

Sharpening the echo: An iterative-resonance model for short-term recognition memory

D. J. K. Mewhort and E. E. Johns

Queen's University at Kingston, Ontario, Canada

We present an iterative-resonance model for recognition memory. On successive iterations, the probe is compared against a feature-by-feature profile of the study set. *Yes* decisions depend on the similarity of the probe to the profile; *No* decisions depend on a count of elements in the probe that are not in the profile. Successive iterations sharpen the evidence, and response latency is a function of the number of iterations needed to obtain a sufficiently clear result. The model successfully simulates classic data as well as recent data problematic for alternate models.

Many popular accounts of recognition memory claim that subjects compare the recognition probe with each of the study items and calculate an overall familiarity value (see Clark & Gronlund, 1996). Although the models differ both in the details of the calculation and in the way that they describe the familiarity value, they agree that decisions are based on a single scalar that is evaluated using signal-detection theory, a decision mechanism borrowed from perceptual psychophysics (see Green & Swets, 1966). Using signal-detection theory, the familiarity value is compared against a criterion; if it exceeds the criterion, the subject will endorse the item; if it falls below the criterion, the subject will reject it.

The decision model implies that subjects seek evidence that the item was in the list, treating *No* as a default response to be used when familiarity is insufficient to support a *Yes* response. Hence, the same factors must affect *Yes* and *No* decisions: Whatever promotes a *Yes* response should hinder a *No* response, and vice versa.

To test the predictions, we manipulated the featural overlap among stimulus items (Johns & Mewhort, 2002; Mewhort & Johns, 2000). Because

we wished to study retrieval processes, we used a fully encoded study set and focused on response latency as the measure of interest (see Sternberg, 1969). In one study (Mewhort & Johns, 2000), study lists comprised four items defined by their values on two dimensions: Aa , Ab , Bc , Cc , where A , B , and C represent the values on one dimension and a , b , and c represent the values on the other. We compared performance on four types of negative probe, defined by the number of times that each feature of the probe occurred in the study set. A 0:0 probe (e.g., Xx) involved unstudied features on both dimensions. A 1:0 probe (e.g., Bx , Cx , Xa , and Xb) included one feature that occurred in the study set and one unstudied feature. A 2:0 probe (e.g., Ax and Xc) included a feature that occurred twice in the study set plus an unstudied feature. A 1:1 probe (e.g., Ba , Ca , Bb , and Cb) included two studied features.

We obtained predictions from two computational instantiations of familiarity theory: Minerva 2 (Hintzman, 1984, 1988) and REM (Shiffrin & Steyvers, 1997). Both models represent stimuli using vectors of elements: Minerva 2 defines elements by sampling at random from a binomial

Correspondence should be addressed to D. J. K. Mewhort, Department of Psychology, Queen's University, Kingston, Ontario, Canada K7L 3N6. Email: mewhortd@post.queensu.ca

The research was supported by grants from the Natural Sciences and Engineering Research Council of Canada and by an AEG grant from SUN Microsystems of Canada. We also acknowledge assistance of the High Performance Computing Virtual Laboratory (HPCVL) for HPC infrastructure and support. We are indebted to Andrew Heathcote for suggesting the smoothing algorithm.

($-1, +1$) distribution, whereas REM samples at random from a geometric distribution. To represent the stimuli used by Mewhort and Johns (2000), we generated eight vectors of random elements to represent the features A, B, C, X, a, b, c and x and vectors to represent the conjunction of A with a , A with b , A with x , and so forth. Each vector had 40 values drawn randomly from a binomial or geometric distribution, as appropriate. We then concatenated the appropriate feature and conjunction vectors to construct stimuli arranged for an Aa, Ab, Bc, Cc study set and the associated probes.

Table 1 summarises predictions for accuracy averaged over 500 independent trials.¹ As shown in Table 1, Minerva 2 and REM predict the same pattern of results for the stimuli used by Mewhort and Johns (2000): With two-dimensional stimuli, the difficulty of rejecting a foil increases with the number of study items that it overlaps. Hence, a 1:0 foil should be harder than a 0:0 foil, a 2:0 foil should be harder than a 1:0 foil, and a 1:1 foil should be equivalent to a 2:0 foil. Although the two models define familiarity in different ways, both predict the same pattern of results because both sum the calculated familiarity across all studied items.

Table 1 also shows the empirical situation (Mewhort & Johns, 2000, Exps. 1–3). In contrast to predictions from Minerva 2 and REM, accuracy was equivalent for 2:0 and 1:0 probes, and it was dramatically higher than for 1:1 probes. As shown in Figure 1 (top panel), the corresponding response-latency data followed the same pattern.

Mewhort and Johns (2000) argued that the relative difficulty of the 0:0, 1:0, 2:0, and 1:1 probes reflects the detectability of contradictory information. The hardest probe (the 1:1 probe)

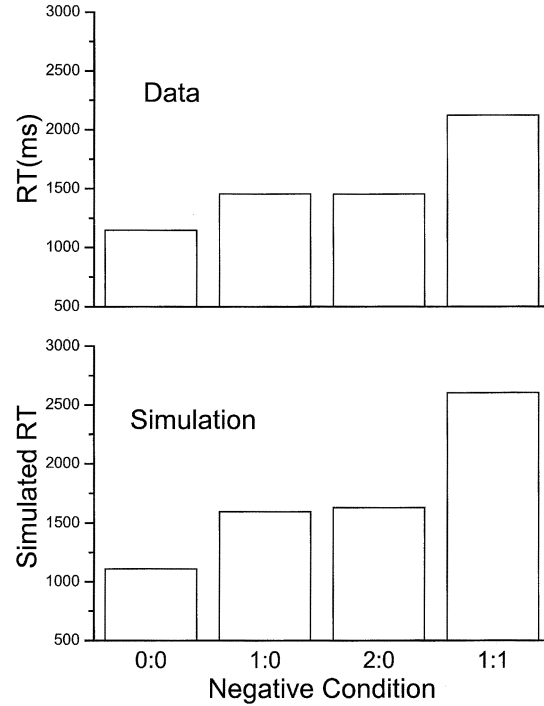


Figure 1. Response latency for correct sure responses as a function of the type of negative probe. The data (top panel) summarise the first three experiments in Mewhort and Johns (2000). The simulation (bottom panel) shows the corresponding data from the IRM averaged over 10000 independent simulations. For the simulations, encoding was perfect, the *Yes* criterion was set at 0.98, the *No* criterion at 5, the running-average parameter was 0.85, a constant of 300 was added to each RT, and the scaling factor was 750.

contains only one source of contradictory information—an unstudied conjunction of studied features. The next most difficult probes (the 1:0 and 2:0 probes) both contain two sources of contradictory information—an unstudied feature and an unstudied conjunction. The easiest probe (the 0:0 probe) contains three sources of contradictory information: two unstudied features and an unstudied conjunction. The more sources of contradictory information, the easier it is to reject the probe.

Subsequent work has confirmed the importance of detecting contradictory information. Johns and Mewhort (2002) contrasted performance on 1:0 and 2:0 probes with two study organisations. For the control condition, each list was organised as before. There were two pairs of items, with one pair sharing a value on one dimension and the other pair sharing a value on the other dimension, i.e., Aa, Ab, Bc, Cc . As before, performance on 1:0 and 2:0 probes was equivalent. For the experimental condition, each

TABLE 1

Probability of correct rejection for four types of foils

Source	Type of negative probe			
	0:0	1:0	2:0	1:1
Minerva	0.99	0.90	0.48	0.48
REM	0.98	0.80	0.63	0.63
Mewhort & Johns (2000)	0.98	0.92	0.92	0.81

¹REM has three main parameters; we used values from simulations described by Shiffrin and Steyvers (1997). Specifically, we used 10 rehearsals; we set $u^* = 0.04$, and we set $c = 0.7$.

list also comprised two pairs of items. The first pair shared a value on one dimension, and the second pair shared a different value on the same dimension, i.e., Aa, Ab, Bc, Bd . For the experimental condition, performance on 2:0 probes was slower than on 1:0 probes (see Figure 2, top panel). The experimental condition exhibited the advantage for 1:0 over 2:0 probes predicted by REM and Minerva, but, because the advantage occurred only in the experimental condition, familiarity theory cannot explain the finding.

Theory based on use of contradictory information anticipates the results. In the control condition, there are always three studied values on the dimension of an extra-list feature; that is, A, B , and C contrast with X , while a, b , and c contrast with x . In the experimental condition, the number of studied values on the dimension of the extra-list feature is confounded with the probe condition. The extra-list feature for the four possible 1:0

probes (Xa, Xb, Xc, Xd) is on the dimension that takes two different values in the study set, A and B . The extra-list feature for the two possible 2:0 probes (Ax, Bx) is on the dimension that takes four different values in the study set, a, b, c , and d . Johns and Mewhort (2002) argued that performance on foils reflects the number of studied values on the extra-list feature; it is more difficult to detect the contradictory information as the number of studied values on that dimension increases.

The experiments summarised in Table 1 and in Figures 1 and 2 are incompatible with familiarity theory. In the next section of the paper, we will sketch a new computational model for recognition and apply it both to the cases that challenge familiarity theory and to some classic phenomena.

THE ITERATIVE-RESONANCE MODEL FOR RECOGNITION MEMORY (IRM)

Representation assumptions

An item comprises several features; the features are both simple features (e.g., colour, shape, initial bigram, final bigram, etc.) and higher-order features (e.g., the combination of a particular colour with a particular shape). Each feature is represented by a vector of binary $[0, 1]$ elements, and an item vector is constructed by concatenating its constituent feature vectors. For example, a coloured shape is defined:

$$item = c_1, c_2 \dots c_i // s_1, s_2, \dots s_j // C_1, C_2, \dots C_k,$$

where the c , s , and C elements define the colour, shape, and conjunction vectors respectively, and $//$ is a standard concatenation operator. For programming convenience, we set $i = j = k = 50$, but there is no principled reason to require that identity.

Current computational models of memory use scalars to represent features instead of vectors. Vectors allow different features from the same dimension to be related. For example, colour is a dimension with many possible values. A vector that represents the feature *purple* may share some elements with vectors that represent the features *blue* and *red*, because *purple* is a mixture of *blue* and *red*. For the simulations reported here, we did not attempt to represent the similarity within features. The vectors used to represent the features were obtained by taking independent sam-

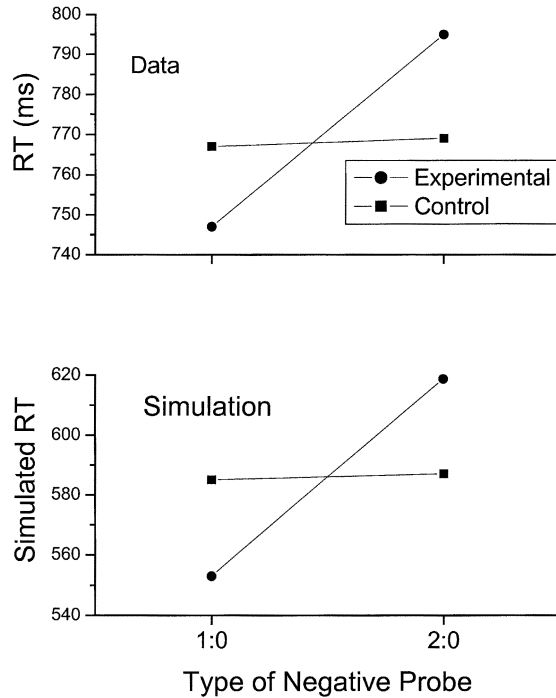


Figure 2. The top panel shows response latency for 1:0 and 2:0 probes under two structures of study list: Aa, Ab, Bc, Cc (Control) and Aa, Ab, Bc, Bd (Experimental). The data were taken from Johns and Mewhort (2002, Experiment 4). The bottom panel shows the corresponding simulated response latency (in arbitrary units) from the IRM averaged over 500 independent simulations. Encoding was perfect, the *Yes* criterion was set at 0.98, the *No* criterion was 5, the running-average parameter was 0.85, a constant of 300 was added to each RT, and the scaling factor was 250.

ples (with replacement) from a binomial distribution with $p = .5$.

Encoding

Following the example of Minerva 2 (Hintzman, 1984, 1988), the model encodes an item in memory by copying the vector to a matrix, one row per item. When the study conditions introduce interference (e.g., with rapid presentation), encoding will be imperfect. We modelled imperfect encoding by weighting the features of the item.

Retrieval and decision

We describe retrieval as a type of resonance (see Ratcliff, 1978). The probe is compared against an element-by-element profile of the set of encoded items. Decision is based on the contents of the profile, and separate evidence is computed for *Yes* and *No* responses.

The model calculates the similarity of the probe to the profile. If the similarity exceeds a criterion, a *Yes* response is given. Elements in the probe not present in the profile provide evidence of contradiction, that is, evidence for an extra-list feature. The model tallies the number of such elements. If the tally exceeds a second criterion, a *No* decision is given.

If the evidence for both responses is insufficient, the subject must try again by reprocessing to obtain better information. Reprocessing is not a matter of re-sampling to average out noise (as in random-walk and diffusion models, see Ratcliff, 2002; Ratcliff, Van Zandt, & McKoon, 1999). Instead, on successive reprocessing steps, the elements of each studied item are weighted by an increasing index of the similarity of that item to the probe. The profile of the study items becomes an echo that reflects the extent to which the study items resonate with the probe. With successive reprocessing steps, the item most similar to the probe is represented more strongly in the echo, and the contribution of the other encoded items is reduced. That is, successive iterations force the profile to hone-in on the studied item most similar to the probe.

The consequence of honing-in on one of the studied items depends on the true status of the probe. If the probe is a positive item, the profile should become closer to that item, and the calculated similarity should increase. If the probe is a foil, the profile should become closer to one of the

studied items, making it easier to find the extra-list elements that distinguish the foil from the nearest studied item.

Response time depends on the number of iterations required to obtain clear evidence for either a *Yes* or *No* response. In our simulations, the count of retrievals rarely exceeds 20. Hence, to estimate response time, we pass the count of retrievals through a smoothing algorithm. For each count, the smoothing function takes an independent sample from an exponential distribution; the exponential samples are summed to obtain a smoothed RT estimate. The mean of the exponential distribution is a scaling parameter.

Table 2 presents pseudo-code describing the algorithm. In the table, *echo* is a vector computed by summing elements across all items (the *number_studied* loop). Before summing into the echo, the elements are weighted by the similarity of the item to the probe (the similarity for each item is stored in the *sim* vector). Each similarity value, in turn, is raised to a power defined by the number of retrieval attempts (*n_cycle*); the series starts with an exponent of 0 and increases across successive retrievals. Increasing the exponent changes the relative contribution of each studied item to the composition of the echo. Evidence supporting a *No* decision (*n_elf*) is obtained by tallying elements of the probe not present in the echo, and evidence for a *Yes* decision (*similarity*) is obtained by computing the normalised inner-product (vector-cosine) of the probe and the current echo. The calculations are embedded in a *repeat-until* loop that iterates until the evidence reaches criterion for either a *Yes* or a *No* decision.

Applying the iterative-resonance model

In Table 1, we showed that existing theory does not capture the effects on correct rejections of similarity to the studied items reported by Mewhort and Johns (2000). In Figure 1 (top panel), we presented the corresponding response times from their study. As in the accuracy data (Table 1), subjects found a 1:0 probe as difficult as a 2:0 probe, and a 2:0 probe easier than a 1:1 probe. Simulations presented in Figure 1 (bottom panel) show that the IRM matches the data from the subjects. In the IRM, performance on the 1:0 and 2:0 probes is equivalent because they contain equivalent contradictory evidence: an extra-list feature and an extra-list conjunction. The 1:0 and

TABLE 2
Pseudo-code for the retrieval and sharpening algorithm used by the IRM

```

n_cycle = 0; echo = 0
Repeat
  do j = 1 to number_studied                ! Retrieve echo
    do i = 1 to number_elements
      echo(i) = echo(i) + item(i, j)*(sim(j)**n_cycle) ! Echo = f(n-cycle)
    enddo ! i-loop
  enddo ! j-loop
  n_elfs = 0
  do i = 1 to number_elements                ! Look for extra-list elements
    if (echo(i) - probe(i) < 0) n_elfs = n_elfs + 1
  enddo ! i-loop
  similarity = Vector_Cosine(probe, echo)    ! Compute familiarity
  n_cycle = n_cycle + 1
  echo = echo * running_average
Until ( (similarity > YES_crit) .or. (n_elfs > NO_crit) )

```

2:0 probes are easier than a 1:1 probe because they contain both an extra-list feature and an extra-list conjunction, whereas a 1:1 probe contains only an extra-list conjunction.

The IRM also captures the effect of varying the number of studied alternatives to the extra-list feature (Figure 2, bottom panel). As the number of studied alternatives increases, the probability of coincidentally encoding elements of the extra-list feature increases. When some of its elements appear in the profile of the study set, it is more difficult to detect the extra-list feature in the probe. Hence, the more studied alternatives to an extra-list feature, the more difficult it is to reject the item.

The experiments summarised in Figures 1 and 2 are based on Sternberg's (1969) classic studies of short-term recognition memory. Like his experiments, they involved fully encoded study sets and they used response latency as the main dependent variable. Because the work so far uses the Sternberg paradigm, it is necessary to show that the model captures standard results from the paradigm.

Sternberg (1969) varied study-set size and found that decision time increased linearly with set size for both positive and negative responses; the serial-position functions were flat. To simulate his experiment, we created vectors by sampling with replacement from a binomial distribution with $p = .5$. Because Sternberg scored only trials on which the subject could recall the study set, we assumed perfect encoding. The top left panel of Figure 3 shows the IRM's simulated set-size functions; as in Sternberg's studies, RT increased at a rate of about 35 ms per item. The top right

panel of Figure 3 shows the IRM's simulated serial-position functions; as in Sternberg's data, the serial-position functions are flat.

In the IRM, the set-size functions reflect the information added to the echo by additional study items. As set size increases, the echo includes data from additional items, and it takes more cycles to hone-in on one item sufficiently to support a response. The serial-position curves are flat because all items were fully encoded.

When the study time and retention interval are reduced relative to the times used by Sternberg (1969), the serial-position curves are not flat but are bow-shaped with a marked advantage for the most recent item (e.g., Corballis, Kirby, & Miller, 1972). Moreover, when expressed as function of lag from the probe, the curves for different set sizes overlap for all but the initial items (McElree & Doshier, 1989; Monsell, 1978).

Reducing the study time should reduce the accuracy of encoding but not by an equal amount at all serial positions. To simulate proactive inhibition, we weighted the elements of all but the first item in the list by a constant (0.81). To simulate retroactive inhibition, we weighted the elements of the composite by a second constant (0.85) as each item was added to the composite.

Figure 3 also shows the predictions from the IRM for the varied-set procedure with reduced study time. We created vectors to represent items as before. The bottom left panel in Figure 3 presents the data as a function of the set size. As before, the model predicts increasing set-size function for both positive and negative items. The bottom right panel in Figure 3 shows response latency as a function of serial position. The serial-

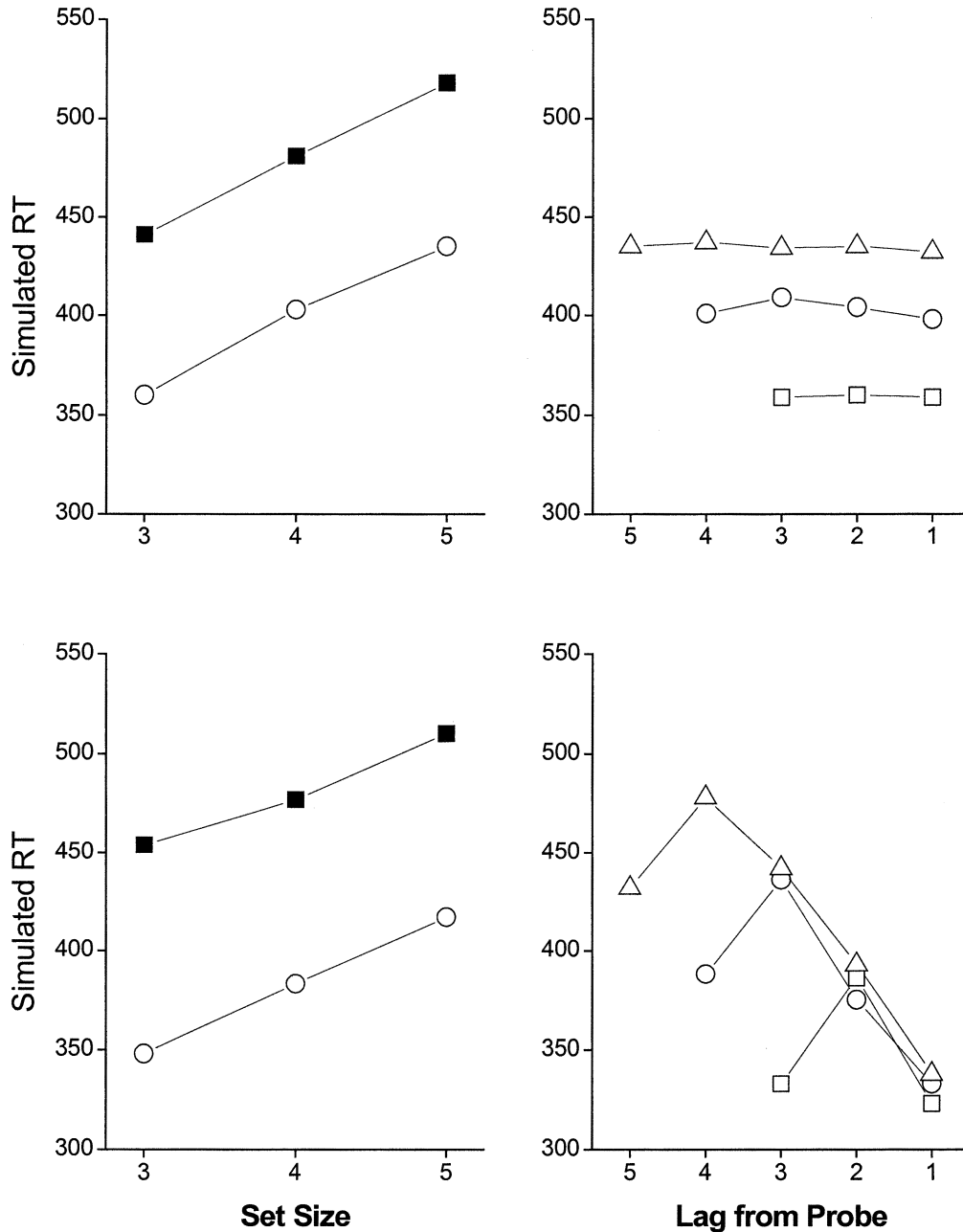


Figure 3. Simulation of a varied-set Sternberg task with fully encoded lists (top panels) and imperfectly encoded lists (bottom panels). The left panels show predicted response latency (in arbitrary units) for positive responses (open symbols) and negative responses (closed symbols) as a function of set size. The right panels show the positive responses as a function of the serial position in the study list. In the simulation of the fully encoded lists, the *Yes* criterion was set at 0.85, the *No* criterion at 22, the running average parameter was 0.85, a constant of 300 was added to each RT, and the scaling factor was 75. For the simulation of the imperfectly encoded lists, the corresponding parameters were 0.81, 21, 0.85, 300, and 75.

position curves are approximately the same shape as the curves reported by Monsell (1978) and by McElree and Doshier (1989).

Because experiments in the Sternberg (1969) tradition are largely sub-span tasks, the work is usually thought to be limited to short-term

memory. There is no principled reason, however, to limit application of the IRM to short-term memory experiments. Burrows and Okada (1975) extended the Sternberg (1969) paradigm to lists of up to 20 items by requiring subjects to pre-memorise the lists. They obtained parallel set-size

functions for positive and negative items and fitted the combined curve successfully (a) to bi-linear functions (separated at span) and (b) to a single increasing log function.

Using the parameters from our simulation of the Sternberg (1969) experiment, we extended the simulation to set size up to 21 items. Following Burrows and Okada (1975), we combined the positive and negative set-size functions. Figure 4 (top panel) shows predicted response latency as a function of set size with a bi-linear fit superimposed on the data. The bottom panel shows the data as a single function expressed on a \log_2 scale. As is clear in the figure, the IRM reproduces Burrows and Okada's results; in particular, a single log function provided an excellent account of the data.

Final comments

We have sketched a new model for recognition memory, the IRM. In the model, retrieval is

treated as a resonance process. An echo is produced when a probe is applied to memory. The echo represents a profile of the features of the studied items. Evidence for *Yes* decisions is based on the similarity of the probe to the echo; evidence for *No* decisions is based on contradictory elements in the probe. Successive iterations sharpen the evidence, and response latency is a function of the number of iterations needed to obtain a clear result.

We have shown that the IRM captures classic data along with recent data that are problematic for familiarity-based accounts. We have not fitted the model to experimental data in the usual way. The conventional strategy is to minimise the root-mean square deviation (RMS) of the predicted values from the obtained data. For a multi-parameter model, minimising the RMS usually requires an iterative algorithm to search the parameter space. The fits shown here were obtained using a less formal procedure. We adjusted the parameters so that the model predicted a small percentage of errors (approximately

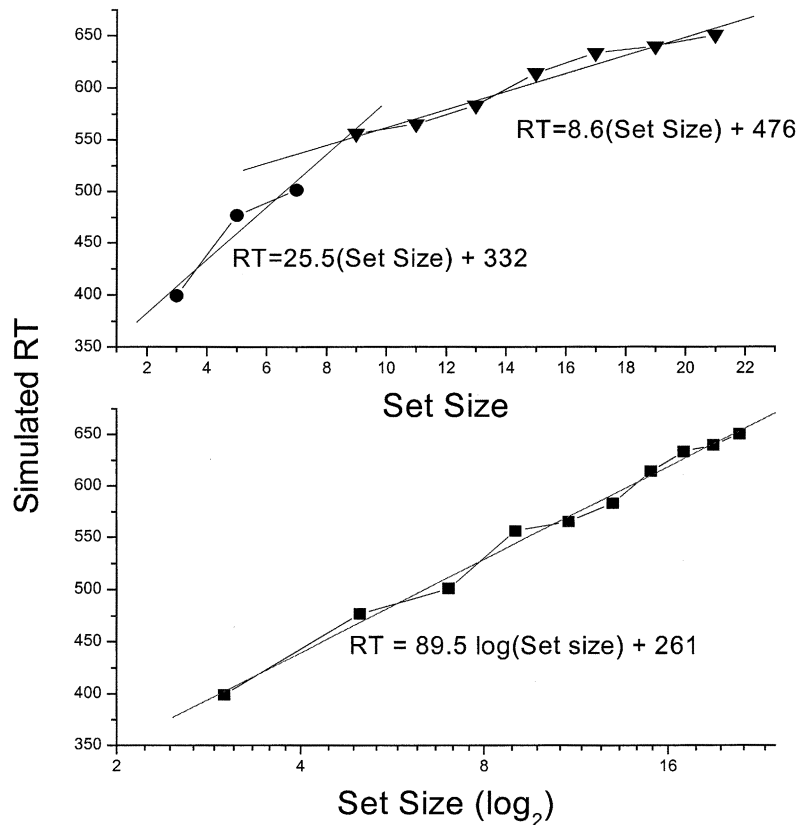


Figure 4. Simulation of the Burrows and Okada (1975) fixed-set paradigm using pre-memorised lists. The top panel shows the simulated RT (in arbitrary units) as a function of the size of the study set fit to bi-linear functions split at memory span. The bottom panel shows the same data fit to a single log function.

2%), about the level of error found in the experiments under consideration. The predicted data fall out without forcing the parameters to yield the desired results when scaled using the smoothing parameter.

The model is a work in progress. In Table 1, for example, the exponent that weights the similarity of the probe with each member of the studied set increases across iterations. That successive iterations depend more heavily on the similarity of the probe to the studied items is a principle of the model, but there is no reason to insist that the change occurs in integer units. Similarly, although it is likely that a smoothing function will be necessary to convert a count of iterations into response latency, the model is not wedded to the sum-of-exponentials function.

Finally, we wish to note some differences between the IRM and alternative models. The item recognition task asks if an item occurred in a specified study set; hence, the IRM constructs a profile that describes the relation of the probe to the study set as a whole. Because the individual items are stored, unlike holographic models such as TODAM (see Murdock, 1982, 1983), we do not preclude different retrieval operations to answer other questions about the encoded items. Second, iterative resonance has a superficial similarity to the popular sequential-sampling models of decision (see Rouder, 2001; Smith, 2000). Both yield an estimate of accuracy and response latency. In the sequential-sampling models, the decision process is noisy, and the accumulation of evidence across iterations is a way of averaging out the noise. In the IRM, by contrast, the evidence is changed across iterations to yield an increasingly clear picture of the relation between the probe and the study set. Variability across simulation runs reflects the stimulus representation; it is not a property of the decision process. Finally, in the IRM, evidence for *Yes* and *No* responses is time-locked to the echo.

REFERENCES

- Burrows, D., & Okada, R. (1975). Memory retrieval from long and short lists. *Science*, *188*, 1031–1033.
- Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review*, *3*, 37–60.
- Corballis, M. C., Kirby, J., & Miller, A. (1972). Access to elements of a memorized list. *Journal of Experimental Psychology*, *94*, 185–190.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*, 96–101.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace model. *Psychological Review*, *95*, 528–551.
- Johns, E. E., & Mewhort, D. J. K. (2002). What information underlies correct rejections in recognition from episodic memory? *Memory & Cognition*, *30*, 46–59.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346–373.
- Mewhort, D. J. K., & Johns, E. E. (2000). The extralist-feature effect: A test of item matching in short-term recognition memory. *Journal of Experimental Psychology: General*, *129*, 262–284.
- Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *4*, 465–501.
- Murdock, B. B. Jr. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*, 609–626.
- Murdock, B. B. Jr. (1983). A distributed memory model for serial-order information. *Psychological Review*, *90*, 316–338.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review*, *9*, 278–291.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connection and diffusion models of reaction time variability. *Psychological Review*, *106*, 261–300.
- Rouder, J. N. (2001). Testing evidence accrual models by manipulating stimulus onset. *Journal of Mathematical Psychology*, *45*, 334–354.
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.
- Smith, P. L. (2000). Stochastic dynamic models of response time and accuracy: A foundational primer. *Journal of Mathematical Psychology*, *44*, 408–463.
- Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, *57*, 421–457.