, Table 8). Also, RJM predicts that misleading feedback has no effect on trials after the immediately following one. On the other hand, SAMBA predicts the effect of misleading feedback will decrease with the difference between the feedback and the magnitude of the next stimulus—because adjustments to the long-term referents are smaller away from the feedback than near. SAMBA also predicts small effects of misleading feedback on trials after the trial that immediately follows misleading feedback because the long-term referents remain affected by feedback until completely corrected by further trials. However, at the level of analysis permitted by Stewart et al. (2005)'s Experiment 2, the effect of false feedback is clearly not a critical test of the class of relative versus absolute judgment theories, and quantitatively, it favors SAMBA over the RJM.

### Sequential Effects

Luce et al. (1982) manipulated the difference between stimuli presented on successive trials (step size) in four conditions (see

<sup>10</sup> If we let  $Y_i$  be the long-term referent for stimulus  $i$ , then, given a magnitude estimate of  $z$  and feedback indicating that this was stimulus  $s$ , the referents for  $i \leq s$  are adjusted according to  $Y_i \rightarrow zY_i/Y_s$  and, for  $i > s$ , according to  $Y_i \rightarrow 1 - [(1 - z)(1 - Y_i)/(1 - Y_s)]$ .

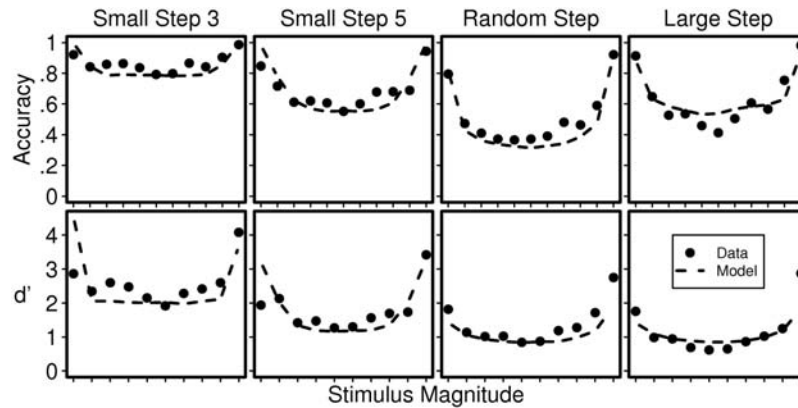


Figure 8. Response accuracy (top row) and sensitivity (bottom) from Luce, Nosofsky, Green, and Smith (1982), with fits of SAMBA (shown by dashed lines). The top-left columns show conditions constrained to have small differences between successive stimuli. The third column shows a standard condition (random differences). The right column has data from a condition where differences between successive stimuli were constrained to be large.

Figure 8). One condition was a conventional 11-stimulus absolute identification design where the stimulus sequence was random—any stimulus could follow any other, with equal probability. This random step condition resulted in typical absolute identification data patterns, shown by the accuracy and  $d'$  graphs in the third column of Figure 8. Luce et al. also used two small step conditions, with constrained differences in the magnitude of successive stimuli. In the Small Step 3 condition, successive stimuli were always very similar: For example, when Stimulus 3 was presented, the next stimulus could only be Stimulus 2, 3, or 4. In the Small Step 5 condition, successive stimuli were moderately similar: For example, if Stimulus 3 was presented, the next stimulus was constrained to be one of Stimulus 1, 2, 3, 4, or 5. Data from the two small step conditions are shown in the left two columns of Figure 8. Finally, Luce et al. used a large step condition in which stimuli were always followed by very dissimilar stimuli. For example, when Stimulus 3 was presented, the next stimulus could only be one of Stimulus 7, 8, or 9. These data are shown on the right of Figure 8.

The manipulation of step size had large effects on response accuracy, shown in the top row of Figure 8. The random step (standard) experiment resulted in the poorest performance, while the Small Step 3 condition resulted in the best performance. The Small Step 5 condition gave better performance than did the large step condition. Some of these effects could be a result of manipulating the number of possible responses, which may have affected accuracy only via response biases. To check this, Luce et al. (1982) also examined their data using a sensitivity measure ( $d'$ , bottom row of Figure 8) designed to take into account response bias effects. The  $d'$  analysis showed that the small step conditions produce greater sensitivity than the large step and random step conditions but also showed that sensitivity in the large step and random step conditions was quite similar, indicating that the accuracy difference between these conditions was mostly due to response bias.

These data may initially be imagined to refute absolute theories because they appear to implicate the previous stimulus in the decision process. Nevertheless, SAMBA accurately captures the

patterns of Luce et al.'s (1982) data and does so without inclusion of the relative component. SAMBA correctly predicts the ordering of the conditions in both response accuracy and sensitivity ( $d'$ ) and also provides a good quantitative description of the data. To fit the data from Luce et al., we added two structural assumptions that reflect the nature of the experiment, and we varied three parameters. Importantly, however, we did not vary any parameters across the four experimental conditions, so different predictions for the different conditions represent purely structural effects. We began with parameters estimated from Lacouture's (1997) data, removed the local process ( $P = 0$ ), and changed the two anchor locations ( $L = 35.5$  dB, and  $U = 90.5$  dB) and the passive decay parameter ( $D = .01$ ) to capture the asymmetry and shape of the observed bow effects. We also increased the rate of decay for accumulators ( $\alpha = .56$ ) in the selective attention stage to match the sequential effects and overall accuracy ( $\eta = 12.8$ ) for Luce et al.'s data.

The two passive decay parameters ( $\alpha$  and  $D$ ) were estimated to have much smaller values in Luce et al.'s (1982) data than in our fits to other comprehensive data sets, presented later. This difference raises an interesting speculation and illustrates the testable nature of SAMBA's architectural assumptions. We have specified the decay parameters in units of trials, rather than real time—each parameter describes how much activation remains after passive decay for the period of one trial. When adjusting parameters from the fits to Lacouture's (1997) data to fit Luce et al.'s data, each decay parameter was reduced by approximately squaring it: from  $\alpha = .75$  for Lacouture and  $\alpha = .75^2 = .56$  for Luce et al.'s data, and  $D = .07$  for Lacouture and  $D = .07^2 = .005$  for Luce et al.'s data. This large change can be parsimoniously interpreted as if the rate of passive decay is constant in real time, if the trial-to-trial interval were twice as long in Luce et al.'s study as in Lacouture's. This hypothesis receives some support from the methodological details of each study: the duration of Lacouture's trials was 1.1 s plus RT; the duration of Luce et al.'s trials was 1.5–2 s plus the slowest of the three subjects' RTs.

Two structural assumptions were made to accommodate the stimulus sequences in Luce et al.'s (1982) experiment. First, we assumed that participants limited their response set; ballistic ac-

accumulators corresponding to any responses other than the allowed responses received no input. For example, suppose Stimulus 4 was presented on the previous trial. In the Small Step 3 condition, the only allowed responses are 3, 4, and 5, so all other response accumulators were given zero input. The second structural modification instantiated a change in the frame of reference for absolute judgments via a change in the distribution of activation in the Poisson process. In the Small Step 3 and Small Step 5 conditions, we assumed that attention was only directed at that range of the accumulators corresponding to allowed responses on the next trial. That is, we assumed that participants were able to focus their attention on the appropriate subrange of stimuli on the next trial.

The second structural assumption parallels suggestions made by Weber et al. (1977), Luce et al. (1982), and Nosofsky (1983) that a roving attention band is moved to focus on appropriate ranges but that this movement is sluggish. The sluggishness of the attention band is manifested in participants' inability to refocus their attention on units corresponding to possible stimuli, particularly in the large step condition. This can be interpreted in SAMBA as due to the relatively slow decay of activity in the selective attention stage, which has a half-life of one trial (around 5 s) in these fits. Our explanation of these data also fits with data from Nosofsky's Experiment 1. Nosofsky ran a similar experiment to Luce et al. but included a discrimination condition in which all 11 stimuli were presented in random order, as in the random condition of Luce et al., but the participants' task was simplified. They were only required to judge whether the current stimulus was the same as, smaller than, or larger than the prior stimulus. Nosofsky observed similar performance in both the discrimination condition and the standard (random) condition for the subset of trials that met the Small Step 3 constraints, with that performance poorer than in the actual Small Step 3 condition. These results support the idea that improved performance in the small step conditions is due to more than just the constrained response set in those condition.

A second sequential effect that appears to support relative accounts over absolute accounts was reported by Rouder et al. (2004; see also Brown et al., 2007; Stewart, 2007). Rouder et al. graphed the probability of a correct response conditional on the signed rank difference between the current and previous stimulus. The filled circles in the left panel of Figure 9 show the results of Rouder et al.'s analysis applied to data from Lacouture (1997). The data show high accuracy for repeated stimuli at the center of the plot, corresponding to zero difference between successive stimuli. The accuracy bonus falls away as the difference between successive stimuli increases, then rises again at the edges of the plot. The rise at the edge of the plot corresponds to improved accuracy for an extremely small stimulus that was preceded by an extremely large stimulus, or vice versa.

The dashed line in the left panel of Figure 9 shows that SAMBA provides a reasonable fit to the data. The model captures the accuracy advantage for repeat and near-repeat stimuli (graph center) as well as the increased accuracy for very large stimulus changes (graph ends). The extreme ends of the figure are influenced only by responses for the very smallest and very largest stimuli, which are accurate in SAMBA because they are near the ends of the selective attention range. Increased accuracy for repeated stimuli (center of the graph) is caused by two mechanisms in SAMBA. First, the assimilation mechanism causes repeated responses to have a slight advantage, as the ballistic accumulator associated with the previous response begins the next trial with a higher activation. Second, when fitting Lacouture's (1997) data, we made partial use of the relative mapping process, which operated on  $P = 62.5\%$  of decisions. The relative mapping process operates by judging the presented stimulus relative to the magnitude estimate of the previous stimulus and the relevant one of the two long-term anchors. The result is increased accuracy for repeat and near-repeat stimuli, as they become effectively judged against a very nearby anchor associated with the prior stimulus. Our model

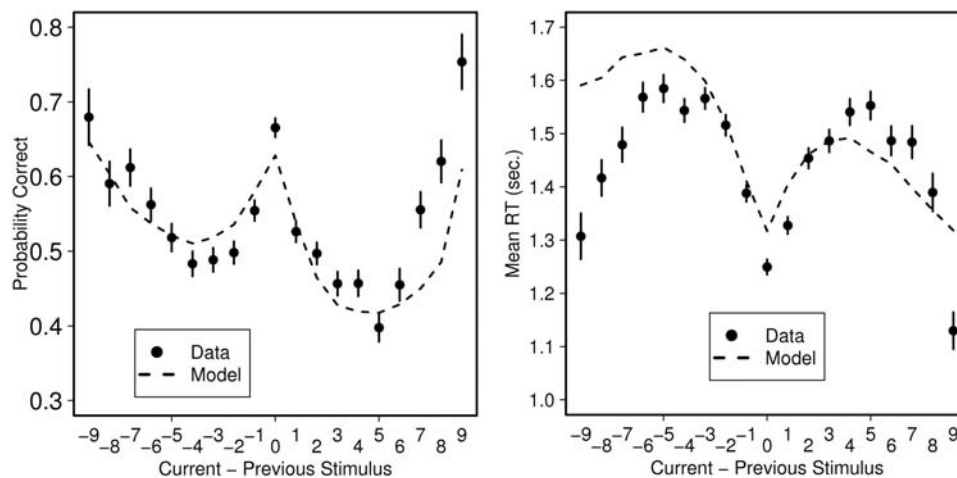


Figure 9. The filled symbols show response accuracy (left panel) and mean response time (RT; right panel) for data from Lacouture (1997) as a function of the difference between stimuli. The graph for probability correct has a local peak at and near the center, signifying an advantage when successive stimuli are similar. The extremes of each graph also show advantages. The graph for mean RT has a parallel dip. Error bars show normal standard errors based on the standard deviation of each point over participants, and dashed lines show predictions from SAMBA. sec. = seconds.

fits to Lacouture’s data set do not distinguish whether individual participants used the relative process for 62.5% of their decisions, or 30 out of 48 of participants always used the relative process and 18 out of 48 never used it, or some combination of the two. Individual data sets were too small for us to carry out individual fits that would reliably differentiate these possibilities.

The right panel of Figure 9 shows that SAMBA captures a similar, but inverted, effect in mean RT. SAMBA’s predictions (dashed lines) capture the M-shaped quantitative trend and match the direction of the asymmetry in the data (faster responses on the right than on the left). However, this fit fails to predict sufficiently fast RTs near the extreme ends of the graph. The failure is mostly due to unusually fast responses at the ends of the range, especially to the largest stimulus in the set (Stimulus 10). These data points are discussed in detail later.

Figure 9 shows that repeated stimuli enjoy large advantages and that near repeats (such as  $\pm 1$  rank-order differences) enjoy smaller advantages. A related question concerns the duration of the advantage for repeated stimuli. Figure 10 shows mean accuracy and RT as functions of how many trials have elapsed since the current stimulus was last presented. A stimulus repeat corresponds to no intervening stimuli (i.e.,  $x = 0$ ). As before, repeated stimuli generate more accurate and faster responses. Figure 10 shows that these advantages do not extend further than immediate repeats. With just one intervening stimulus (e.g., the stimulus sequence . . . 3, 4, 3, . . .), response accuracy and latency are at baseline levels. The dashed lines in Figure 10 show that SAMBA provides a quantitatively accurate account of both the advantage for repeated stimuli (Lag 0) and the lack of advantage for other lags. The locus of this account in the model is as described for Figure 9—a combination of passive decay in the ballistic accumulator stage and partial use of the relative mapping process. Note that SAMBA overestimates total RT in the left panel of Figure 10—once again, this was caused by unusually fast responses in the data to just one stimulus (Stimulus 10), discussed later.

An effect similar to those described by Luce et al. (1982) and Rouder et al. (2004) was also observed by Stewart et al. (2005).

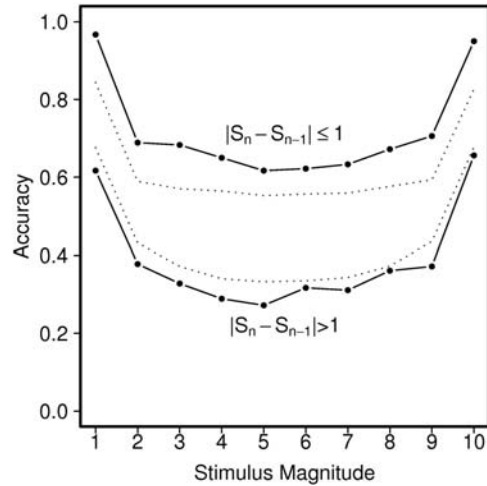


Figure 11. Accuracy graphed separately for stimuli ( $S$ ) that were preceded by near stimuli (either repeats or  $\pm 1$  rank-order difference) versus stimuli that were preceded by far stimuli (more than one rank different). The solid lines with filled circles are data from Stewart, Brown, and Chater’s (2005) Experiment 1. The dashed lines are predictions from SAMBA.

Stewart et al.’s Experiment 1 was a standard absolute identification task using approximately equal-loudness tones of different frequencies, while also manipulating set size and stimulus spacing (we analyze these data in detail below). Performance was much better for stimuli that were close to the stimulus presented on the previous trial. Figure 11 reproduces Stewart et al.’s Figure 26, and graphs accuracy against stimulus magnitude. The graph has separate lines for those stimuli that were preceded by a close stimulus (either an identical stimulus or  $\pm 1$  rank-order difference) and by a far stimulus (all others). This analysis is similar to Luce et al.’s Small Step 3 condition and also is equivalent to taking the central three points in Rouder et al.’s analyses or in our Figure 9.

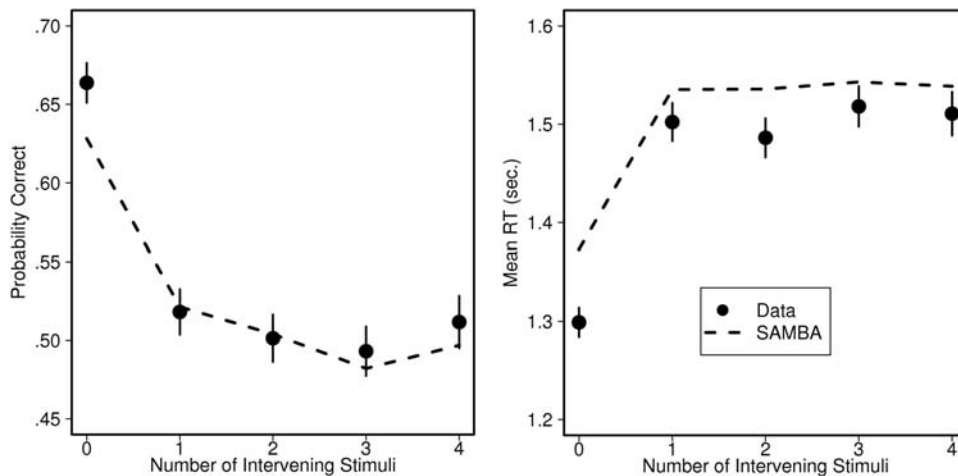


Figure 10. The symbols show response accuracy (left panel) and mean response time (RT; right panel) for data from Lacouture (1997). The  $x$ -axis shows the number of trials that have intervened since the current stimulus was last presented. Error bars show normal standard errors based on the standard deviation of each point over participants. The dashed lines show predictions from SAMBA. sec. = seconds.

In contrast to the results of Rouder et al. (2004), Luce et al. (1982), and Lacouture (1997), Stewart et al. (2005) observed a very large performance bonus for stimuli that were close to the preceding stimulus—accuracy nearly doubled, from 39% to 72%, and response sensitivity ( $d'$ , not shown) more than tripled, from 0.78 to 2.7. By contrast, Luce et al. found a  $d'$  advantage of about 1.1 units, Lacouture found a  $d'$  advantage of only 0.25 units, and Rouder et al. observed an increase in accuracy of only about 12% after their participants were well practiced. Furthermore, in an analysis almost identical to Stewart et al.'s, Purks, Callahan, Braida, and Durlach (1980) found no significant difference in  $d'$ . The predictions from SAMBA are shown in Figure 11 by dotted lines. SAMBA provides a reasonable account of the advantage for repeat and near-repeat stimuli but cannot quite predict the very large bonus observed in the data: SAMBA predicts that response to repeated and near-repeated stimuli are about 1.5 times as accurate as other stimuli, whereas the data show an effect of nearly double accuracy. It seems that SAMBA is sufficiently constrained that it cannot quite predict the extraordinarily large bonus for repeated stimuli observed in Stewart et al.'s Experiment 1. In particular, SAMBA does not include a process specific to the identification of repeated stimuli, which RJM does.

Our preceding analyses suggest that the effects of stimulus repetitions require further investigation, both theoretically and empirically. On the empirical side, previous research has observed a very wide range of effect sizes, from no significant difference (e.g., Purks et al., 1980) to extremely large effects on response accuracy and  $d'$  (Stewart et al., 2005). Numerous data sets seem to show small but reliable effects, for example, Kent and Lamberts (2005), Luce et al. (1982), Rouder et al. (2004), Petrov and Anderson (2005), and our analyses of Lacouture's (1997) data in Figures 9 and 10. Two data sets show large effects: Neath and Brown (2006) and Stewart et al. (2005). The causes of such wide variability in effect size observations are unclear but may be due to the stimuli chosen in each experiment. The experiments that demonstrated very large effects of stimulus repetition were the only ones to use equal-loudness tones of differing frequency. The identification and discrimination of tone frequency may be rather different than for other stimuli because of the existence of critical bands (Green & Swets, 1966, Table 10.1, p. 280). These critical bands may allow participants to perform very accurately on stimulus repetitions, effectively using a more powerful than usual sensory memory. Green and Swets (1966) further indicated that the width of the critical bands for frequencies around those used by Stewart et al. are in the same range as some of the frequency separations in Stewart et al.'s experiment.

Our review of the critical tests of absolute versus relative theories of absolute identification has shown that almost all of the data can be accommodated by a purely absolute version of SAMBA. The only phenomenon that implicates a relative judgment mechanism relates to improved accuracy for repeat and near-repeat stimuli observed in analyses of Lacouture's (1997) data and Stewart et al.'s (2005) Experiment 1. These analyses suggest that the detection of repeated stimuli requires further study, both empirical and theoretical. On the empirical side, it appears that the detection of repeated stimuli is particularly privileged when stimuli are defined by frequency, as opposed to loudness or line length, for example. As for theory, SAMBA's account of stimulus repetition is clearly incomplete and deserves

further development. Preliminary work suggests that expanding SAMBA to include a more detailed account of learning, by continual adjustment of the long-term referents held in memory, helps to accommodate the effects of stimulus repetition without assuming that judgments are relative. We return to this point in the General Discussion.

### Comprehensive Data Sets

In this section, we fit two data sets—from Lacouture (1997) and Stewart et al. (2005)—in great detail. Together, these two data sets exhibit almost all of the benchmark findings in absolute identification. Hence, our analyses of these data sets test whether SAMBA is able to simultaneously accommodate the various patterns in absolute identification data with fixed parameter values. We use two data sets because they provide complimentary coverage of the domain. Lacouture's data (Session 1 from his Experiments 2–5) were collected using  $N = 10$  lines as stimuli and include accurate measurements of RT for each decision. Stewart et al.'s data, taken from their Experiment 1, were collected using equally loud tones of differing frequencies as stimuli. This experiment did not include measurement of RTs, but did have two important stimulus manipulations: The number of tones in the stimulus set was varied ( $N = 6, 8, \text{ or } 10$ ), and the spacing of the tones was either wide (each tone was 12% higher in frequency than the one below) or narrow (6% spacing).

In both Lacouture's (1997) and Stewart et al.'s (2005) experiments, there were many participants, each of whom contributed around 800 data points. Our modeling of each data set stressed parsimony, with each fit using a fixed set of parameters (different for the two data sets) to generate all model predictions for data averaged across participants. This approach provides stringent model tests. For example, in Stewart et al.'s data, set size and stimulus spacing effects must be generated by the model's architecture, rather than by different parameter settings. Our fits to these data sets include the effects of stimulus magnitude, set size, and stimulus spacing on response measures including choice probabilities, sensitivity ( $d'$ ), and full RT distributions. We also examine sequential effects on both choices and RT. This is the first time that an absolute identification model has had the explanatory range to be tested so comprehensively.

Parameter estimates for the two data sets are shown in Table 2. For Stewart et al.'s (2005) experiment, no parameters were changed to model the different set size conditions or the different stimulus spacings. Instead, the predictions of SAMBA change with set size and stimulus spacing simply because the stimulus representations reflect the physical elements of the experimental design. Since Stewart et al.'s data did not contain RT measurements, the parameters of the ballistic accumulator stage of SAMBA were not estimated (except for the assimilation parameter,  $D$ , which can be estimated from choice data). As a result, there were seven model parameters for Stewart et al.'s data and 11 for Lacouture's (1997) data. We do not claim that these parameter estimates are optimal. No formal optimization was performed, and only a relatively coarse grid of parameter estimates was tried. Better fitting solutions likely exist, but the values displayed in Table 2 provide a sufficiently good account.

### Data From Lacouture (1997)

Several phenomena observed in Lacouture's (1997) data have already been discussed in the earlier section on critical tests because they featured in our attempts to distinguish absolute from relative models of absolute identification. These phenomena included sequential effects on response accuracy, decision sensitivity, and mean RT, as well as the effect of response repetition on accuracy and mean RT. Five more phenomena are discussed in this section, including the effects of stimulus magnitude on response accuracy, mean RT and RT distributions, and assimilation and contrast. For all graphs, we display both data (using symbols) and model fits (using dashed lines).

Figure 12 shows response accuracy as a function of stimulus magnitude. This is the standard bow effect plot, and SAMBA accounts for the bow shape and the asymmetry evident in the data, except for the largest stimulus. SAMBA posits that asymmetry is due to unequal placing of the stimulus anchors, with the lower anchor (L) placed much closer to the smallest stimulus than the upper anchor (U) is to the largest stimulus. When fitting Lacouture's (1997) data, we estimated the lower anchor for the selective attention process at 91 pixels—just one pixel smaller than the smallest stimulus—and the upper anchor at 520 pixels—100 pixels larger than the largest stimulus. The asymmetry in the anchor placements (in both pixel and log units) accounts for the strong asymmetry in the data: Responses to smaller lines were faster and more accurate than responses to larger lines. SAMBA is the first absolute identification model that accounts for asymmetry in choice and RT, which are often observed in data but have been almost uniformly ignored. However, SAMBA's account is still quite constrained and cannot accommodate one unusual aspect of the asymmetry in Lacouture's data—namely, responses to Line 10 are both very accurate and extremely fast relative to other responses.

Figure 12 shows that SAMBA provides a good account of differences in mean RT for the different stimulus magnitudes in Lacouture's (1997) data. Figure 13 extends this analysis in two ways—

using full RT distributions, rather than just means, and showing both correct and incorrect responses. Figure 13 uses a similar format to Figure 5 (for Kent & Lamberts's, 2005, data). The left panel shows correct responses (7,324 data points), and the right panel shows data for errors of 1 response (2,072 data points) and  $-1$  response (3,407 data points). Data for 1 and  $-1$  errors were too noisy to present separately, so we have averaged them together, after flipping response magnitudes 1–10 for the  $-1$  errors.

SAMBA accurately predicts RT distributions for the correct responses apart from overestimating RT for the largest stimulus (as discussed regarding mean RT). The model captures the shape of these RT distributions, as illustrated by the relative spacing of the quantiles, and the variance of the distributions, as illustrated by the absolute spacing of the quantiles. SAMBA also captures the changes in shape and variability across the range of stimulus magnitudes. For the incorrect responses, SAMBA provides a qualitatively reasonable account but shows systematic quantitative errors. For example, SAMBA accurately matches the entire RT distributions for errors on the smallest and largest stimuli and for all stimuli for the middle quantiles. However, it predicts too large a spread in the tails (i.e., 10% and 90% quantiles) of incorrect RT distributions for the middle range of stimuli.

Figure 14 shows assimilation and contrast effects in Lacouture's (1997) data, along with model fits. For this figure, we use the same layout as Figure 4 (see also Ward & Lockhead, 1970; Stewart et al., 2005) and show SAMBA's predictions using solid lines, rather than the usual dashes. The plot shows response biases, measured as average errors. For example, if Stimulus 5 is presented, the error will be zero if the participant responds correctly with Response 5 but will be 2 if the participant gives Response 7 and  $-1$  if the participant gives Response 4. The graph shows average errors on Trial  $N + X$ , for  $X = 1, 2, \dots, 8$ , conditional on the magnitude of the stimulus presented on Trial  $N$ . Filled symbols show average error when the preceding stimulus was small; open symbols show the error when the preceding stimulus was large.

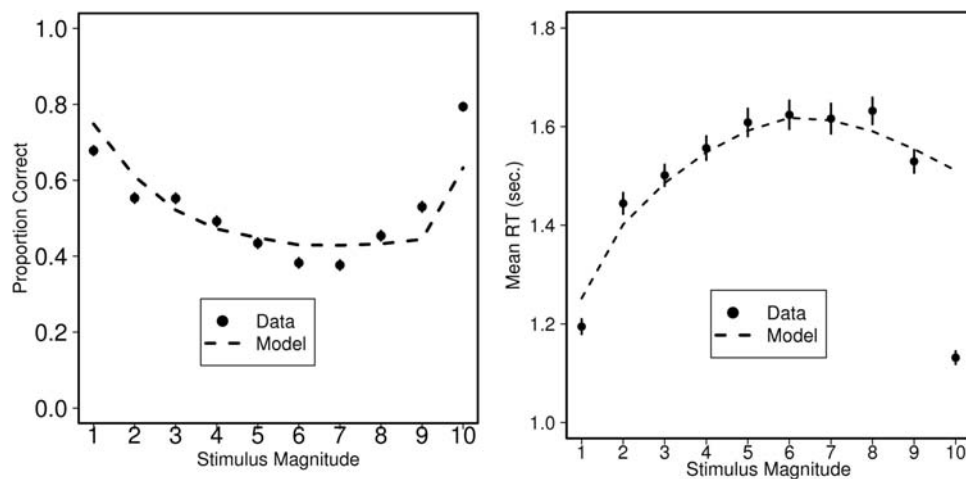


Figure 12. Accuracy and mean response time (RT) as functions of the ordinal stimulus magnitude ( $x$ -axis) for data from Lacouture (1997) and SAMBA model fits. The vertical lines on each data symbol show normal standard errors based on the standard deviation of each point over participants. sec. = seconds.

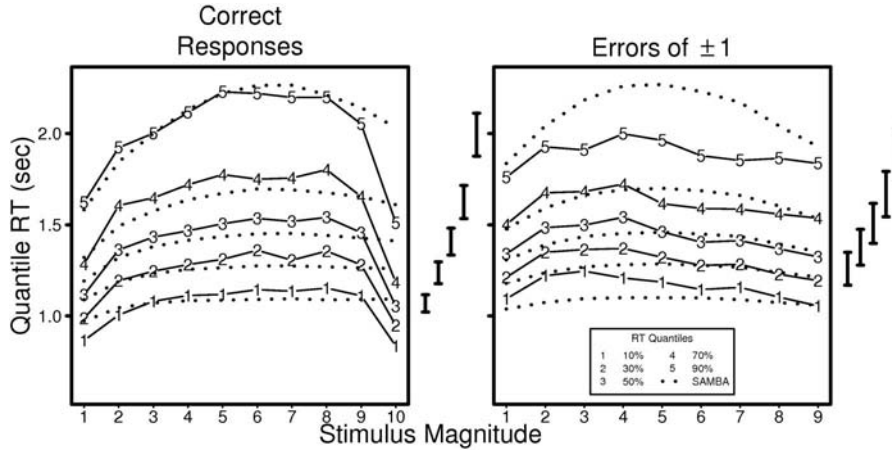


Figure 13. Response time (RT) distributions from Lacouture (1997) for correct responses (left panel) and for errors of 1 or -1 response category (right panel). The lines numbered 1-5 in each panel show 10%, 30%, 50%, 70%, and 90% quantiles estimated from the data for each of the 10 stimulus magnitudes (x-axis). Dashed lines show predicted quantiles from SAMBA. Bars to the left of each plot show average standard errors for each quantile, calculated by bootstrap from the raw data (see, e.g., Ratcliff, Gomez, & McKoon, 2004). Note that standard errors are larger for the longer quantiles and for the incorrect response data. sec. = seconds.

The characteristic pattern of assimilation at lag = 1 and contrast at lag > 1 was observed in these data. That is, when a large stimulus was presented just one trial previously, errors were positive, and vice versa for small stimuli. When a large stimulus was presented several trials previously (X = 2-8), average error was negative, and vice versa for small stimuli. SAMBA successfully fits the qualitative pattern of assimilation at short lags and contrast at longer lags. In Lacouture's (1997) data, contrast is strongest at lag = 3, while, in many other data

sets (including Ward & Lockhead, 1970, and Holland & Lockhead, 1968), contrast is strongest at lag = 2 and decays monotonically thereafter. To fit the peak at lag = 3 observed in Lacouture's data, we chose K = 4, so the redirection of attention in the selective attention stage lasts for four trials.

Data From Stewart et al.'s (2005) Experiment 1

Stewart et al.'s (2005) Experiment 1 provides a valuable resource as it allows a direct comparison between the goodness-of-fit to choices of two competing models. Stewart et al. provided graphs of many observed data patterns and simultaneously provided predictions from their RJM. By fitting SAMBA to the same data patterns, we can compare the two models, although we are unable to compare them on aspects of absolute identification that the RJM does not cover (such as RT). Comparing goodness of fit is complicated by the varied nature of the phenomena we examine, including response probabilities, average biases, and sensitivity measures.

In the absence of an agreed statistical model, we can do no better than compare how closely the predicted values from each model (RJM and SAMBA) match the observed values. To do this, we use root-mean-square error (RMSE). The magnitude of RMSE is, of course, without statistical meaning. Nevertheless, the relative size of RMSE for the two models is informative. The most intractable problem with this approach is model complexity—it is possible for a false model to more accurately fit observed data than a true model if the false model is the more complex of the two. Stewart et al. used eight free parameters to fit the RJM to the data from their Experiment 1, whereas the fits of SAMBA have only seven, suggesting that RJM is more complex. Although the number of parameters does not give a complete measure of model complexity, this difference indicates that complexity is unlikely to explain cases where

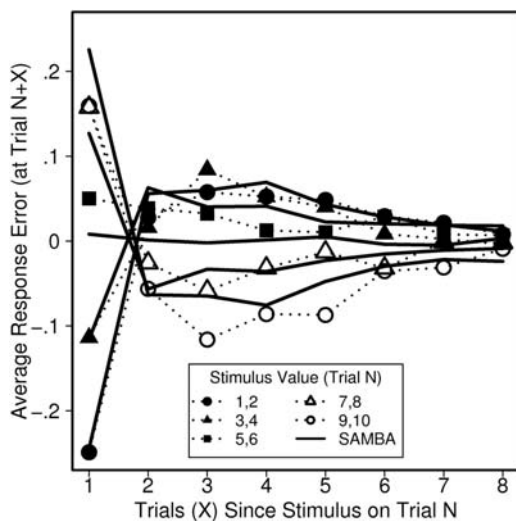


Figure 14. Average error as a function of preceding stimulus (different symbols—see legend) and number of trials since that stimulus (x-axis), from Lacouture (1997). Assimilation occurs to stimuli presented one trial previously: Errors are positive at x = 1 for large previous stimuli (filled symbols) and negative for small previous stimuli (open symbols). At longer lags (x = 2-5), contrast, that is, the opposite pattern, is observed.

This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.

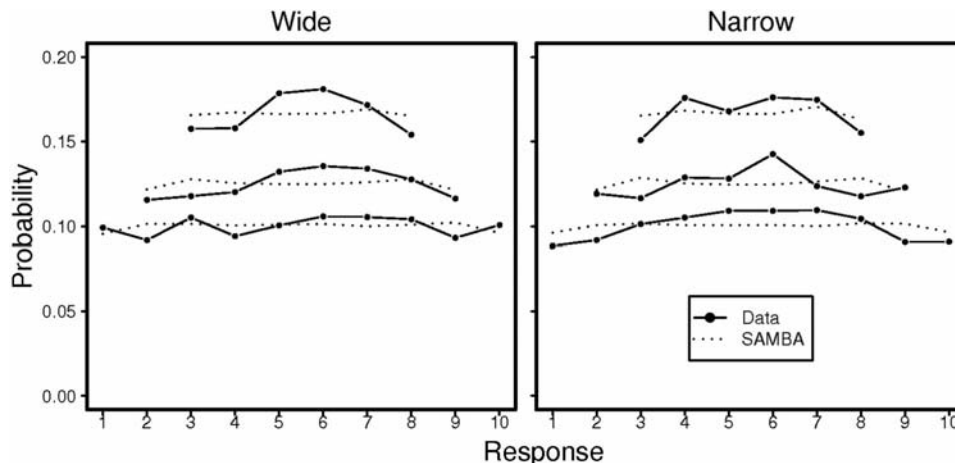


Figure 15. The probability of using each response, for narrow and wide conditions and set size  $N = 6, 8,$  and  $10$  (top to bottom, respectively). Central responses were used more frequently than extreme responses. The dotted lines show the predictions of SAMBA; lines with filled circles are data from Stewart, Brown, and Chater (2005).

SAMBA provides an equal or better fit than RJM.<sup>11</sup> To foreshadow our results, we found that both SAMBA and the RJM provide quite good fits to the data, which do not include RTs.

The stimuli for Stewart et al.'s (2005) Experiment 1 were equal-loudness tones of different frequencies. In the narrow condition, the tones were separated by 6% (e.g., the lowest tone was 768.7 Hz, the next tone was 6% higher at 814.8 Hz, and so on), while, in the wide condition, tones were separated by 12%. Set size was manipulated independently of stimulus spacing, by using either all  $N = 10$  tones, or just the central  $N = 8$  or  $N = 6$ . Both set size and stimulus spacing were manipulated between subjects. The data do not include RTs, so several of SAMBA's parameters can be omitted. Below, we compare SAMBA's predictions with predictions from Stewart et al.'s RJM. As presented in Stewart et al., the RJM does not accommodate asymmetry in the data (i.e., more accurate responses to very small than very large stimuli, or vice versa), although SAMBA handles such data naturally, as in our fits to Lacouture's (1997) experiment. However, to keep the flexibility of SAMBA similar to that of the RJM, we constrained our fits to be symmetric (i.e., we used  $L = U$ ).

Stewart et al. (2005) also did not model the effects of the wide and narrow stimulus conditions separately, even though the RJM has a perceptual noise parameter that might be able to account for this manipulation. Certainly, SAMBA's perceptual noise process ( $\sigma_p$ ) can account for differences between the narrow and wide conditions, and so, we estimated it in our fits to this data. We could have ignored the differences in the data from the wide and narrow spacing conditions and therefore dropped the perceptual noise parameter from our fits. This would have had the advantage of allowing SAMBA and RJM's predictions to be compared more easily but the disadvantage of forcing a greater disparity in model complexity: Without the stimulus noise parameter, SAMBA would use just six parameters compared with RJM's eight.

We assumed that frequency is represented on a simple logarithmic scale and that the magnitude of psychophysical noise was  $\sigma_p = 1.9\%$  of stimulus magnitude. The psychophysical noise was more influential in the narrow than in the wide condition because

the stimuli were only 6% apart in the narrow condition, so the 1.9% standard deviation more often resulted in confusion between adjacent stimuli. For parsimony, we also treated the narrow and wide conditions equally: The lower anchor was always 10% of the total stimulus range lower than the smallest stimulus, and the upper anchor was always 10% of the stimulus range above the largest stimulus. To provide even greater model constraint, we assumed that this anchor placement was constant across set sizes ( $N = 6, 8,$  and  $10$ ) as well as the stimulus spacing conditions. For example, in physical terms for the narrow stimulus spacing condition with set size  $N = 10$ , the lower anchor was set at  $L = 773$  Hz and the upper anchor at  $U = 1,369$  Hz. Finally, we also assumed that participants in Stewart et al.'s (2005) experiment always used the locally relative judgment mechanism in SAMBA. This assumption follows from the extraordinarily large accuracy bonus previously observed for repeat and near-repeat stimuli in these data.

In what follows, we illustrate the effects of set size and stimulus spacing on response frequency, accuracy, sensitivity ( $d'$ ), and average bias (contrast and assimilation). Note that we have already discussed the effects of stimulus repetition and near repetition on accuracy in the critical tests section. All of the following different data fits, as well as the fits to repetition effects already reported, were modeled in SAMBA using a single set of parameters (shown in Table 2).

The first new analysis examines response frequency. Figure 15 shows how often each response was given, separately for the wide and narrow conditions and for the set sizes  $N = 6, 8,$  and  $10$ . There is little difference in the shape of the response probability curves

<sup>11</sup> We calculated RMSE values by reading data from the published graphs in Stewart et al. (2005) and separately comparing the predictions of the RJM to the data from the wide and narrow stimulus spacing conditions. We took the same approach for calculating RMSE values from SAMBA's predictions. Slight differences in the RMSE values might arise if a different approach were used (e.g., if one compared with the average of the wide and narrow data conditions or if one differentially weighted the different set sizes).



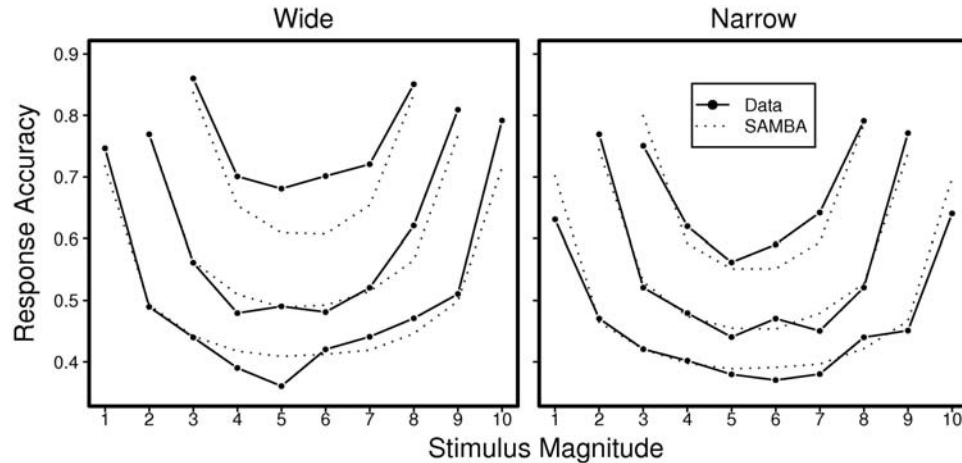


Figure 16. Response accuracy as a function of set size and stimulus spacing and set sizes  $N = 6, 8,$  and  $10$  (top to bottom, respectively). The dotted lines show the predictions of SAMBA; lines with filled circles are data from Stewart, Brown, and Chater (2005).

between the wide and narrow conditions. SAMBA captures the changes in response probability with set size and the lack of change with stimulus spacing. Both SAMBA and RJM (see Stewart et al.'s, 2005, Figure 20) fit these data very well, and RJM performs slightly better, with an RMSE of 0.0062 compared with 0.0079 for SAMBA.

Figure 16 shows response accuracy for wide and narrow conditions separately for set sizes  $N = 6, 8,$  and  $10$ . Note that the bow effect is apparent at every set size and increases in depth with increasing set size. Performance was also worse for narrow than widely spaced stimuli. SAMBA accommodates the bow effect, the changes in accuracy with set size, and even the improved accuracy for wide over narrow spaced stimuli. The only apparent misfit is to the central four stimuli from the set size  $N = 6$  data under the wide spacing condition, where SAMBA underestimates performance.

Quantitatively, SAMBA fits the data better than the RJM (RMSEs of .035 and .05, respectively).

Figure 17 shows the accuracy data transformed to response sensitivity ( $d'$ ). As for accuracy, SAMBA underpredicted performance on the smallest set size, and this effect is greatly exaggerated in the  $d'$  data because of the nonlinear stretching effect of the inverse normal transformation for probabilities close to one. Indeed, the depth of the  $d'$  bow for the data for set size  $N = 6$  (and, to a lesser extent,  $N = 8$ ) is remarkably large in Figure 17 and not well fit by our model. The better account of stimulus spacing effects and the  $d'$  bow result in SAMBA fitting the data somewhat better than RJM (RMSEs = 0.39 and 0.43, respectively).

The response accuracy data (see Figure 16) describe only correct responses, while the sensitivity data (see Figure 17) use all responses but collapse them into a single summary statistic. Figure

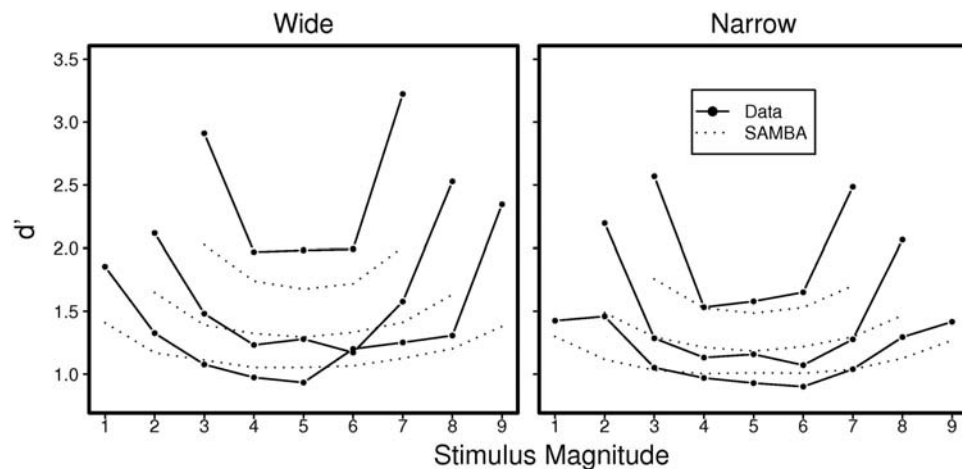


Figure 17. Response sensitivity ( $d'$ ) as a function of set size and stimulus spacing and set sizes  $N = 6, 8,$  and  $10$  (top to bottom, respectively). The dotted lines show the predictions of SAMBA; lines with filled circles are data from Stewart, Brown, and Chater (2005).

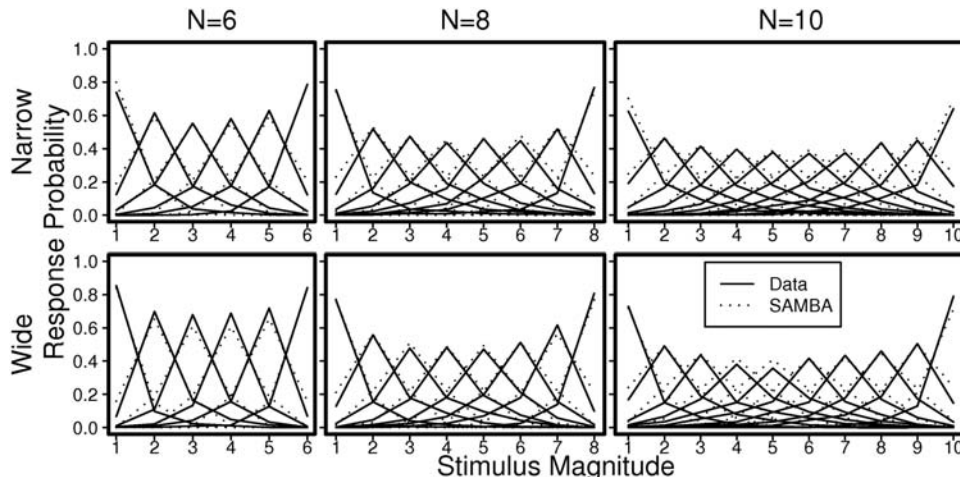


Figure 18. Response matrices for the three different set sizes and the wide–narrow stimulus spacing conditions from Stewart, Brown, and Chater (2005). Each line represents the probability of a particular response, conditioned on the various stimuli.

18 shows the probability of all responses to all stimuli—that is, full confusion matrices—along with SAMBA’s predicted values. As observed in Figures 15 and 16, the probability of a correct response (the highest peak for each line in each graph) decreases with increasing set size and in the narrow compared with the wide stimulus spacing conditions. SAMBA captures the complete distribution of error and correct responses, including the effects of stimulus magnitude and narrow versus wide stimulus spacing. However, for set size  $N = 10$ , the model overpredicts the proportion of  $\pm 1$  responses to Stimuli 2–5, and for set size  $N = 6$ , it underpredicts overall accuracy in the wide spacing condition. These same effects are apparent in the accuracy graph (see Figure 16). RMSE for the confusion matrices favors RJM very slightly over SAMBA (0.038 vs. 0.041).

Finally, we turn to the sequential effects of assimilation and contrast. Figure 19 shows assimilation and contrast effects in Stewart et al.’s (2005) data, using the format introduced for Figures 4 and 14. Each graph plots the average error on the current trial as a function of the size of a preceding stimulus (separate lines). The  $x$ -axis shows the number of trials that have elapsed since that preceding stimulus was presented (lag). As usual, there is an assimilation effect at lag = 1, with responses being biased toward the previous stimulus: if the previous stimulus was large, average error is positive, and vice versa. At longer lags (2+), contrast is observed, with responses biased away from the previously seen stimuli. Most trends in the sequential data are captured well. SAMBA predicts assimilation at lag = 1 followed by contrast at the longer lags. It also predicts larger effects in the narrow

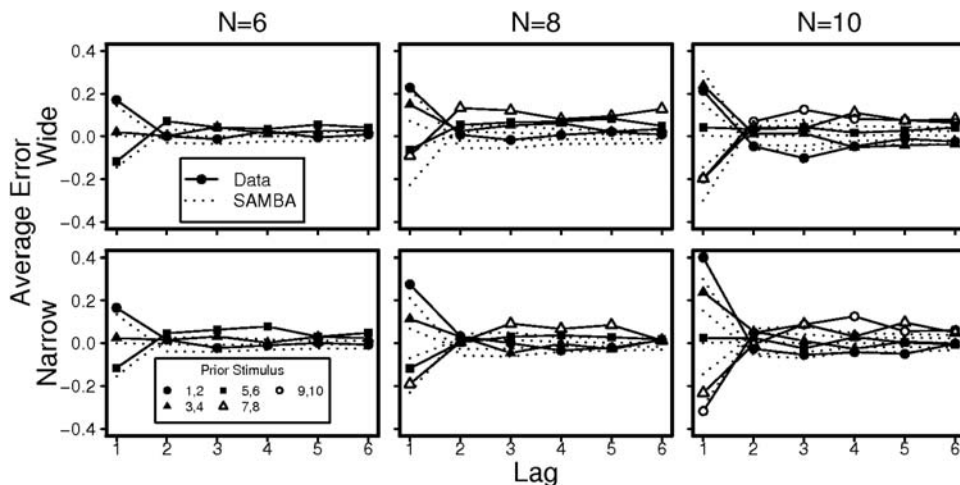


Figure 19. Assimilation and contrast effects in data from Stewart, Brown, and Chater (2005) separately for the three set sizes  $N = 6, 8,$  and  $10$  and for wide and narrow stimulus spacing conditions. The lines in each plot show different magnitudes for preceding stimuli (see legends).

than in the wide stimulus spacing and larger effects with increasing set size. SAMBA's predictions match the data just as well as the RJM's (both RMSEs = 0.044).

### General Discussion

In interpreting data from absolute identification, a distinction has been proposed between local and global phenomena. Local phenomena are those with short temporal duration, particularly effects of recent stimuli and responses on current decisions. Global phenomena are relatively stable over time, such as the effects of stimulus ranges and set sizes. Some theoretical accounts have focused on local processes (Holland & Lockhead, 1968; Laming, 1968, 1984; Lockhead & King, 1983), while others have focused on global processes (e.g., Braida et al., 1984; Lacouture & Marley, 1991, 1995, 2004; Marley & Cook, 1984). Several more recent models have incorporated both local and global processes (e.g., Nosofsky, 1997; Nosofsky & Palmeri, 1997; Petrov & Anderson, 2005). Recently, Stewart et al. (2005) took a more extreme position against global processing, affirming that absolute identification is based only on local processes of a particular type, namely, relative judgment, with no absolute or global processing whatsoever (see also Laming, 1984; Lockhead, 2004).

We have described an extension of global, restricted-capacity models, developed in various studies by Cook, Karpiuk, Lacouture and Marley, to include local processes. SAMBA provides a comprehensive account of absolute identification because it predicts not only choices but also the time taken to make them. Also, SAMBA includes an account of multiple sources of variability affecting decision processes, including sequential effects, and its predictions for RT are not restricted to just mean RT bow effects. Our analyses of data from Kent and Lamberts (2005), Lacouture (1997), and Lacouture and Marley (2004) show that SAMBA provides an accurate account of the entire distribution of RTs as a function of stimulus magnitude. The model also provides an accurate account of asymmetries and sequential effects on RT distributions and response choices. The choice and RT effects are predicted by the same set of parameters, providing a more stringent test than fitting choice or RT data alone.

Although SAMBA provides a close quantitative fit to dozens of phenomena in absolute identification, there are some places where it fails to fit the data, underlining the point that SAMBA is sufficiently constrained in its predictions to be falsifiable. One of the failures, overprediction of Stewart et al.'s (2005) false-feedback effect, was relatively small, and SAMBA still performed better than the only other model tested against this effect. Given that a natural extension of SAMBA's learning mechanism (partial correction for feedback) predicts an appropriately smaller effect, this failure is not troubling. Two other small failures were associated with underprediction of bow effects involving the very largest stimulus for some conditions in Luce et al.'s (1982) and Lacouture's (1997) data. The cause of such failures is unclear, although the limitation to one stimulus in each case suggests some idiosyncratic factor may be in play, with increased performance perhaps related to external referents that were most salient for the largest stimuli.

Two larger and more systematic failures concern responses to repeated stimuli in Stewart et al.'s (2005) Experiment 1 (see Figure 11) and the distribution of RTs for incorrect responses in Lacouture's (1997) data (see Figure 13). The latter failure is striking because our RT distribution fits for correct responses are of almost the same

quality as achieved by the leading accounts of two-choice RT (e.g., Ratcliff & Smith, 2004). One explanation for our poorer fit to incorrect responses is that two-choice RT models are typically not constrained to fit as wide a range of phenomena as were our fits to Lacouture's data (e.g., sequential effects). A second explanation concerns our use of quantiles averaged over participants, necessitated by small sample sizes per participant. Averaging can spread quantiles when individual differences are present, which might particularly be a factor for incorrect responses, as participants can vary markedly in their speed-accuracy tradeoff setting (see, e.g., Brown & Heathcote, 2003). Different settings are associated with systematic changes in the speed of error responses. In support of this explanation, we note that our fits to the error RT distributions for the individual participant in Lacouture and Marley's (2004) Experiment 2 were quite good.

SAMBA performs well in fitting the accuracy and RT advantages for repeated stimuli in many data sets (e.g., Lacouture, 1997; Luce et al., 1982). In Stewart et al.'s (2005) data, however, repetitions have a much larger advantage than SAMBA can predict. We noted earlier that SAMBA has no special mechanism for repeated stimuli. Given the difference between Stewart et al.'s results and others in the literature, this failure may not indicate a direct falsification. A possible explanation is that the stimuli in Stewart et al.'s experiment differed in tone frequency, whereas Lacouture (1997) used line lengths and Luce et al. (1982) used tone intensity (loudness). It is possible that stimulus repetitions are more easily detected in frequency than in other continua; this suggestion is compatible with the existence of critical bands in the perception of frequency (Green & Swets, 1966). Further experimentation is required to investigate the effects of response repetitions on accuracy in a range of different stimulus modalities.

SAMBA accounts for all standard sequence effects, global effects, and the effect of misleading feedback without any relative judgment process. Short-term memory in the decision stage predicts repetition and assimilation effects, whereas short-term memory in the selective attention stage predicts contrast effects. A combination of these processes, along with restrictions on the available responses, explains the effects of manipulating the stimulus sequences found by Luce et al. (1982). Hence, SAMBA demonstrates that none of these effects necessarily implicates relative judgment. There is just one data pattern that does implicate relative judgment in SAMBA: the accuracy bonus sometimes observed for repeat and near-repeat stimuli. Although just a small part of the data, this effect has proven theoretically important, and not just for SAMBA. The same analyses proved critical for the RJM, prompting Stewart (2007) to modify the RJM to allow the stimulus two trials back to be used as the basis for relative judgment on some trials instead of the memory for the stimulus one trial back. This modification is one step toward an extension where "any previous stimulus could be used [that would] introduce long-term representation of magnitudes into the model" (Stewart, 2007, p. 536).

### Future Developments

There are several aspects of SAMBA that warrant further development. That development will be aided by the clear and testable predictions that arise from several of SAMBA's assumptions about the cause and locus of certain phenomena. We do not provide an exhaustive discussion on these topics here but instead restrict ourselves to three aspects requiring further development and to three novel predictions.

1. *The selective attention mechanism.* Like its precursors,

SAMBA is based on the notion of global processing—the central tenet of the model is that stimuli are judged relative to a context that changes slowly over time. The precise nature of this selective attention context requires further study. For example, more experiments are required to determine which experimental factors cause participants to change their anchor positions and how quickly these changes occur. Other experiments are required to investigate what factors affect the size, duration, and peak timing of contrast effects; these results will illuminate what factors affect the Poisson process. For example, it is an open question whether the dynamic aspects of that process operate in units of real time or trial time. If contrast effects operate on a time (rather than trial) basis, the peak magnitude for contrast effects could be manipulated by changing the response-to-stimulus delay interval (RSI). If the RSI is made very short, the number of trials before peak contrast is reached will increase, shifting the location of peak contrast further to the right (i.e., lag = 3 or lag = 4). We cannot tell from the published details the precise RSI values used in previous data sets, so we are currently pursuing new experiments in which the intertrial interval is manipulated either within or across experiments.

2. *Learning the stimulus representations.* The mechanics of the mapping stage are completely determined by long-term representations of average stimulus magnitude estimates. For example, if there are three evenly spaced stimuli, with anchors fixed at the lower and upper magnitudes, the average magnitude estimates are  $\{0, \frac{1}{2}, 1\}$ . Throughout this article, except when addressing misleading feedback, we have assumed that these values are unvarying and accurate estimates of the true expected values. This assumption is clearly too strong, suggesting the need to develop a model of how the magnitude estimates are learned and maintained. We developed the beginnings of this model when addressing the false-feedback data from Stewart et al.'s (2005) Experiment 2. For those data, we suggested a mechanism that adjusts magnitude estimates to align with feedback. This mechanism is similar to one of two mechanisms proposed by Treisman and Williams (1984). In future work, we will examine the addition of Treisman and Williams's second mechanism (an assimilative process) into SAMBA. Together, these two processes provide a testable system for accommodating the effects of correct feedback, false feedback, and the absence of feedback (see also Mori & Ward's, 1995, discussion of a similar use of Treisman & Williams's, 1984, mechanisms). Our preliminary investigations have offered the intriguing prospect that the use of Treisman and Williams's adjustment (learning) mechanisms may eliminate altogether the need for the relative judgment process in SAMBA. This suggests that those effects previously considered indicative of local relative judgment (particularly the accuracy advantage for repeat and near-repeat stimuli) may be alternatively considered as evidence of a learning process that maintains and adjusts long-term referents for magnitude estimates.

3. *The causes of incorrect responses.* We assume that the mapping stage is error free. The decision stage can produce errors because of the influence of responses on previous trials on the starting points of the ballistic accumulators, but our estimated parameters suggest that this effect is smaller than that of the selective attention stage. For example, in fits to Lacouture's (1997) data, only 23% of incipient choices were changed by the action of the decision phase (i.e., in 23% of cases, the final response generated by the decision stage was different from the response corresponding to the maximum output of the selective attention and mapping stages). These assumptions stand in contrast to several prior models, which have attributed errors more directly to processing after a magnitude estimate is produced, such as in a decision mechanism (e.g., Kent & Lamberts, 2005; Lacouture & Marley, 1995, 2004; Nosofsky & Palmeri, 1997; Treisman & Williams, 1984).

### Predictions

The current version of SAMBA leads to some novel empirical predictions, of which we present three.

1. *Unequal stimulus presentation frequency (a test of the selective attention stage).* The contrast mechanism associated with the selective attention stage makes the perhaps surprising prediction that if stimuli within a subinterval of the range are presented more frequently, then those stimuli will eventually be identified more accurately at the expense of the remaining stimuli outside the subinterval. We are aware of two experiments that tested the above prediction, one of which found changes in response bias but not in  $d'$  (Chase, Bugnacki, Braid, & Durlach, 1983), whereas the other found a small but significant effect on  $d'$  (Nosofsky, 1983). We are currently working on sharper experimental tests of this prediction.
2. *Nonuniform stimulus spacing (a test of the mapping stage).* The mapping stage of SAMBA predicts that stimuli that are spaced closer together, relative to other stimuli, will be more poorly identified. This prediction sounds trivial but is not because it holds even when the closely spaced stimuli are still well above threshold in comparative judgment tasks. We have interpreted Lockhead and Hinson's (1986) data as supporting this prediction, although other interpretations have been advanced, for example, the use of appropriately selected cutpoints on a suitably constructed response variable (Stewart et al., 2005). Data from Lacouture (1997) illustrate that stimuli that are more widely spaced, relative to other stimuli, are more accurately identified.
3. *The effect of ITI on assimilation (a test of the decision stage).* Our assumption that assimilation is caused by passive decay in the ballistic accumulators leads to the prediction that assimilation effects will be smaller when the ITI is larger. To our knowledge, relevant absolute identification experiments have not been performed (but see DeCarlo, 1992, for related supporting theoretical and empirical work in magnitude estimation). Many absolute

identification models would be able to accommodate this prediction, for example, by simply adjusting parameter values associated with assimilation. However, SAMBA is currently the only model of absolute identification that makes a quantitative prediction for this effect as an a priori consequence its architecture.

These three predictions follow from the current version of the SAMBA model, in which we have assumed that the distribution of selective attention is mostly uncontrolled by the participant. Instead, accumulators are incremented randomly, except for the redirection that results in contrast. With these assumptions, SAMBA makes starkly different predictions for the effect of unequal stimulus presentations, in Prediction 1 above, and for nonuniform stimulus spacing, in Prediction 2. That is, unequal stimulus presentations can improve discrimination for the more frequently presented stimuli, whereas nonuniform spacing is likely to decrease discrimination for the stimuli that have a decreased relative spacing in the nonuniform case. However, if attention were under sufficient strategic control, this differential prediction for the two tasks would be reduced or eliminated. Therefore, stronger tests of the degree to which selective attention is under strategic control are required and are being pursued.

### Concluding Remarks

In closing, we note that our results have implications beyond absolute identification. Although strong sequential effects are present in choice and RT data from other paradigms (e.g., Gilden, 1997, 2001; Gilden, Thornton, & Mallon, 1995; Kelly, Heathcote, Heath, & Longstaff, 2001; Laming, 1968; Wagenmakers, Farrell, & Ratcliff, 2004, 2005), these sequential effects are not accommodated by existing choice RT models (but see Wagenmakers et al., 2004, for some possible directions). We have suggested a mechanism for assimilation effects that could be as easily implemented in any number of sequential sampling models of choice RT (such as those in Ratcliff & Smith, 2004) as it can be in Brown and Heathcote's (2005) ballistic accumulation model—namely, by making the starting points of accumulation dependent on the response made on the previous trial. This mechanism moves choice RT models toward an explanation, rather than an assumption, of variability (i.e., start points are usually assumed to be variable, with no testable explanation of the source of that variability).

We have also suggested that longer term sequential effects (e.g., contrast) can arise as a consequence of changes over trials in the inputs to the decision stage and have provided a mechanism that predicts the magnitudes and variability of these inputs, at least in the context of absolute identification. At present, most choice RT models (see Smith, 1995; Smith, Ratcliff, & Wolfgang, 2004, for exceptions) simply assume the appropriate magnitudes and variability of inputs to fit data, with little motivation for how these inputs arise. The latter approach, although perhaps viable for two-choice data and a limited number of stimulus conditions, breaks down as the number of stimuli and responses increases because of parameter proliferation. The SAMBA absolute identification model, in contrast, requires no extra parameters to accommodate an increase in the number of stimuli and responses. Similar approaches to modeling inputs to decision processes in other paradigms would greatly simplify and expand the power of models of choice RT.

### References

- Berliner, J. E. (1973). *Intensity perception in audition*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.
- Berliner, J. E., Durlach, N. I., & Braida, L. D. (1977). Intensity perception: VII. Further data on roving-level discrimination and the resolution and bias edge effects. *Journal of the Acoustical Society of America*, *61*, 1577–1585.
- Braida, L. D., & Durlach, N. I. (1972). Intensity perception: II. Resolution in one-interval paradigms. *Journal of the Acoustical Society of America*, *51*, 483–502.
- Braida, L. D., Lim, J. S., Berliner, J. E., Durlach, N. I., Rabinowitz, W. M., & Purks, S. R. (1984). Intensity perception: XIII. Perceptual anchor model of context-coding. *Journal of the Acoustical Society of America*, *76*, 722–731.
- Brown, S., & Heathcote, A. (2003). Averaging learning curves across and within participants. *Behavior Research Methods, Instruments, and Computers*, *35*, 11–21.
- Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, *112*, 117–128.
- Brown, S., & Heathcote, A. (2007). *The ballistic accumulator: Efficient computation for any number of choices*. Manuscript submitted for publication.
- Brown, S., & Heathcote, A. (in press). The simplest complete model of choice reaction time: Linear ballistic accumulation. *Cognitive Psychology*.
- Brown, S., Marley, A. A. J., & Lacouture, Y. (2007). Is absolute identification always relative? Comment on Stewart, Brown, and Chater (2005). *Psychological Review*, *114*, 528–532.
- Busemeyer, J. R., & Townsend, J. T. (1992). Fundamental derivations for decision field theory. *Mathematical Social Sciences*, *23*, 255–282.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic cognition approach to decision theory. *Psychological Review*, *100*, 432–459.
- Chase, S., Bugnacki, P., Braida, L. D., & Durlach, N. I. (1983). Intensity perception: XII. Effect of presentation probability on absolute identification. *Journal of the Acoustical Society of America*, *73*, 279–284.
- DeCarlo, L. T. (1992). Intertrial interval and sequential effects in magnitude scaling. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 1080–1088.
- DeCarlo, L. T. (1994). A dynamic theory of proportional judgment: Context and judgment of length, heaviness, and roughness. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 372–381.
- DeCarlo, L. T., & Cross, D. V. (1990). Sequential effects in magnitude scaling: Models and theory. *Journal of Experimental Psychology: General*, *119*, 375–396.
- Durlach, N. I., & Braida, L. D. (1969). Intensity perception: I. Preliminary theory of intensity resolution. *Journal of the Acoustical Society of America*, *46*, 372–383.
- Garner, W. R. (1953). An informational analysis of absolute judgments of loudness. *Journal of Experimental Psychology*, *46*, 373–380.
- Gilden, D. L. (1997). Fluctuations in the time required for elementary decision making. *Psychological Science*, *8*, 296–301.
- Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological Review*, *108*, 33–56.
- Gilden, D. L., Thornton, T., & Mallon, M. W. (1995, March 24). 1/f noise in human cognition. *Science*, *267*, 1837–1839.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Holland, M. K., & Lockhead, G. R. (1968). Sequential effects in absolute judgments of loudness. *Perception & Psychophysics*, *3*, 409–414.
- Karpiuk, P., Jr., Lacouture, Y., & Marley, A. A. J. (1997). A limited capacity, wave equality, random walk model of absolute identification. In A. A. J. Marley (Ed.), *Choice, decision and measurement: Essays in honor of R. Duncan Luce* (pp. 279–299). Mahwah, NJ: Erlbaum.
- Kelly, A., Heathcote, A., Heath, R. A., & Longstaff, M. (2001). Response time dynamics: Evidence for linear and low-dimensional nonlinear

- structure in human choice sequences. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 54(A), 805–840.
- Kent, C., & Lamberts, L. (2005). An exemplar account of the bow and set size effects in absolute identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 289–305.
- Krantz, D. H. (1972). A theory of magnitude estimation and cross-modality matching. *Journal of Mathematical Psychology*, 9, 168–199.
- Lacouture, Y. (1997). Bow, range, and sequential effects in absolute identification: A response-time analysis. *Psychological Research*, 60, 121–133.
- Lacouture, Y., & Marley, A. A. J. (1991). A connectionist model of choice and reaction time in absolute identification. *Connection Science*, 3, 401–433.
- Lacouture, Y., & Marley, A. A. J. (1995). A mapping model of bow effects in absolute identification. *Journal of Mathematical Psychology*, 39, 383–395.
- Lacouture, Y., & Marley, A. A. J. (2004). Choice and response time processes in the identification and categorization of unidimensional stimuli. *Perception & Psychophysics*, 66, 1206–1226.
- Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.
- Laming, D. R. J. (1984). The relativity of “absolute” judgements. *British Journal of Mathematical and Statistical Psychology*, 37, 152–183.
- Link, S. W. (1992). *The wave theory of difference and similarity*. Hillsdale, NJ: Erlbaum.
- Lockhead, G. R. (2004). Absolute judgments are relative: A reinterpretation of some psychophysical ideas. *Review of General Psychology*, 8, 265–272.
- Lockhead, G. R., & Hinson, J. (1986). Range and sequence effects in judgment. *Perception & Psychophysics*, 40, 53–61.
- Lockhead, G. R., & King, M. C. (1983). A memory model of sequential effects in memory tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 461–473.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
- Luce, R. D., Nosofsky, R. M., Green, D. M., & Smith, A. F. (1982). The bow and sequential effects in absolute identification. *Perception & Psychophysics*, 32, 397–408.
- Marley, A. A. J., & Cook, V. T. (1984). A fixed rehearsal capacity interpretation of limits on absolute identification performance. *British Journal of Mathematical and Statistical Psychology*, 37, 136–151.
- Marley, A. A. J., & Cook, V. T. (1986). A limited capacity rehearsal model for psychological judgments applied to magnitude estimation. *Journal of Mathematical Psychology*, 30, 339–390.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for information processing. *Psychological Review*, 63, 81–97.
- Mori, S., & Ward, L. M. (1995). Pure feedback effects in absolute identification. *Perception & Psychophysics*, 57, 1065–1079.
- Neath, I., & Brown, G. D. A. (2006). Further applications of a local distinctiveness model of memory. *Psychology of Learning and Memory*, 46, 201–243.
- Nosofsky, R. M. (1983). Shifts of attention in the identification and discrimination of intensity. *Perception & Psychophysics*, 33, 103–112.
- Nosofsky, R. M. (1997). An exemplar-based random-walk model of speeded categorization and absolute judgment. In A. A. J. Marley (Ed.), *Choice, decision, and measurement* (pp. 347–365). Hillsdale, NJ: Erlbaum.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266–300.
- Petrov, A. A., & Anderson, J. R. (2005). The dynamics of scaling: A memory-based anchor model of category rating and absolute identification. *Psychological Review*, 112, 383–416.
- Pollack, I. (1952). The information of elementary auditory displays. *Journal of the Acoustical Society of America*, 24, 745–749.
- Pollack, I. (1953). The information of elementary auditory displays: II. *Journal of the Acoustical Society of America*, 25, 765–769.
- Purks, S. R., Callahan, D. J., Braida, L. D., & Durlach, N. I. (1980). Intensity perception: X. Effect of preceding stimulus on identification performance. *Journal of the Acoustical Society of America*, 67, 634–637.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review*, 88, 552–572.
- Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review*, 111, 159–182.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347–356.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, 111, 333–367.
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, 107, 358–367.
- Romani, G. L., Williamson, S. J., & Kaufman, L. (1982, June 18). Tonotopic organization of the human auditory cortex. *Science*, 216, 1339–1340.
- Rouder, J. N., Morey, R. D., Cowan, N., & Pfaltz, M. (2004). Learning in a unidimensional absolute identification task. *Psychonomic Bulletin & Review*, 11, 938–944.
- Shiffrin, R. M., & Nosofsky, R. M. (1994). Seven plus or minus two: A commentary on capacity limitations. *Psychological Review*, 101, 357–361.
- Smith, P. L. (1995). Psychophysically principled models of simple visual reaction time. *Psychological Review*, 102, 567–593.
- Smith, P. L., Ratcliff, R., & Wolfgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: Response time distributions with masked and unmasked displays. *Vision Research*, 44, 1297–1320.
- Stewart, N. (2007). Absolute identification is relative: A reply to Brown, Marley, and Lacouture (2007). *Psychological Review*, 114, 533–538.
- Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, 112, 881–911.
- Teghtsoonian, R. (1973). Range effects in psychophysical scaling and a revision of Steven’s law. *American Journal of Psychology*, 86, 3–27.
- Teghtsoonian, R., & Teghtsoonian, M. (1978). Range and regression effects in magnitude estimation. *Perception & Psychophysics*, 24, 305–314.
- Teghtsoonian, R., & Teghtsoonian, M. (1997). Range of acceptable stimulus intensities: An estimator of dynamic range for intensive perceptual continua. *Perception & Psychophysics*, 59, 721–728.
- Treisman, M., & Williams, T. C. (1984). A theory of criterion setting with an application to sequential dependencies. *Psychological Review*, 91, 68–111.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592.
- Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of  $1/f^n$  noise in human cognition. *Psychonomic Bulletin & Review*, 11, 579–615.
- Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2005). Human cognition and a pile of sand: A discussion on serial correlations and self-organized criticality. *Journal of Experimental Psychology: General*, 134, 108–116.
- Ward, L. M. (1987). Remembrance of sounds past: Memory and psychophysical scaling. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 216–227.
- Ward, L. M., & Lockhead, G. R. (1970). Sequential effect and memory in category judgment. *Journal of Experimental Psychology*, 84, 27–34.
- Ward, L. M., & Lockhead, G. R. (1971). Response system processes in absolute judgment. *Perception & Psychophysics*, 9, 73–78.
- Weber, D. L., Green, D. M., & Luce, R. D. (1977). Effects of practice and distribution of auditory signals on absolute identification. *Perception & Psychophysics*, 22, 223–231.
- Wiemer, J. C., & von Seelen, W. (2002). Topography from time-to-space transformations. *Neurocomputing*, 44–46, 1017–1022.

Received May 31, 2006

Revision received December 5, 2007

Accepted December 6, 2007 ■