

Dissociating speed and accuracy in absolute identification: the effect of unequal stimulus spacing

Christopher Donkin · Scott D. Brown ·
Andrew Heathcote · A. A. J. Marley

Received: 15 April 2008 / Accepted: 17 July 2008
© Springer-Verlag 2008

Abstract Identification accuracy for sets of perceptually discriminable stimuli ordered on a single dimension (e.g., line length) is remarkably low, indicating a fundamental limit on information processing capacity. This surprising limit has naturally led to a focus on measuring and modeling choice probability in absolute identification research. We show that choice response time (RT) results can enrich our understanding of absolute identification by investigating dissociation between RT and accuracy as a function of stimulus spacing. The dissociation is predicted by the SAMBA model of absolute identification (Brown, Marley, Dockin, & Heathcote, 2008), but cannot easily be accommodated by other theories. We show that SAMBA provides an accurate, parameter free, account of the dissociation that emerges from the architecture of the model and the physical attributes of the stimuli, rather than through numerical adjustment. This violation of the pervasive monotonic relationship between RT and accuracy has implications for model development, which are discussed.

Introduction

In many choice paradigms, more accurate responses are associated with faster response times, and vice versa: e.g., in Stroop-interference tasks (e.g., Kane & Engle, 2003;

Wuehr & Frings, 2008), various naming tasks (Duyck, Lagrou, Gevers, & Fias, 2008; Roelofs, 2006) and absolute identification (Kent & Lamberts, 2005; Lacouture & Marley, 1995, 2004; Petrov & Anderson, 2005). As a result, it might be thought that models predicting choice probability can also account for RT through a simple monotonic transformation (e.g., inversion). In this paper, we focus on the relationship between RT and accuracy in the absolute identification of unidimensional stimuli, where on each trial, participants identify a randomly chosen stimulus from a set of stimuli varying on only one dimension. For example, the stimulus set might consist of a set of ten lines of varying lengths which are given the labels #1 through #10 from shortest to longest.

We expand upon a previous finding that RT is not always a simple monotonic function of accuracy in absolute identification, when stimulus spacing is manipulated (Lacouture, 1997). We demonstrate the reliability of this result, and show that it provides a powerful test of different theoretical accounts of absolute identification. This dissociation is predicted by the SAMBA theory (Brown et al., 2008), which models both choice probability and choice response time. SAMBA predicts the violation of the pervasive negative correlation between RT and accuracy because of one of its components—the mapping model developed by Lacouture and Marley (1995). We show that SAMBA accounts for the dissociation without parameter adjustment because its account emerges from the architecture of the model and the physical attributes of the stimuli.

There are numerous benchmark phenomena for the absolute identification of unidimensional stimuli, when the stimuli are equally spaced. For instance, when mean RT and accuracy are plotted as functions of each stimulus' ordinal position within the set, one observes the ubiquitous

C. Donkin (✉) · S. D. Brown · A. Heathcote
School of Psychology, University of Newcastle, Callaghan,
Newcastle, NSW 2308, Australia
e-mail: christopher.donkin@newcastle.edu.au

A. A. J. Marley
University of Victoria, Victoria, BC, Canada

“bow effect”—a U-shape for accuracy and an inverted U for RT. In such plots, stimuli associated with shorter RTs are always associated with higher accuracy, and vice versa. The inverse relationship between accuracy and RT is also observed in many other kinds of plots of absolute identification data, for example, if accuracy and mean RT are plotted as functions of the difference between successive stimuli; the number of trials that have intervened since the current stimulus was last presented; or the number of stimuli within the set (Brown et al., 2008; Kent & Lamberts, 2005; Lacouture & Marley, 1995, 2004).

Brown et al. (2008) developed SAMBA to account for the benchmark empirical choice and RT phenomena. SAMBA was intended to be a complete account of absolute identification, including all stages from a psychophysical stimulus representation through to response selection, and modeling all of the important benchmark phenomena from the field. These benchmark phenomena in absolute identification can be divided into two types: global and local effects. Local effects refer to the influence of previous stimuli and responses on the current decision, whereas global effects are the phenomena observed irrespective of sequential dependencies. Stewart, Brown, and Chater (2005) give a good summary of existing models of absolute identification and their respective ability to capture the important empirical phenomena. They demonstrated that their Relative Judgment Model (RJM) was capable of capturing both global and local benchmark choice phenomena (they did not model RT), whereas other models only capture either local (e.g., Lockhead, 2004) or global effects (e.g. Lacouture & Marley, 2004). Brown et al. showed that SAMBA was capable of predicting both global and local benchmark phenomena, not only in choice, but also in RT.

It could be argued that the ability to account for RT in absolute identification is not a large advantage over other models that predict choice phenomena because RT effects are simply the inverse of accuracy. If this were the case, it may be reasoned that any satisfactory model of choice in absolute identification could obtain an equally good account of RT effects in this domain by inverting accuracy. However, Brown et al. (2008) also pointed out that SAMBA makes the surprising and testable prediction that a larger relative spacing between two adjacent stimuli in the stimulus set will result in increased accuracy for those two stimuli, but will have little impact on RT. In the following sections we first provide a brief overview of SAMBA, followed by a detailed account of this particular prediction. We then examine data from Lacouture (1997) that confirm the prediction, and show SAMBA’s fit to the data. We conclude by discussing the implications of these results for theoretical development in the field of absolute identification.

SAMBA

SAMBA (Brown et al., 2008) is composed of three stages: a selective attention stage, a mapping stage and a decision stage. The selective attention stage begins with an impoverished psychophysical representation—one without an associated numerical magnitude—of a stimulus and produces an estimate of its magnitude. This estimate is constructed using Marley and Cook’s (1984, 1986) selective attention theory, which posits that stimulus magnitudes are judged relative to a context defined by upper and lower “anchors” (call these L and U), which are placed beyond the smallest and largest stimuli. On each trial, the magnitude of the stimulus is judged relative to the overall context determined by the interval $[L, U]$, and the estimate is noisy due to the selective attention process. This magnitude estimate falls in the interval $[0, 1]$, and the average magnitude estimate for any particular stimulus (over repeated presentations) is given by the linear function which maps the interval $[L, U]$ onto the interval $[0, 1]$. For example, in an experiment with ten equally spaced stimuli, if stimulus #5 were presented, a magnitude estimate of close 0.45 might be expected, but on any given trial the estimate will vary somewhat from this average.

SAMBA’s mapping stage transforms the magnitude estimate produced by the selective attention stage into response strengths, one for each possible response. The mapping operates like a highly constrained set of tuning curves, where each curve produces a response strength that depends on how closely the observed magnitude estimate matches a referent for the given response. SAMBA assumes the referents are obtained by averaging magnitude estimates associated with repeated presentations of each stimulus. The mapping phase operates similarly to all tuning curve systems, in that the largest response strength is always assigned to the response whose referent most closely matches the observed magnitude estimate. For example, a magnitude estimate of 0.45 might be closest to the long term referent for stimulus #5 and so the largest response strength will be assigned to response #5. The outputs from the mapping phase are the inputs to SAMBA’s decision stage, which consists of a set of ballistic accumulators, one for each possible response (Brown & Heathcote, 2005). The ballistic accumulators instantiate a noisy max-picking algorithm. The chosen response will usually be the one with largest response strength (from the mapping stage), but not always. Larger response strengths are associated with faster responses, and bigger differences between response strengths are associated with more accurate “pick-the-max” behaviour.

The dynamics of SAMBA’s basic architectural elements also account for sequential effects. For example, activity in the decision phase ballistic accumulators decays slowly

between trials. Amongst other things, this means that the response selected on the previous trial will have an advantage on the current trial, as observed in data. However, for our purposes the critical element of SAMBA is the bow mapping phase. Lacouture and Marley (1995) developed the mapping from a theoretical viewpoint. That is, they started out with a list of mathematical properties that any reasonable set of tuning curves should have. For example, for any reasonable set of tuning curves the greatest response strength should always be assigned to the response whose long-term referent most closely matches the incoming magnitude estimate. Obviously, a great variety of tuning curves would satisfy this property, so other properties were included to constrain and simplify the solution, including: all response strengths should always be positive; the tuning curves should use the simplest functional form possible—a straight line; and the set of curves should be symmetric, as long as the referents are symmetric in the interval $[L,U]$. Lacouture and Marley developed their bow mapping as a set of linear tuning “curves” which satisfied all these constraints. Their solution is very parsimonious because it is also parameter free, being entirely specified by the values of the long term referents for each stimulus’ average magnitude estimate. The mapping solution also predicted the ubiquitous bow effects observed in both response time and accuracy for absolute identification, even though these properties were not included as constraints for its development.

Of course, other solutions to the basic tuning curve problem could be developed. In particular, one may relax the simplifying constraints imposed by Lacouture and Marley (1995), by allowing more complex nonlinear forms for the tuning curves. Some such solutions would probably also be able to accommodate the accuracy-RT dissociation discussed below, but it is difficult to justify their extra complexity. From this point of view, one may consider the bow mapping as the simplest and most constrained set of tuning curves available, with the added benefit that they allow SAMBA to accommodate not only the empirical data addressed by Brown et al. (2008), but also to make testable predictions.

SAMBA’s predictions for unequally spaced stimuli

SAMBA makes the prediction that if the spacing between two adjacent stimuli is increased, with other stimulus spacings unchanged, then these particular stimuli are identified with higher accuracy, but RT is relatively unaffected. This prediction is a consequence of Lacouture and Marley’s (1995) mapping solution. When stimuli are unequally spaced, the long-term referents (average magnitude estimates) that define the mapping stage will reflect the unequal spacing. For example, first consider a standard

absolute identification experiment with 10 equally spaced stimuli, and suppose that participants place their lower and upper anchors at a distance equivalent to one stimulus separation above and below the stimuli at the upper and lower end of the range, respectively. In this case, the selective attention phase of SAMBA produces average magnitude estimates given by the linear mapping of the stimulus magnitudes onto the unit interval, namely: $\{\frac{1}{11}, \frac{2}{11}, \dots, \frac{10}{11}\}$. Now imagine that a set of ten unequally spaced stimuli is constructed by first taking 14 equally spaced stimuli, then removing the central four. This stimulus set has a large central gap between stimuli #5 and #6. The selective attention phase of SAMBA then produces average magnitude estimates that respect the unequal stimulus spacing, namely $\{\frac{1}{15}, \frac{2}{15}, \dots, \frac{5}{15}, \frac{10}{15}, \dots, \frac{14}{15}\}$. Since the average estimates define the mapping solution, the spacing of the stimulus set is naturally encoded into the operation of the model.

On each trial of an absolute identification experiment, the selective attention phase produces a noisy magnitude estimate, say z . The mapping solution transforms this estimate into a response strength R_j for each of the j possible responses according to the formula $R_j = (2Y_j - 1)z - Y_j^2 + 1$, which is linear in the magnitude estimate, z . The function is completely defined by Y_j , which is the average magnitude estimate for the j th stimulus—the long term referent for response j . Figure 1 illustrates the mapping solutions that arise from both the equally spaced and unequally spaced sets of ten stimuli. In both cases, the mapping supports the basic property, that if the observed magnitude estimate is close to the average magnitude estimate for stimulus j , the highest response strength will be assigned to response j . For example, suppose on a particular trial stimulus #5 is presented, and the selective attention phase produces a magnitude estimate of 0.45 units. In both the equal and unequal spacing cases, the large black dot shows that the greatest response strength in this case is assigned to response #5 (i.e., the highest line

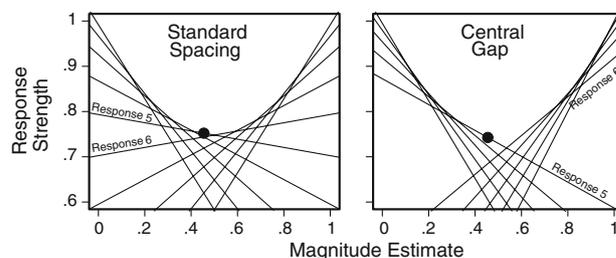


Fig. 1 Mapping solution for equally spaced stimuli (*left panel*) and a set of ten stimuli with a central gap in stimulus spacing equivalent to four stimuli (*right panel*). Each *line* shows how the response strength varies with input magnitude estimate, for one of the ten possible responses. The mapping is from $[0,1]$ to $[0,1]$, with the ordinate truncated here for display purposes

above $x = 0.45$ is the one corresponding to the fifth response). It may strike the reader as surprising, on first glance, that the greatest response strength for each response always occurs at one extreme or the other ($x = 0$ or 1). For example, response #5 is the maximum-strength response at $x = 0.45$, but the *greatest* response strength is assigned to response #5 when $x = 0$. This property arises from the severe simplifying constraints imposed by Lacouture and Marley (1995), in particular that the tuning curves should be linear. It is testament to the power of their solution that it still fits the data so well, even with such constraint.

The difference in the mapping solutions for equally and unequally spaced stimuli leads to the prediction that is our focus. Consider again stimulus #5, which is adjacent to the large central gap in the unequal stimulus set (a similar argument applies to stimulus #6). In the equally spaced condition, response #5 is the maximum-strength response¹ for any magnitude estimates in the interval $z \in [\frac{4.5}{11}, \frac{5.5}{11}]$. However, in the unequally spaced condition, response #5 is the maximum-strength response for a larger range of magnitude estimates: $z \in [\frac{4.5}{15}, \frac{7.5}{15}]$. Given that the response with the largest strength is usually the response made by SAMBA's decision phase, accuracy is predicted to be higher for stimulus #5 in the unequally spaced condition than in the equally spaced condition. This is mostly due to the prediction that stimulus #5 will not often be confused with stimulus #6 (and vice versa). For example, for the unequally spaced stimuli in Fig. 1, the average magnitude estimate for stimulus #5 is $\frac{5}{15}$. Due to the properties of the rehearsal stage, it is rare that if stimulus #5 were presented that the magnitude estimate would be greater than $\frac{7.5}{15}$, and hence fall in the region where response #6 would receive the largest response strength.

Turning now to predictions for response time, Fig. 1 shows that the size of the response strength produced for stimulus #5 is about the same in both the equal and unequal stimulus spacing conditions. In the decision stage of SAMBA, the response strength for response #5 determines the rate of increase of activation in the corresponding ballistic accumulator. All other parameters being equal, response time is inversely related to the response strength, so SAMBA predicts about the same response times for stimulus #5 in both the equal and unequal spacing conditions. To be numerically precise, the average magnitude estimate associated with stimulus #5 in the equally spaced condition is $\frac{5}{11}$, and this results in a maximum response strength being assigned to response #5, a strength of $(2\frac{5}{11} - 1)\frac{5}{11} - (\frac{5}{11})^2 + 1 = 0.752$. In the unequally spaced

case, stimulus #5 generates an average magnitude estimate of $\frac{5}{15}$, or $\frac{1}{3}$, but again the maximum response strength is assigned to response #5, $(2\frac{1}{3} - 1)\frac{1}{3} - (\frac{1}{3})^2 + 1 = 0.778$. Critically, in both cases, the response strengths assigned to the correct response for neighbouring stimuli is larger, so they will have a faster RT. For example, when stimulus #4 is presented, the average response strength assigned to response #4 is 0.769 in the equally spaced condition and 0.804 in the unequally spaced condition. Hence, SAMBA predicts a dissociation that responses to stimuli near a large gap will be more accurate than for neighbouring stimuli, but the corresponding response times will be slower.

Empirical evidence

Several researchers have manipulated stimulus spacing, including Lockhead and Hinson (1986) and Lacouture (1997). Brown et al. (2008) demonstrated that SAMBA provides a parsimonious account of the choice probabilities reported by Lockhead and Hinson (RTs were not recorded). Lacouture's data set included RT measurements, allowing us to test SAMBA's predictions about the effect of stimulus spacing on both choice probabilities and RT. Participants in Lacouture's experiment spent the first hour in a standard absolute identification experiment, with ten equally spaced stimuli. Each participant then spent a second hour in one of several conditions in which physical properties of the stimuli were manipulated. Brown et al. presented fits of SAMBA to data from the first session (equal spacing), but until now the unequal spacing conditions have never been modeled. Since Lacouture published his findings, several important, integrative theories of absolute identification have been published, some of which have even addressed the effects of unequal stimulus spacing on response choices, but none have addressed the effects of stimulus spacing on response times. This leaves a gap in theoretical development, especially because, as we now show, Lacouture's data presents a challenging test for models.

Participants in the second session of Lacouture's (1997) experiment experienced one of six conditions, four of which employed unequally spaced stimuli. These four conditions had larger gaps introduced either in the centre (between stimuli #5 and #6) or at the edges (between stimuli #2 and #3 and stimuli #8 and #9). The gap location was crossed with a manipulation of gap size (large or small) to create the four conditions: a large central gap (C-L); a small central gap (C-S); large extreme gaps (E-L); and small extreme-gaps (E-S). The top row of Fig. 2 provides a schematic illustration of the stimuli from these four conditions (for actual stimulus lengths, see Lacouture's Table 1). The second and third rows of Fig. 2 show

¹ It is elementary to show that responses j and $j + 1$ have equal response strengths at the point that is midway between the long term referents for stimuli j and $j + 1$, and this holds for both equally spaced and unequally spaced stimulus sets.

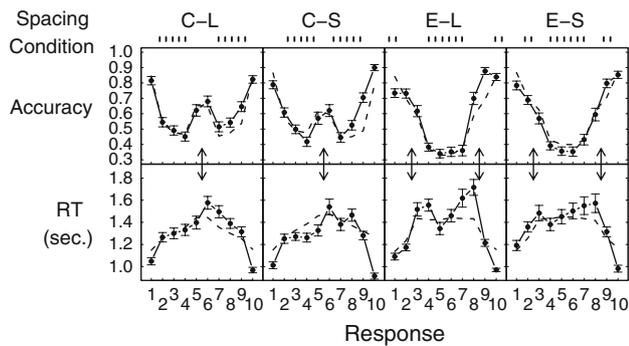


Fig. 2 The *top* row shows a schematic representation of the stimuli used in Lacouture’s (1997) second session. *C–L* refers to the ‘large central-gap’ and *C–S* to the ‘small central-gap’ condition, *E–L* refers to the ‘large extreme-gaps’ condition, and *E–S* ‘small extreme-gaps’. The *second* row shows response accuracy and the *third* row shows mean RT for correct responses, both as functions of response. Data are shown as points with *standard error bars* that are joined by *solid lines*, and SAMBA’s fits are shown with *dotted lines*

data from the four spacing conditions, replicating Lacouture’s Figure 4. The data are represented by solid circles with ± 1 standard error bars, calculated assuming normal distributions across subjects for mean RT and binomial distributions for accuracy. In each graph, vertical arrows show the locations of the larger gaps, and the dashed lines show predictions generated by SAMBA. The second row of Fig. 2 shows response accuracy separately for each response and the third row shows mean correct response times. Notice that accuracy is greater for stimuli adjacent to gaps, and the effect is more pronounced in the large spacing conditions than the small spacing conditions. However, the improved accuracy is never accompanied by faster response times relative to neighbouring stimuli, contrary to the typical inverse RT-accuracy relationship.

When analyzing response choices from absolute identification tasks, it is customary to calculate sensitivity (d') instead of raw percent correct (Luce, Nosofsky, Green, & Smith, 1982). Sensitivity provides a bias-free measure of how often successive pairs of stimuli are confused, that is, how often stimuli #1 and #2 are confused, and stimuli #2 and #3, and so on up to stimuli #9 and #10. For any given pair, say stimuli #4 and #5, d' is calculated in the usual manner, using hit and false alarm rates, where ‘hits’ are defined as responses #5 or greater, when stimulus #5 is presented, and ‘false alarms’ are defined as responses #5 or greater, when stimulus #4 is presented. To ensure that the effects observed in Lacouture’s (1997) data were not due to a response bias effect, we calculated d' values for each stimulus pair, shown in Fig. 3. Graphing the data using d' shows an even more pronounced effect of stimulus spacing—stimuli that are separated by large gaps were almost never confused with one another.

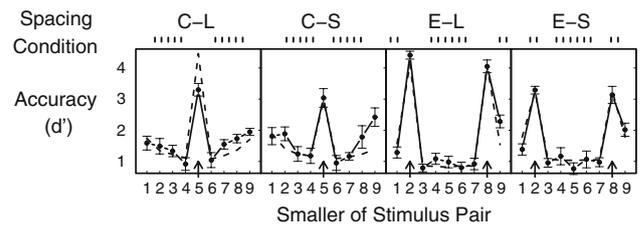


Fig. 3 Sensitivity (d') for each stimulus pair in the four unequally spaced stimulus conditions from Lacouture’s (1997) second session. Data are shown as points with *standard error bars* (calculated assuming d' is normally distributed across participants) that are joined by *solid lines*, and SAMBA’s fits are shown with *dotted lines*

Figures 2 and 3 demonstrate a clear dissociation—stimuli separated by large gaps enjoy an accuracy (and sensitivity) bonus, but no corresponding RT bonus. To confirm the statistical reliability of this dissociation, we calculated binomial tests. We used binomial tests because they provide robust analyses that directly test the ordinal hypotheses we entertain, without potentially problematic distributional assumptions. We carried out two tests, one for the dissociation of response times and raw response accuracy and the other for the dissociation of response times and d' . We examined the accuracy (or d') and RT values separately for each participant, and counted how frequently the dissociation in question was observed on a single-participant basis—that is, how often we observed improved accuracy for stimuli on either side of the larger gaps (relative to neighbouring stimuli) without observing faster RT for those same stimuli. Take, for example, one of the central gap conditions (*C–L* or *C–S*) in which stimuli #5 and #6 are near the gap. For any single participants’ data, we would say we observed the dissociation whenever we found four different ordinal constraints to be satisfied, namely that:

- Response time was not smaller for stimulus #5 than its neighbour (#4).
- Response time was not smaller for stimulus #6 than its neighbour (#7).
- Accuracy was greater for stimulus #5 than its neighbour (#4).
- Accuracy was greater for stimulus #6 than its neighbour (#7).

For the extreme gap conditions (*E–L* and *E–S*) the same ordering constraints were tested twice for each participant, once for the gap between stimuli #2 and #3, and once for the gap between stimuli #8 and #9. The dissociation between d' and RT was assessed slightly differently, since d' measures the discriminability of each pair of stimuli. Thus, constraints (a) and (b) from the above list remained the same, but constraints (c) and (d) were replaced with:

- c. Discriminability of stimuli #5 and #6 was greater than for stimuli #4 and #5.
- d. Discriminability of stimuli #5 and #6 was greater than for stimuli #6 and #7.

For each of the 16 participants in the E–S and E–L conditions, there were two opportunities to observe the dissociation—corresponding to the two stimulus gaps in each condition—and for each of the other 16 participants in the C–S and C–L conditions there was one opportunity. For the raw accuracy data, under a null hypothesis of no relationship, the probability of observing the dissociation by chance is one in 16 at each opportunity, but we observed the dissociation ten out of 48 times, significantly more than the three that would be expected by chance ($p = 0.0015$). For the d' data, the probability of observing the dissociative ordering by chance is one in 12, but we observed the dissociation 20 out of 48 times, again significantly more than would be expected by chance ($p < 10^{-10}$). These tests are quite convincing, especially given there were only eight participants in each condition and the reduced power afforded by robust non-parametric statistical tests. Note that those participants who did not demonstrate the critical dissociation on single-participant level did not necessarily demonstrate the opposite (i.e., the usual inverse accuracy-RT relationship). In fact, of the 96 opportunities to observe the usual inverse relationship in these critical tests, we observed only one instance where stimuli adjacent to a gap exhibited both increased accuracy and a decrease in RT relative to their neighbouring stimuli—for the other 95 opportunities, we either observed random ordering due to noise (65 times) or the accuracy-RT dissociations counted above (30 times).

Turning now to predictions from SAMBA, we can see from the dashed lines in Figs. 2 and 3 that the model provides a good qualitative account of the data, capturing the observed dissociation between accuracy and RT. The model also provides a very close quantitative fit to the data, which is all the more surprising given the strong constraints we imposed on the parameters. To fit SAMBA to Lacouture's (1997) unequal spacing conditions, we began with the parameters reported by Brown et al. (2008) that were used to fit the standard (equal spacing) condition from the first session of Lacouture's experiment. Only three parameters were adjusted for the fits presented in Fig. 2, and even these parameters were irrelevant in capturing the critical dissociation between accuracy and RT—all three parameters were instead related to the effects of practice, capturing differences in the data between the first and second experimental sessions. Firstly, we decreased the response threshold parameter for SAMBA's decision phase to be 90% of the value it took for the first experimental session, reflecting that participants may have become a

little less careful in the second session of the experiment. Secondly, we had previously noted asymmetry in the data: in the first experimental session, responses were slower and less accurate for the large stimuli than the small stimuli. SAMBA accounted for the asymmetry by setting the lower anchor close to the smallest stimulus (L was set at 95% of the magnitude of the smallest stimulus) but the upper anchor quite far away from the largest stimulus (U was 62% larger than the largest stimulus). In the second experimental session the asymmetry disappeared: note that in Fig. 2 accuracy and RT are about the same for the smaller stimuli as for the larger stimuli. To capture this return to symmetry, we set the lower anchor to 99% of the magnitude of the smallest stimulus, and the upper anchor to 101% of the magnitude of the largest stimulus.

One possible interpretation for the change in symmetry between sessions could be improvement due to practice. Although absolute identification is mostly immune to practice effects, Rouder, Morey, Cowan, and Pfaltz (2004) showed that learning in absolute identification is possible. Donkin, Dodds, Brown and Heathcote (2008) have shown that this is especially true for lines of varying length, as used by Lacouture. This explanation is consistent with the change in parameters of SAMBA used to achieve the reported fits. The upper and lower anchors, U and L , were moved closer to stimuli #1 and #10 in the unequal spacing conditions, indicating that participants improved their knowledge of the task in the second session relative to the first.

It is notable that exactly the same parameters were used to fit all four different spacing conditions. The differences between spacing conditions are completely determined by the properties of the stimulus spacing, which in turn determine the referents. For example, referents for the large central-gap condition are based on long-term averages of magnitude estimates produced by SAMBA's selective attention phase, and these magnitude estimates naturally reflect the large gap between stimuli #5 and #6. The same mechanism applies to the other stimulus spacing conditions. Given the constraints we imposed on the model parameters, the quantitative fits are quite good, although SAMBA overpredicts the improvement in d' in the C–L condition. This misfit is quite small in terms of probability, but is exaggerated by the inverse cumulative normal transformation used to calculate d' when accuracy is high.

Alternative models

There are four recent models of absolute identification, besides SAMBA, that make predictions for both choice and RT. Two of these models are exemplar based accounts of general categorization behavior, applied to absolute

identification, which can be seen as a special case of categorization. These two models are the exemplar-based random walk (EBRW: Nosofsky, 1997; Nosofsky & Palmeri, 1997) and the extended generalized context model for response times (EGCM-RT: Kent & Lamberts, 2005, Lamberts, 2000). Both models predict that increased accuracy should always be associated with faster RT, at least when parameters unrelated to stimulus properties are kept constant. In Lacouture's (1997) data, the dissociation of RT and accuracy was observed within blocks in which only stimulus magnitude was manipulated, so it would seem that these theories are incapable of accounting for the dissociation between accuracy and RT with unequal stimulus spacing (in particular, see Equation 5 in Nosofsky, 1997, and Equation 12 in Lamberts, 2000). The EBRW and EGCM-RT both predict the observed increase in accuracy with increased spacing between stimuli, caused by reduced similarity between stimulus representations. However, both models also predict an associated decrease in RT, which was not observed in Lacouture's data. It is possible that, with carefully chosen parameter values, these models could decrease the size of the predicted misfit. That is, there may exist parameter values that allow the models to predict increased accuracy near large gaps, accompanied by only a small decrease in RT for those same stimuli. Even if these parameter values exist, the models still make incorrect predictions about the (statistically reliable) ordering of the data values observed above.

Another absolute identification model that predicts RT is Karpiuk, Lacouture, and Marley's (1997) limited capacity, wave equality, random-walk model. This model is similar to SAMBA, in that it uses Marley and Cook's (1984) rehearsal model, but in place of SAMBA's mapping stage, Karpiuk et al. used a set of tuning curves for each response. Tuning curves, specified by free parameters, operate like SAMBA's mapping stage but with less constraint and greater flexibility. For this reason, it is quite likely that Karpiuk et al.'s model is capable of capturing (but not predicting) the observed dissociation between RT and accuracy. Lacouture and Marley's (1995, 2004) mapping model employs the same mapping functions as SAMBA, and so it also predicts the dissociation between RT and accuracy.

Ashby (2000) developed a theory of categorization that includes predictions for RT as well as choices. Other versions of this theory have been applied to absolute identification data (Ashby & Lee, 1991), although the RT-inclusive version has not. Similarly to the exemplar-based categorization models, Ashby's theory generally predicts a monotonic relationship between mean RT and accuracy in categorization (see e.g., Ashby, 2000, p. 321 for a summary of the extensive successes, and limited failures, of this prediction) and, therefore, does not accommodate the

observed dissociation. Ashby's model "sometimes predicts violations of the RT-distance hypothesis" (p. 322), but under complex assumptions that are unlikely to be satisfied in absolute identification data. We also note that extant absolute identification models (other than SAMBA) that predict both accuracy and RT fail to predict other key phenomena. For example, none of the models described above predict the well-known sequential effects in absolute identification data, such as assimilation and contrast.

Discussion

We have presented and tested a prediction arising from the mapping stage of the SAMBA model of absolute identification (Brown et al., 2008). The predicted dissociation between accuracy and RT is surprising due to the regularity with which a monotonic relationship has been observed in a range of speeded choice paradigms. Nevertheless, SAMBA's prediction was confirmed by data from Lacouture's (1997) unequal spacing experiments, data which has not previously been accounted for by any model. SAMBA predicts the dissociation between RT and accuracy under different spacing conditions, and provides an impressive quantitative fit, given that no parameter changes were made between conditions, and all but three parameters were fixed at values estimated using data from a different condition.

In most empirical and theoretical work on absolute identification, response times have received much less attention than response choices. Despite an empirical research history pre-dating Miller's (1956) seminal review, and theoretical accounts existing for at least 50 years, models have only begun to address RT in the last 15 years. The disinterest in RT is underlined in Stewart et al.'s (2005) model summary table (p. 886), where only three out of the 14 models reviewed made predictions about RT. This neglect is most likely due to the belief that RT has little utility for discriminating models, which might have been true if a systematic monotonic inverse relationship between RT and accuracy always held. However, Lacouture's (1997) results show that this is not the case, and that RT and accuracy data together provide greater model constraint than accuracy data alone. In particular, Lacouture's data provide a strong test for any theoretical account of absolute identification that attempts to account for both choice and RT. SAMBA passes this test, confirming a prediction made by Lacouture and Marley's (1995) highly constrained method of obtaining tuning curves, which was adopted by SAMBA. Hence, Lacouture and Marley's method, motivated on entirely independent theoretical grounds, not only predicts the ubiquitous bow effects found in absolute identification, but also a heretofore unexplored dissociation between speed and accuracy.

Appendix: The latencies of incorrect responses

The relative speeds of correct and incorrect responses have proven very illuminating in the development of theories of choice response time (see, e.g., Brown & Heathcote, 2005, 2008). Theoretical accounts of response times in absolute identification are less well developed, so the fine model discrimination afforded by the analysis of error RT may yet be premature. Nevertheless, we note here two interesting phenomena related to incorrect RTs in Lacouture's (1997) data. Firstly, response times were slightly, but reliably, slower for incorrect responses than correct responses in the unequal spacing conditions of Lacouture's experiment (mean difference 29 ms, $t(38) = 2.5$, $p < 0.01$). Secondly, the relative speed of correct and incorrect responses changed systematically with stimulus magnitude. For extreme stimuli (#1 and #10), incorrect responses were much slower than correct responses (mean difference 269 ms, $t(46) = 7.5$, $p < 0.001$) but for central stimuli (#5 and #6) there was almost no difference (mean difference 10 ms, $t(46) = 0.3$, $p > 0.05$). The relative speeds of correct and incorrect responses are captured well by SAMBA—a brief illustration, Fig. 4 shows mean error response times along with SAMBA's predictions using the same format as Fig. 2. The model captures the global qualitative trends in the data, but misses some of the finer quantitative properties, such as the tendency for some extreme responses to be associated with very fast errors (e.g., #1 in C-L condition and #10 in E-S and C-S). It also fails to capture a tendency for participants in the central gap conditions (C-L and C-S) to make fast errors when responding with #5. The reader might suppose that these faster errors are due to stimulus #5 lying adjacent to a gap. However, this does not explain why the same pattern is not shown for response #6 or in the E-L or E-S conditions.

We do not take the observed goodness of fit to be as impressive as SAMBA's ability to fit our main focus, the effects of stimulus spacing. Although the patterns of fast

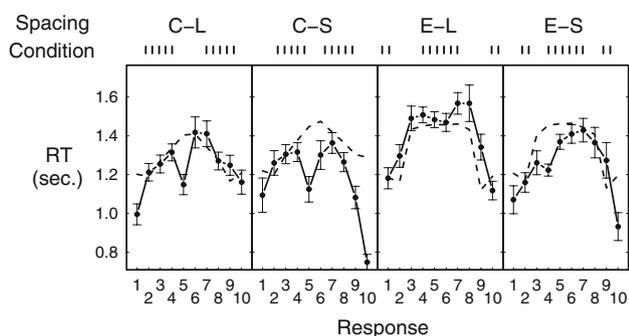


Fig. 4 Mean response times for incorrect responses, along with predictions from SAMBA. Error bars show ± 1 standard error assuming that mean RT is normally distributed across participants

and slow errors may appear complex at first glance, they are less theoretically challenging than might be imagined. For example, incorrect response times were slower than correct response times, as predicted by SAMBA. Other models of absolute identification do not predict this in their current forms. For example, Kent and Lamberts' (2005) model uses a random walk, which is constrained to predict equal response times for correct and incorrect responses (see e.g., Ratcliff, 1978). However, this limitation is not central to Kent and Lamberts' model, and can easily be remedied by the addition of certain variance components to its decision phase (as described by Ratcliff). Similarly, any model of absolute identification that predicts the ubiquitous bow effects—longer RT for central responses and shorter RT for extreme responses—will necessarily produce slower mean error than correct RTs for extreme stimuli. This is because those incorrect responses are less extreme (usually #2 and #9, rather than #1 and #10, for example) and hence slower due to the bow effect. For these reasons, we think that a detailed comparison of empirical results with theoretical predictions for incorrect response times may be premature for models of absolute identification.

References

- Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, 44, 310–329.
- Ashby, F. G., & Lee, W. W. (1991). Predicting similarity and categorization from identification. *Journal of Experimental Psychology: General*, 120, 150–172.
- Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, 112(1), 117–128.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice reaction time: linear ballistic accumulation. *Cognitive Psychology*.
- Brown, S., Marley, A. A. J., Donkin, C., & Heathcote, A. (2008). An integrated architecture for absolute identification. *Psychological Review*, 115, 396–425.
- Donkin, C., Dodds, P., Brown, S. D., & Heathcote, A. (2008). Revisiting the Limits of Human Information Processing Capacity. (submitted).
- Duyck, W., Lagrou, E., Gevers, W., & Fias, W. (2008). Roman digit naming: evidence for a semantic route. *Experimental Psychology*, 55, 73–81.
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to stroop interference. *Journal of Experimental Psychology*, 132, 47–70.
- 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 honour of R. Duncan Luce* (pp. 279–299). Mahwah: Erlbaum.
- 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.
- 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. (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.
- Lamberts, K. (2000). Information-accumulation theory of speeded categorization. *Psychological Review*, *107*, 227–260.
- 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.
- 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.
- 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: essays in honor of R. Duncan Luce* (pp. 347–365). Mahwah: Erlbaum.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded categorization. *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.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Roelofs, A. (2006). Functional architecture of naming dice, digits, and number words. *Language and Cognitive Processes*, *21*, 78–111.
- Rouder, J. N., Morey, R. D., Cowan, N., & Pfaltz, M. (2004). Learning in a unidimensional absolute identification task. *Psychonomic Bulletin & Review*, *11*, 938–944.
- Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, *112*, 881–911.
- Wuehr, P., & Frings, C. (2008). A Case for Inhibition: visual attention suppresses the processing of irrelevant objects. *Journal of Experimental Psychology: General*, *137*, 116–130.