Absolute Identification is Surprisingly Faster with More Closely Spaced Stimuli

James S. Adelman (J.S.Adelman@warwick.ac.uk)
Neil Stewart (Neil.Stewart@warwick.ac.uk)
Department of Psychology, University of Warwick,
Gibbet Hill Road, Coventry, CV4 7AL, UK

Abstract

Bow and set size effects on response times in absolute identification mirror the effects on accuracy: Central stimuli and stimuli in large sets are responded to more slowly and less accurately. In an analysis of the response time data from Experiment 1 of N. Stewart, G. D. A. Brown and N. Chater (2005), involving the absolute identification of tone frequency (pitch), we find that in contrast to the accuracy data, where widely spaced stimuli are responded to slightly more accurately than narrowly spaced stimuli, widely spaced stimuli receive slower responses than narrowly spaced stimuli. This result poses an additional challenge for models of absolute identification, as accuracy and response times are not jointly linked to some unified difficulty factor.

Introduction

In absolute identification, participants learn unique associations between stimuli that vary on a simple perceptual dimension (e.g., pitch) and ordered numerical responses. Behavioral data that may inform us about the representations and processes involved in the performance of this task come in the form of the responses (i.e., their accuracy) and the time they take. Absolute identification is particularly intriguing for psychophysicist, given the surprising finding that although observers can discriminate perfectly between stimuli drawn from the set in pairwise comparisons, they make a great deal of error in identifying the same stimuli. Further, the limit in identifying stimuli holds across a wide range of sensory continua - such as tones varying in pitch, lines varying in length, smells varying in intensity, tastes varying in their saltiness, and cutaneous electric shocks varying in their intensity - suggesting that the limit is a fundamental property of the cognitive system (see Stewart, Brown, & Chater, 2005, for a recent review).

Key effects in the accuracy data are the bow effect, that stimuli nearer the center of the range are responded to less accurately; the set size effect, that accuracy is reduced when there are more stimuli (and hence responses) in the experimental set; and sequential effects — assimilation to recent responses and contrast to less recent responses (again see Stewart et al., 2005). At least with regard to the bow and set size effects, response times mirror the accuracy data (in terms of a difficulty interpretation): Central stimuli and stimuli in large sets are responded to slower (e.g. Karpiuk, Lacouture, & Marley, 1997; Kent & Lamberts, 2005).

Models have been developed in attempts to give integrated explanations of accuracy and response time components of absolute identification performance, and with surprising apparent difficulty. Karpiuk et al. (1997), for instance, linked a rehearsal-based limited capacity model to a counter system for this purpose, but to simulate differences between set sizes, the parameters needed to be modified from set size to set size, and so this model does not intrinsically explain set size effects. Nosofsky (1997) linked an exemplar-similarity system to counters (as in Nosofsky & Palmieri, 1997); although this model produced basically correct effects, it did not account well for effects of set size on edge stimuli.

An alternative exemplar-similarity model, based on the EGCM-RT (Lamberts, 2000), was produced by Kent and Lamberts (2005). With a small modification to the original assumptions about information accumulation of elements of features, this model was able to account for differences in set size with completely common assumptions and parameters. The basic operation of the model proceeds by the sampling of elements from the stimulus dimension at random intervals and, with increased sampling of elements, similarity to the presented stimulus of stored exemplars that mismatch this stimulus decreases. A response may be made upon information sampling, depending probabilistically on the equivocality of the current exemplar-stimulus similarities. Should a response be made, the choice of the response is based on the choice rule (Shepard, 1957; Luce, 1963).

The relative judgment model (RJM) of Stewart et al. (2005) differs greatly in operational principles — it assumes that differences to the previous stimulus are the main source of information in absolute identification, and 'confusions' are not (primarily) due to similarity (between stimuli), but rather limited capacity on the response scale — but it also fits accuracy data from different set sizes without an arbitrary change in parameters (although the model has yet to be extended to response time data).

One piece of evidence that has been adduced in favor of relative models is a relative insensitivity to the spacing of the stimuli — making stimuli more different makes little or no improvement in accuracy despite gross changes in similarity (Brown, Neath, & Chater, 2002). A traditional exemplar model therefore needs to be augmented by a relative-judgment assumption that sensitivity scales with stimulus range to capture this effect (Brown et al., 2002, but see Lockhead, 2004, for arguments that this does not occur).

The relevant prediction of similarity-based models is that there should be a large advantage for more widely spaced stimuli as the reduction in similarity reduces confusion. The advantage in accuracy is however meagre, as illustrated with the data of Stewart et al. (2005) in Figure 3. In a basic exemplar model, the lowest accuracies should be almost doubled by the doubling in spacing between narrow and wide conditions.

It might nonetheless be the case that there is an advantage for widely spaced stimuli, but that it obtains most strongly in response times, rather than accuracy. We therefore examined previously collected but unanalysed response times from the study of Stewart et al. (2005) for this effect.

Method

Accuracy data from this experiment were previously reported as Experiment 1 in Stewart et al. (2005); here we focus on response times.

Participants

A total of 120 undergraduates from the University of Warwick participated in this experiment.

Stimuli

Two sets of ten tones varying in frequency (pitch) were used. The set for the *narrow* condition had lowest frequency 768.7 Hz, and the ratio between adjacent tones was 1.06:1. The set for the *wide* condition had lowest frequency 600 Hz, and the ratio between tones was 1.12:1. Tones were thus equally spaced in log-Hz, and therefore approximately equally spaced in psychological space. The two sets have equal means in log space.

Tones were 500 ms in duration. Over the first 50 ms, amplitude linearly increased from zero to maximum, and over the final 50 ms, linearly decreased from maximum to zero.

Design

Spacing of tones (narrow and wide) was crossed with set sizes of six, eight and ten stimuli (chosen from the center of the range) to produce six between-participants conditions.

Procedure

Within each of seven blocks that were 120 trials long, each tone was presented equally often, and ordering was randomized was within blocks.

On each trial, a tone was presented over headphones simultaneously with a visually presented “?” that remained on the computer screen until the participant responded with a key from those labeled “1” to “10”. The ordering of the correct assignment was counterbalanced over participants. Although a computer keyboard is not ideal for the comparison of different responses, the primary comparisons here are of responses with the same key, and between stimuli, so this should not affect the pattern of results. The response triggered the replacement of the “?” with the correct response for 750 ms, before a blank inter-stimulus interval of 500 ms.

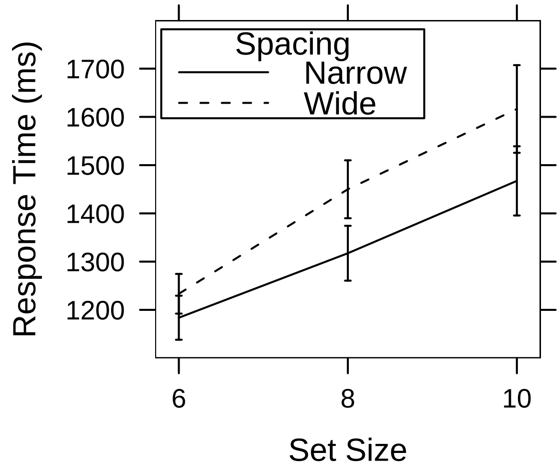

Figure 1: Mean response time as a function of set size and spacing. Error bars indicate $\pm 1SE$.

Results

Slow responses (greater than ten seconds) and fast reponses (less than 150 ms) were excluded from all analyses. Incorrect responses were not excluded, although doing so gives a similar pattern of results.

The relationship between overall mean response time, set size and spacing is illustrated in Figure 1. An ANOVA confirmed that there were significant effects of set size ($F(2, 114) = 13.89$, $p < .0001$, $\eta^2 = .137$) and spacing ($F(1, 114) = 4.57$, $p < .05$, $\eta^2 = .027$), but no interaction ($F(2, 114) = 0.28$, $p > .7$, $\eta^2 = .003$).

These results are broken down by stimulus in Figure 2, with the corresponding accuracy statistics in Figure 3 (cf. Figure 11 of Stewart et al., 2005). Whilst the cost in response is apparent in this break-down, the edge (lowest and highest) stimuli within each set size appear to be immune to the effect. An ANOVA with a further factor of the two edge stimuli vs. all other stimuli continued to demonstrate a significant set size effect ($F(2, 114) = 11.86$, $p < .0001$, $\eta^2 = .136$) and a spacing effect in the margin ($F(1, 114) = 2.91$, $p = .09$, $\eta^2 = .017$), and critically showed a significant interaction between spacing and stimulus endness ($F(1, 114) = 15.71$, $p < .001$, $\eta^2 = .008$). There was also evidence for a main effect of edgeness of stimuli, an interaction between stimulus edgeness and set size, and the three-way interaction. ANOVAs conducted separately for end and non-end stimuli confirmed that non-end stimuli showed an effect of spacing ($F(1, 114) = 6.00$, $p < .05$, $\eta^2 = .042$), but there was no evidence of this for end stimuli ($F(1, 144) = 0.38$, $p > .5$, $\eta^2 = .003$).

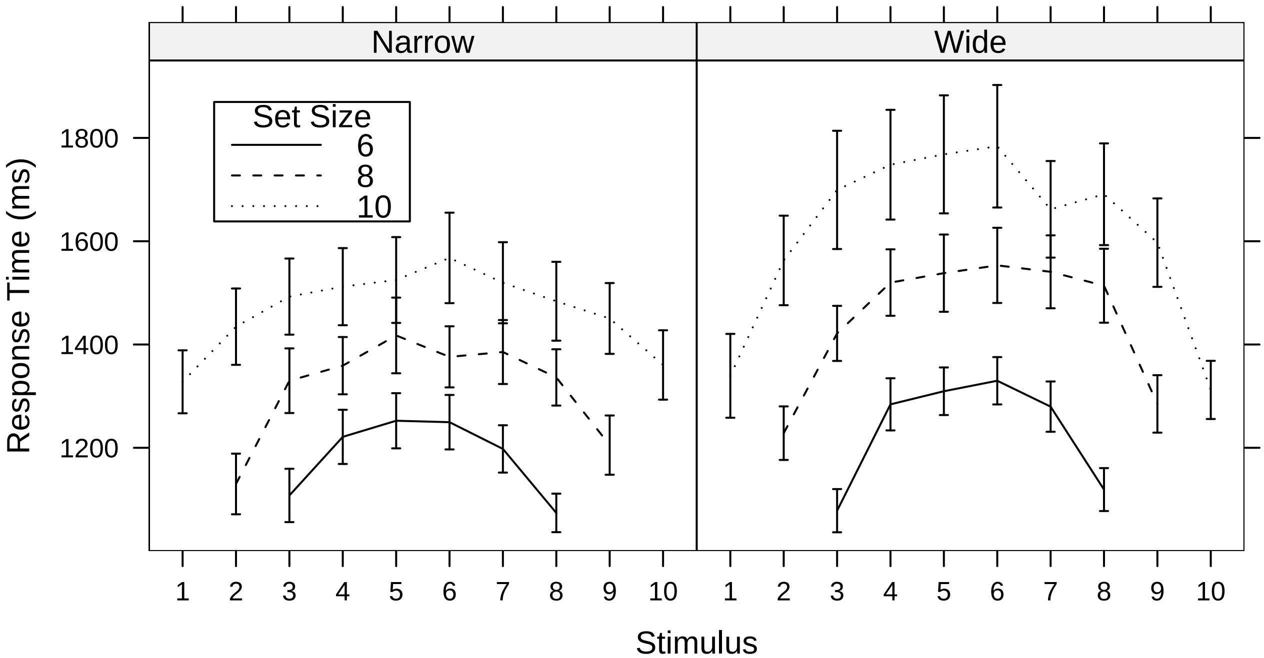

Figure 2: Mean response time as a function of set size, spacing and stimulus. Error bars indicate $\pm 1SE$.

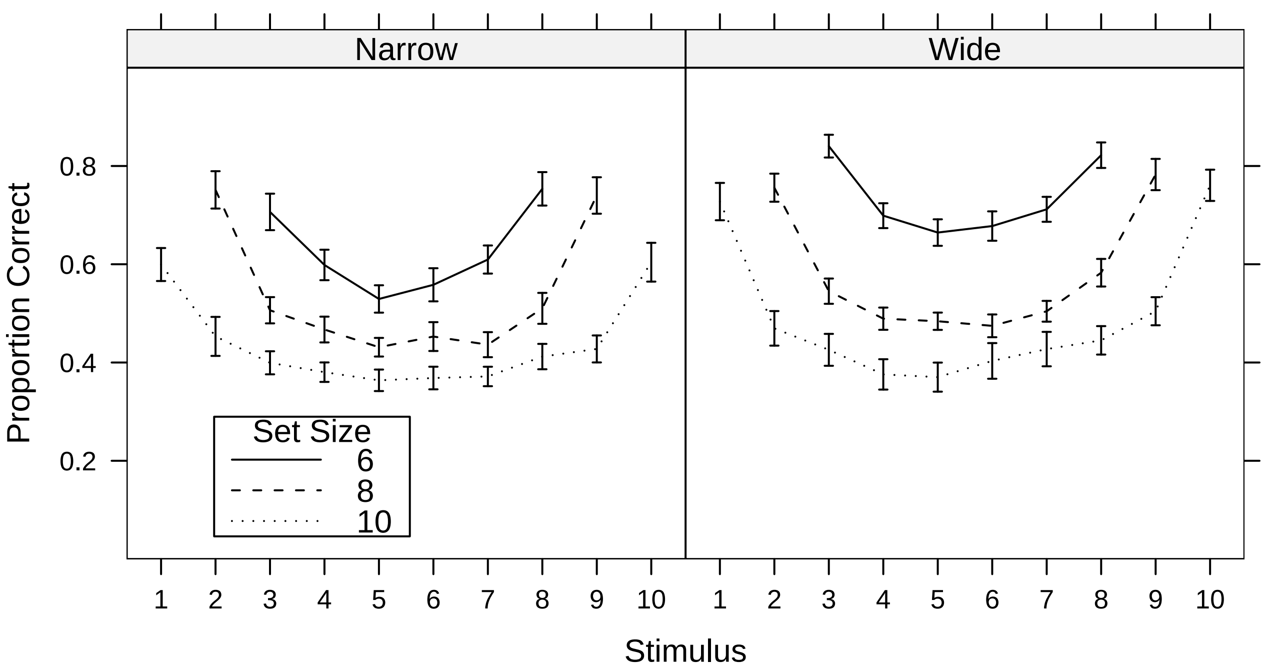

Figure 3: Mean proportion correct responses as a function of set size, spacing and stimulus. Error bars indicate $\pm 1SE$.

Modeling

We optimised parameters for the EGCM-RT for absolute identification (Kent & Lamberts, 2005) by the minimization of the sum squared deviation from the points in Figures 2 and 3, with the deviation measured in standard errors of the mean to approximately equalize the influence of response time and accuracy. Only the sensitivity parameter (c) was permitted to vary between spacings, to allow for adjustment of scale. The residual time parameter (t_{res}) was constrained to be at least 150 ms as this is intended to include the motor component of response (Kent and Lamberts use 250 ms), and other parameters were constrained to be non-negative.

The best fits obtained from this procedure are illustrated in Figures 4 and 5 (parameters were $q = 0.0005$, $c_{\text{narrow}} = 0.55$, $c_{\text{wide}} = 0.35$, $\theta = 1.73$, $t_{res} = 150$, $\alpha = 0.64$, and $\lambda = 0.00$). The ratio between c for narrow and wide conditions was close to 2, the value suggested by Brown et al.'s (2002) relative-judgment assumption (although this cannot be range-based for the EGCM-RT as this would be inconsistent with its explanation of the set size effect) but the wide spacing condition was nevertheless predicted to be more accurate and quicker than the narrow condition. To examine whether this was due to under-weighting of the response time component, we conducted new optimizations weighting the response time squared discrepancies by a factor of 10. This simply reversed the predictions: Wide spacing was predicted to be slightly less accurate and slower than the narrow condition. Permitting t_{res} to vary between conditions (reflecting, for example, differing rescaling costs) resulted in essentially no improvement in fit, because this parameter was already straining at the bound.

Discussion

The new analysis showed an effect of spacing on response times such that responses were slower for widely spaced stimuli despite the responses to these stimuli being more accurate. To the best of our knowledge, this is the first examination of the effect on response times of the manipulation of the spacing of the stimulus set independently of set size and practice[1]. The effect on response time did not appear to be a uniform cost of, for instance, rescaling, because edge stimuli were immune to the effect.

The EGCM-RT could not account for these results when only sensitivity (c) and residual time (t_{res}) were permitted to differ between spacings. The reason for this problem is that both stopping probability (and hence response time) and accuracy are linked to the equivocality of the current similarities ratings: Stopping is more likely to occur (sooner) and responses are more likely to be correct when the similarities strongly favor the single correct response.

The remaining parameters are supposed either to relate to basic perceptual rates or similarity calculations (q, α or λ), or to the criterion for responding (θ). Although a different stopping criterion for each spacing would almost certainly give the correct form of results, Kent and Lamberts (2005) have correctly avoided varying parameters freely between conditions because doing so begs the question: Why do participants behave differently in different conditions?

The effect of stimulus spacing on response time would also seem to be inconsistent with Stewart et al.'s (2005) RJM. In its basic form, without allowance for stimulus noise, the model treats different stimulus spacings identically. Any extended version of this model intended to account for response times must include a mechanism that causes response times for central stimuli to be sensitive to stimulus spacing.

Even if accuracy differences are due to greater perceptual noise in the narrow condition, this does not explain why this condition is faster. If accuracy differences are in fact due to the reduction in time spent on the narrow condition, this does not explain why participants systematically choose to spend less time on this condition; otherwise said, what rule relates condition to stopping criterion?

Suppose that, for instance, judgments in the narrow are more difficult and hence would be slower and less accurate. Then, participants might attempt to compensate by performing faster in the narrow condition but overcompensate to be faster than in the wide condition. With more levels of spacing, it may be possible to test a specific model of such a threshold (over-)adjustment. However, if the adjustment in threshold is so severe, why is accuracy relatively preserved? Alternatively, participants may (possibly incorrectly) think the narrow condition too difficult and expend less effort for a gain in speed with only a small accuracy penalty. If this is the case, a full account will explain why differences in set sizes do not follow the pattern, since larger set sizes must be slower and subjectively more difficult, and yet do not yield faster responses.

Reversing the trade-off to be a slowing to compensate for inadequate accuracy would have to mean that the wide condition is intrinsically more difficult, and accuracy is overcompensated. This might be the case within the Luce, Green, and Weber (1976) attention-band model of absolute identification, as the attention band will cover a smaller region and so is less likely to cover the current stimulus. This model does not, however, extend to response times nor account for biases due to sequential effects. Another possibility is that the wider condition is more difficult because it is further removed from the natural psychological number scale for tone frequency (see Gescheider, 1988, for a discussion of evidence that such a scale exists).

In any case, none of these possibilities readily explains the interaction between the bow and spacing effects: Edge stimuli are no slower in a wide spacing condition, but they are more accurate, whereas center stimuli are both slower and more accurate in the wide spacing condition. It is unclear why this should be the case. Overall, the relationship between response time and spacing

[1] Lacouture (1997) compared these conditions, but participants always participated in the narrow condition before the wide, and so the reverse effect in his data may be due to practice.

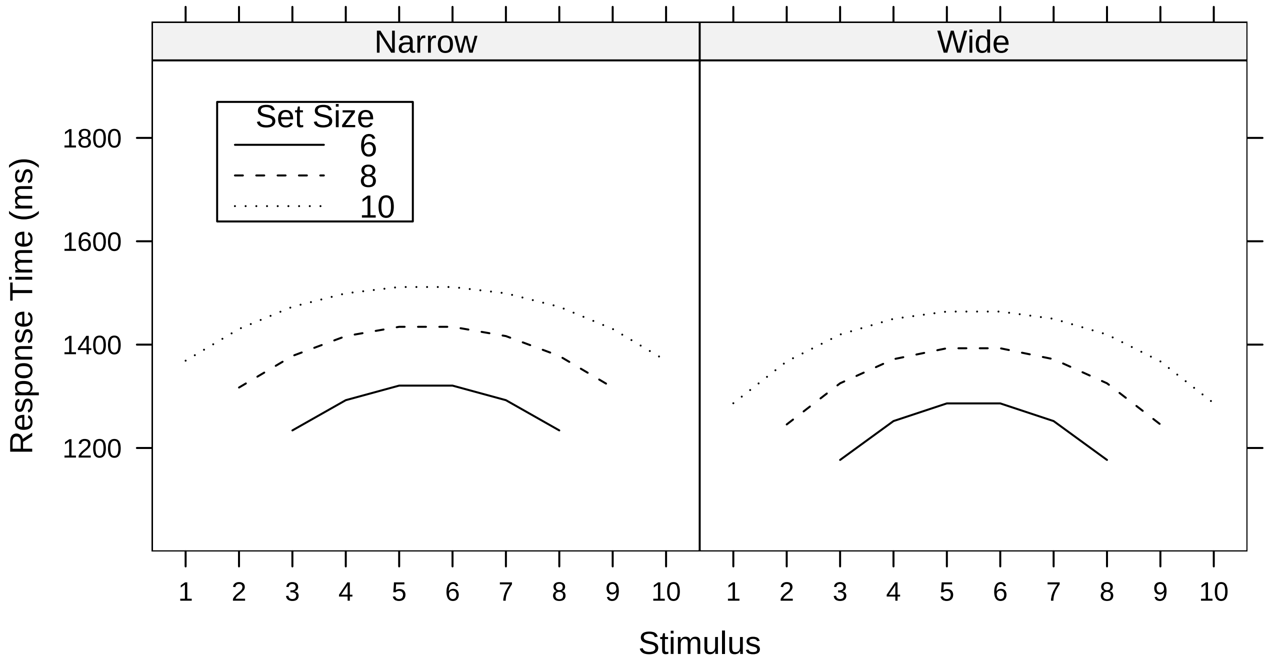

Figure 4: EGCM-RT predicted mean response time as a function of set size, spacing and stimulus.

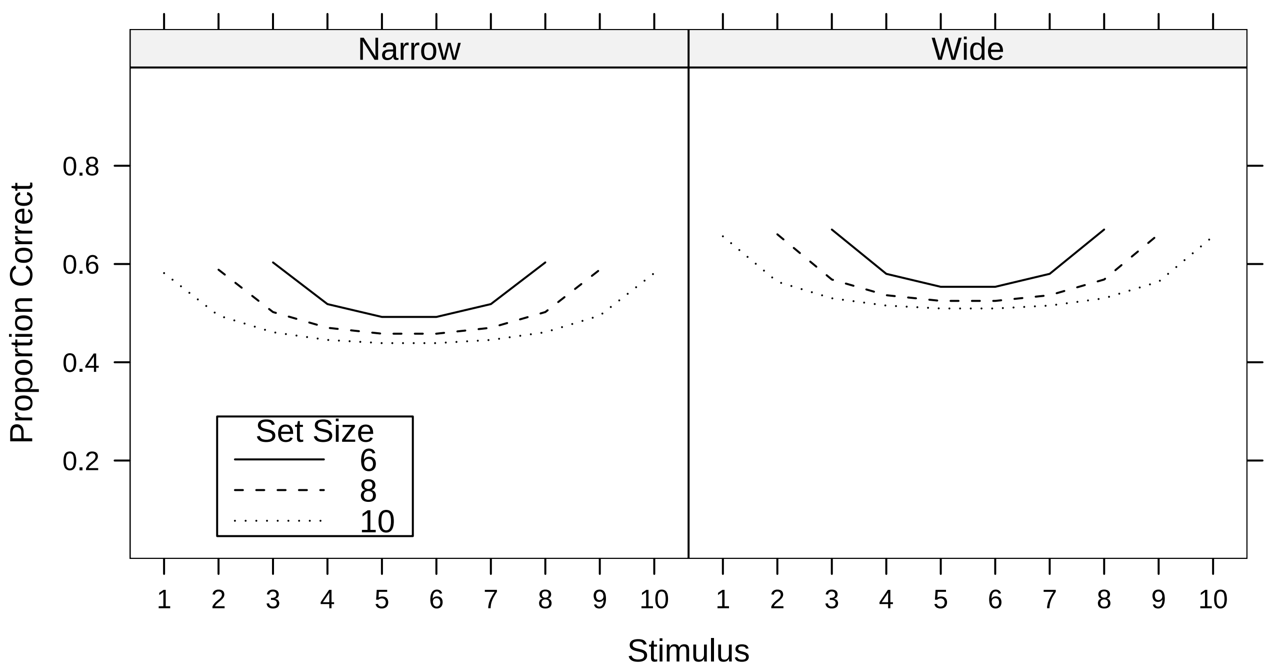

Figure 5: EGCM-RT predicted mean proportion correct responses as a function of set size, spacing and stimulus.

identified here poses a difficult problem for resolution in theoretical accounts of absolute identification.

Acknowledgements

This work was supported by grants RES 000221558 and RES 000231372 from the Economic and Social Research Council (U.K.). We thank Duncan Guest, Elisabeth Blagrove, Chris Kent and Gordon Brown for discussion of the data and manuscript.

References

Brown, G. D. A., Neath, I., & Chater, N. (2002). *A ratio model of scale-invariant memory and identification.* Manuscript submitted for publication.

Gescheider, G. A. (1988). Psychophysical scaling. *Annual Review of Psychology*, *39*, 169–200.

Karpiuk, Jr., P., 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.

Kent, C., & Lamberts, K. (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*, 121–123.

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.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. 1 (pp. 103–189). London: Wiley.

Luce, R. D., Green, D. M., & Weber, D. L. (1976). Attention bands for absolute identification. *Perception and Psychophysics*, *20*, 49–54.

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, NJ: Erlbaum.

Nosofsky, R. M., & Palmieri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika*, *22*, 325–345.

Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review*, *112*, 881–911.