

# 2

## Modeling Distinctiveness: Implications for General Memory Theory

JAMES S. NAIRNE

The capacity to remember, to use the past in the service of the present, is a highly adaptive component of cognitive functioning. Although one need not reproduce the past, either consciously or unconsciously, in order to benefit from the service of memory, reproduction is clearly an important design feature (Anderson & Schooler, 2000; Nairne, 2005). Telephone numbers, street addresses, medication times, passwords—each needs to be recovered exactly, with the components in sequence, and inferential or reconstructive processing is unlikely to suffice.

To explain the specificity of retention, students of memory appeal often to the concept of distinctiveness, the focus of the present volume. Mnemonic distinctiveness can be defined in various ways—for example, as a property of a stored trace, a retrieval cue, or as a type of processing (see Hunt, Chapter 1 this volume; Schmidt, 1991). I define it here as the extent to which a particular cue (or set of cues) specifies a particular stored event (or target response) to the exclusion of others. Framed in this way, distinctiveness is not a fixed property of a cue, or a target trace, or even of an interaction between a given cue and a given target. It is a property of a cue in context: given a fixed set of alternatives, a measure of distinctiveness can be assigned to a particular cue with respect to a particular alternative. Change the context—for example, by changing how the cue is perceived or the range of possible responses—and the measure of distinctiveness changes as well.

To facilitate our discussion, and to add some formality to the preceding definition, I introduce a simple retrieval model below (borrowed from my feature model of immediate retention; Nairne, 1990a) and show how it helps account for some of the phenomena classically associated with the study of distinctiveness. For example, I show how the model informs us about the particulars of the von Restorff effect (Hunt, 1995;

von Restorff, 1933) and about the paradoxical effects of processing similarity and difference on episodic retrieval (Hunt & McDaniel, 1993). I then consider the role of time in the calculation of distinctiveness and contrast the retrieval model with certain extant models of temporal distinctiveness (e.g., Brown, Neath, & Chater, 2002; Neath, 1993). Finally, I end the chapter by discussing how the retrieval model forces us to reassess some widely held beliefs about memory, particularly the notion that memory is directly related to the match between an encoded cue and an encoded target.

## A SIMPLE MODEL

Directed retrieval reduces ultimately to a matter of response selection. There is a vast storehouse of information in the brain; the retrieval problem is to select appropriate content based on information available in the present. When we forget an item from a memory list, we are not really forgetting the item—we are forgetting that it occurred in a particular space and time defined by the memory list; when we forget where we parked our car, we are not forgetting our car, we are forgetting the position our car occupied today as opposed to yesterday or the day before. Retrieval cues help us solve these kinds of discrimination problems. They provide us with the information we need to pick and choose from the wide variety of responses that are potentially available.

To formalize the response selection process, I adopt a simple retrieval, or choice, rule of the type often found in categorization and some memory models (e.g., Nosofsky, 1986; Nairne, 1990a, 2001). Under this formulation, an item is chosen for recall by comparing, or matching, the operative retrieval cue(s) to possible candidates in long-term memory (see also Raaijmakers & Shiffrin, 1980). The probability that any particular event,  $E_1$ , will be selected as the recall candidate depends on how well the retrieval cue,  $X_1$ , matches  $E_1$  to the exclusion of other possible recall candidates (e.g.,  $E_2, E_3, \dots, E_N$ ):

$$P_r(E_1|X_1) = \frac{s(X_1, E_1)}{\sum s(X_1, E_i)} \quad (1)$$

The quantity  $s(X_1, E_1)$  refers to the similarity of  $X_1$  to  $E_1$ , which in turn varies as a function of the number of matching or mismatching features between the two terms (a distance measure). Shepard (1987) recommends relating distance ( $d$ ) to similarity in the following manner:

$$s(X_1, E_1) = e^{-d(X_1, E_1)} \quad (2)$$

This means that nearby items in psychological space (e.g., those that contain few mismatching features) will be deemed the most similar (and thereby

produce the largest effects), and similarity will fall off rapidly with increasing distance.

Equations 1 and 2 are not meant to suffice as a complete model of memory. Among other things, one needs to specify how event traces are represented in memory, how probabilities translate into actual output (Nairne, 1990a; Raaijmakers & Shiffrin, 1980), and the similarity and distance measures need to be scaled appropriately as well (Nosofsky, 1986; Shepard, 1987). However, as I demonstrate below, this simple ratio model provides a nice conceptual framework for interpreting the empirical patterns of concern in the distinctiveness literature. Note that equation 1, which expresses the probability that a particular target event will be selected, doubles as our measure of distinctiveness. Distinctiveness is therefore a property of a cue, but only with respect to a particular retrieval candidate. By itself, the measure tells us nothing about whether the retrieval candidate is correct or incorrect, or good or bad from a mnemonic standpoint.

If the goal is to recover event  $E_1$  in the presence of a particular cue  $X_1$ , then equation 1 isolates the factors that promote successful sampling. To maximize the probability of selecting  $E_1$ , it needs to be similar to the cue,  $X_1$ , and dissimilar to other possible retrieval candidates ( $E_2, E_3, \dots, E_N$ ). The numerator of equation 1 tells us that retrieval will depend importantly on the match between the retrieval cue and the target (Thomson & Tulving, 1970); the denominator quantifies cue overload, or the extent to which a cue is predictive of many things (Earhard, 1967; Watkins & Watkins, 1975). Successful recovery, put generally, will be proportional to the cue–target match and inversely proportional to the amount of cue overload. Note that because of the ratio form, neither the cue–target match nor the amount of cue overload, alone, will be sufficient to predict successful retention; successful recovery of a target will always depend on both. As I discuss later, this conclusion has a number of implications for general memory theory.

### The von Restorff Effect

To illustrate how the retrieval rule works, I begin by applying it to the von Restorff effect—the so-called mother of all distinctiveness effects (Hunt & Lamb, 2001). The *von Restorff effect* (or isolation effect) refers to the memory enhancement that is found for events that differ, or deviate, from their context. In von Restorff's original experiments, participants recalled 10-item lists containing either 10 unrelated items (list 1), nine numbers and one nonsense syllable (list 2), or nine nonsense syllables and one number (list 3). The discrepant items were remembered best (e.g., the number in the list of syllables)—even better, in fact, than the unrelated items occupying similar list positions (e.g., items from list 1), or the “background” homogenous items (e.g., the syllables in list 3).

In experiments of this type, items become distinctive by virtue of their list context; that is, items are “isolated” only relative to particular back-

grounds. To consider a specific case, if the number 43 was presented in each of the three von Restorff lists, we would expect its retention to be enhanced only in list 3, where it stands out from the other list items. For the effect to emerge, typically, the nonisolated (or background) items need to share some measure of similarity—that is, the detection of “difference” depends on a background of similarity (see Hunt, Chapter 1 this volume; Smith & Hunt, 2000). As I will discuss later, it is possible to reduce or eliminate the isolation advantage simply by asking people to focus their processing on how items differ from one another in a typical von Restorff list (Hunt & Lamb, 2001).

The isolation advantage also remains robust when the isolate occurs early in the list, even in the first serial position (Kelley & Nairne, 2001; Pillsbury & Rausch, 1943). This is an important finding because it suggests that the locus of the effect should be placed at retrieval. Encoding-centered accounts have been proposed over the years, and it seems reasonable to argue that isolates sometimes do capture more attentional resources (Schmidt, 1991), but encoding-centered accounts have difficulty explaining why the effect is found when the isolate occurs in the first or second serial position. At this point, no list context has been established, so there is no background of similarity against which the item can be considered unusual or especially salient. Instead, as embodied in the retrieval model, it makes more sense to assume that the isolate leads to the encoding of features that potentially help one discriminate its prior occurrence at retrieval, after all of the list items have been presented.<sup>1</sup>

To implement the model, it is necessary to make some assumptions about how items are represented in memory, about how similarity is calculated, and about the nature and generation of retrieval cues. Following Nairne (1990a), one can represent items as lists of features and distance derived by comparing features across each position. The number of mismatching features is summed and the total is then divided by the number of compared features. For example, suppose memory trace A is represented by a vector of five features, [C C 2 3 1], and memory trace B by a second vector, [C X 2 2 1]. A feature-by-feature comparison reveals two mismatching features—in positions 2 and 4. Dividing the number of mismatching features (2) by the number of compared features (5) gives us the distance measure (.40). This distance measure is then plugged into equation 2, yielding a similarity value of .67. (For further numerical examples, see Nairne, 1990, 2001.).

In the retrieval model, the critical similarity comparisons are between cues and viable retrieval candidates stored in long-term memory. Like most memory theorists, I assume that the immediate present is used to recover the past—that is, memories do not spontaneously appear, but rather are cue-driven (Tulving, 1983). In the feature model, which deals primarily with remembering over the short term, the operative retrieval cues are lingering records of the immediate past, which can be accessed directly (from primary memory) or recovered through context. When one is remember-

ing over the longer term, a comparable process occurs: some version of the original encoding record is recovered via context and “interpreted” by sampling from a candidate set of possible responses. Equation 2 specifies how the recovery process proceeds: the record is compared to each possible item in the candidate set and, based on the relative similarity values, a candidate is selected for recall (see Nairne, 2001, 2002a).

### A Numerical Example

Table 2.1 shows similarity and sampling probabilities for some hypothetical three-item lists. Encoded list items are represented by trace vectors of five features (e.g., a sequence of letters and digits). The first list, labeled “Control,” is meant to contain three unrelated items, although, importantly, some measure of similarity is assumed (e.g., overlapping contextual features). The cue, shown to the left, is an intact version of the second list item; under normal conditions, this cue would presumably be a blurry or degraded record of the encoding, but it is presented intact here for the sake of simplicity. The last two columns show the similarity and sampling calculations based on the comparisons between this cue and each of the three list traces. Correct recall, given this cue, occurs when the second list item is sampled and successfully recovered (see Nairne, 1990a, for details).

The second list, labeled “Isolate,” instantiates the isolation manipulation: A new nonoverlapping feature, X, is represented in the trace for item 2, replacing one of the shared contextual features (we might assume, for example, that the second list item was presented in a unique color or voice). In all other respects, items remain the same. Note that the similarity value between the cue and its long-term memory representation remains the same, 1.0, but the probability of sampling that target increases. This is the isola-

TABLE 2.1 Similarity Values and Sampling Probabilities Generated by the Retrieval Model for a Hypothetical Three-Item List

	<i>Cue</i>	<i>Traces</i>	<i>Similarity</i>	<i>Samp. Prob.</i>
Control	[C C 2 3 1]	[C C 1 2 3]	.55	.26
		[C C 2 3 1]	1.0	.48
		[C C 3 1 2]	.55	.26
Isolate	[C X 2 3 1]	[C C 1 2 3]	.45	.24
		[C X 2 3 1]	1.0	.53
		[C C 3 1 2]	.45	.24
Iso/Sim	[C C 2 3 1]	[B B B B 3]	.37	.21
		[C C 2 3 1]	1.0	.58
		[B B B B 2]	.37	.21

*Note:* All of the calculations are based on a cue vector representing the second item on the list. Sampling probabilities may not add to one because of rounding.

tion effect, and it is caused here by a reduction in cue overload: the correct target is more likely to be sampled because the cue is now less similar to other candidates. The addition of feature X, which is unique to the encoding of the second list item in this list context, reduces the number of matching features between the cue and the target's competitors.

The third list, labeled "Iso/Sim," shows what happens when the control item from the first list is presented against a background of highly similar items. This is the same target representation and cue as in the first list, and the cue–target match remains perfect, but the probability of correctly sampling the target increases substantially. Once again, what determines performance is the overlap between the features of the target item and those of the background items. As the similarity among the background items increases, their match with the operative retrieval cue decreases. Note, however, that it is not background similarity per se that mediates performance; what matters is the overlap between the cue and the nontarget competitors. If the similarity of the background items increases, but in a way that also increases their match with the retrieval cue, then performance would suffer rather than improve.

Of course, recovery of the isolated item in the model also depends on how well the cue matches the relevant target. The cue–target match is held constant in Table 2.1, but it is easy to imagine isolation improving the functional cue–target match. For example, by definition an isolated item contains features that are unusual in that list context; consequently, those features, once encoded, are probably less susceptible to interference (i.e., overwriting) from subsequently presented items. This should help guarantee an intact cue at retrieval, one that better matches its representation in long-term memory. Moreover, when the isolated item occurs after the list context has been established, its appearance is surprising, which in turn could enrich the overall encoding (or hurt encoding in some circumstances—see Schmidt, Chapter 3 this volume). Richer or more elaborate encodings tend to be matched better by relevant retrieval cues and more protected from interference. In any given situation, it will be difficult to disentangle the relative contributions of the cue–target match and changes in the amount of cue overload; the presence or absence of unusual features is likely to affect both.

## Background Recall

It is also of interest to consider how the presence of an isolated item affects recall of the nonisolated (background) items in the list. In principle, one can conceive of the isolate acting in several ways: enhancing recall of the isolate itself, reducing memory for the background items, or leading to both outcomes. The literature is somewhat equivocal in regard to background recall; sometimes the presence of an isolate hurts the retention of the other list items (e.g., Schmidt, 2002; Schulz, 1971), sometimes recall of those items improves (Farrell & Lewandowsky, 2003), and often there is no effect (e.g., Kelley & Nairne, 2001).

Theoretically, it is easy to justify any of these outcomes. From an organizational perspective, some theorists have argued that the isolate promotes the formation of two list-based categories, one containing the isolated item and a second category comprising the background items (Bruce & Gaines, 1976; Fabiani & Donchin, 1995). Because it is easier to recall items from smaller categories, better memory is expected for both the isolate and the background items. Alternatively, if the isolate captures more attentional resources, or is more likely to be rehearsed, then recall of the background items should suffer because they receive a smaller proportion of the allocated resources. One could argue as well that isolated items, because of their superior mnemonic value, will tend to be recalled early during output, rendering the remaining items subject to more output interference (e.g., Cunningham, Marmie, & Healy, 1998; Schmidt, 1985).

The retrieval model makes no explicit assumptions about encoding, organizational processing, or selective rehearsal; it merely assumes that the recall of an item (isolate or background) will depend on the cue, its match to the relevant target, and the composition of the competitor set. Table 2.2 shows the similarity and sampling values for a background item in our three hypothetical lists. In this case, the cue is for the first list item instead of the isolate, and the correct response is to sample the first of the three list vectors. (Identical values hold for the third item.) Note that the sampling probabilities for this background item change across the three conditions.

Of initial interest is the comparison between lists with and without an isolate. The probability of correctly sampling the first list item in the Control condition is .48 compared to .50 in the Isolate condition. This slight increase, which is predicted by a grouping or organizational account, is caused here by a net decrease in the amount of cue overload (the overall

TABLE 2.2 Similarity Values and Sampling Probabilities for the Background Items in a Hypothetical Three-Item List

	<i>Cue</i>	<i>Traces</i>	<i>Similarity</i>	<i>Samp. Prob.</i>
Control	[C C 1 2 3]	[C C 1 2 3]	1.0	.48
		[C C 2 3 1]	.55	.26
		[C C 3 1 2]	.55	.26
Isolate	[C C 1 2 3]	[C C 1 2 3]	1.0	.50
		[C X 2 3 1]	.45	.23
		[C C 3 1 2]	.55	.28
Iso/Sim	[B B B B 3]	[B B B B 3]	1.0	.46
		[C C 2 3 1]	.37	.17
		[B B B B 2]	.82	.37

*Note:* All of the calculations are based on a cue vector representing the first item on the list. Sampling probabilities may not add to one because of rounding.

value of the denominator in equation 1 goes down). Because there are fewer overlapping features between the isolate and everything else, the isolate is less likely to be sampled when cued by traces left by any of the other list items. Interestingly, the presence of an isolate actually increases the distinctiveness of the remaining items on the list. The effect is small because the decrease in the denominator is caused only by the comparison between the cue and the isolate; with longer lists, this contribution is less important—it is proportionally smaller—which may help explain why a null effect of the isolate on background recall is often reported.

The same reasoning applies to the third condition, Iso/Sim, although background recall is low relative to the other two conditions. Cues for the background items are less distinctive in this condition because the members of the competitor set (with the exception of the isolate) share lots of features. Despite the low sampling probability, however, background performance is not actually hurt by the presence of the isolated item; in fact, for the same reasons discussed in the preceding paragraph, having an isolate in the list improves the sampling probabilities for background items, at least relative to a list containing all similar items. From the model's perspective, any manipulation that decreases the overlap among traces will increase the likelihood of correct cue-target sampling (assuming the cue-target match remains constant). Consequently, whether background items will show a benefit, no effect, or a loss will depend on the control condition and on other factors such as the length of the lists employed.

### Processing Effects

As noted, in its simplest form the retrieval model makes no assumptions about encoding or processing. The isolation effect, as well as background recall, is determined solely by the state of cues, targets, and competitors at the point of recall. However, any encoding manipulation that affects trace composition is likely to influence performance. As discussed earlier, when an isolate occurs late in a list, its surprise value could easily lead to additional processing (or more rehearsal), which in turn could produce a more elaborate memory trace. Moreover, any orienting task that causes participants to focus on common or unique features across to-be-remembered items should have significant effects on performance as well.

In one relevant study, Hunt and Lamb (2001) examined how various orienting tasks affect the isolation advantage. Participants were given lists containing either 10 related items (e.g., a list of vegetables) or 9 related items and 1 item from a different category (e.g., a tool). The standard isolation advantage was produced across the lists—that is, the tool in the list of vegetables was remembered best. Of main interest, however, were several orienting tasks that induced participants to compare item characteristics during presentation. In one condition, participants were asked to state how each item differed from the one immediately preceding it in the list;

in another condition, judgments of similarity were required. Once again, these judgments were made on lists either containing an isolate or not.

Two major findings emerged. First, when participants were asked to focus on item differences, the isolation effect was eliminated; second, when the orienting task was similarity-based, a robust isolation effect occurred. If we assume that the "difference-based" orienting task created list traces with little or no feature overlap, then the results follow nicely from the retrieval model. If traces already contain few, if any, matching features, then inserting an isolate—that is, an item with little or no matching features—should not enhance retention compared to a control. On the other hand, if the orienting task substantially increases the amount of feature overlap, by focusing attention on item similarities, then the isolate should be remembered especially well. It is interesting to note that recall of the background items followed the pattern predicted by the model as well. Difference processing led to a significant increase in background recall compared to the condition requiring similarity processing. Again, difference processing reduces the amount of feature overlap, and therefore the amount of cue overload, for both isolates and background items.

Other empirical patterns in the isolation-effect literature can be explained by reasoning of this sort. For example, it has been reported that the isolation effect is sometimes reduced or eliminated when participants report using elaborative rehearsal strategies during study (see Fabiani & Donchin, 1995). To the extent that elaborative processing leads to richer traces, ones that contain unique individual item information, then the isolation advantage should be reduced for the reasons discussed above. However, the predicted pattern will depend on the type of elaborative processing engaged. If participants relate items together, such as linking them into a cohesive story, then a very different pattern might well emerge. If the net result is an increase in feature overlap, because the processing emphasis has been placed on similarity rather than difference, the isolation advantage could increase.

## THE PARADOX OF SIMILARITY AND DIFFERENCE

Our discussion up to this point has centered on the importance of difference. For a given retrieval cue, sampling probabilities are inversely proportional to the amount of cue overload; consequently, manipulations that reduce feature overlap will increase the chances of correct target sampling. However, analysts of memory have known for decades that memory often benefits from the processing of similarities as well. For example, items from categorized lists are usually recalled better than items from unrelated lists (Tulving & Pearlstone, 1966); for unrelated word lists, relational processing, or the processing of commonalities among list items, can benefit recall substantially (e.g., Hunt & Einstein, 1981).

