Quarterly Journal of Experimental Psychology

ISSN: 0033-555X (Print) (Online) Journal homepage: http://www.tandfonline.com/loi/pqje19

Memorability, word frequency and negative recognition

J. Brown , V. J. Lewis & A. F. Monk

To cite this article: J. Brown , V. J. Lewis & A. F. Monk (1977) Memorability, word frequency and negative recognition , Quarterly Journal of Experimental Psychology, 29:3, 461-473, DOI: 10.1080/14640747708400622

To link to this article: http://dx.doi.org/10.1080/14640747708400622

Published online: 29 May 2007.

Quarterly Journal of Experimental Psychology (1977) **29**, 461–473

MEMORABILITY, WORD FREQUENCY AND NEGATIVE RECOGNITION*

J. BROWN, V. J. LEWIS† AND A. F. MONK‡

Department of Psychology, University of Bristol, 8–10 Berkeley Square, Bristol BS8 1HH, U.K.

Brown (1976) has provided an analysis of the effect of the memorability of an item on the confidence with which it is accepted or rejected in a test of recognition or recall. When the subject has no clear recollection of the inclusion of an item in an input list, he is assumed to evaluate its memorability in the context of the experiment before he decides whether to accept or reject it. If the judged memorability is high, the absence of a clear recollection is stronger evidence against the item than if it is low. A specific prediction is that memorable distractors in a recognition test will be more confidently rejected than non-memorable ones. This prediction was tested and confirmed in three experiments in which recognition was tested by 4-category rating. Except in Experiment I, items memorable to individual subjects were identified by administering a questionnaire. For example, in Experiment III forenames of immediate family were assumed to have high memorability. This experiment also included word frequency as a variable. Low-frequency distractors were rejected significantly more firmly than high-frequency distractors: extraction of memorable names enhanced this effect. The relationship of memorability to word frequency is discussed.

Introduction

One neglected problem of memory is how we can be certain of negative memories. All three authors of this paper are quite confident that they have not been introduced to Harold Wilson. One possible explanation is that the basis for a negative memory is simply the absence of a positive memory. This explanation is inadequate. We forget so much that the mere absence of a positive memory for an event cannot be reliable evidence that the event did not occur. There must be many thousands of introductions to people which each of us has forgotten. A related explanation can be formulated in signal detection theory terms. A positive memory results if trace strength is above a critical value and a negative memory if it is below this value. One problem with this explanation is how to interpret "trace strength". It is natural to assume that one's own name has a

* Based on a paper delivered at a meeting of The Experimental Psychology Society held in Reading, July 1976.

† Now at the Medical Research Council Applied Psychology Unit, 15 Chaucer Road, Cambridge, CB2 2EF, U.K.

‡ Now at the Department of Psychology, York University, Heslington, York, U.K.

high trace strength. Yet negative memories concerning one's own name are often very confident. (This is formally demonstrated in the experiments to be described.) Another possible explanation for confident negative memories is that they are based on inferences from positive memories. If a man remembers he has a weak heart, he will be confident that he did not run upstairs ten minutes ago. If a subject remembers that a list of words consisted only of animal names, he will then confidently reject distractors in a recognition test which are not animals' names. However it seems difficult to account for all confident memories in this way. Thus there are no obvious positive memories which enabled the present authors to be sure that they had not met Harold Wilson.

We have confident negative memories only about some possible events. These are events which we would have expected to remember if they had occurred. For example, meeting Harold Wilson would have been a memorable event. Accordingly, it is the absence of a positive memory for a subjectively *memorable* event which often forms the basis for a confident negative memory. This has been suggested as a possible basis for negative memories by both Brown (1976) and Groninger (1976).

Brown (1976) offered an analysis of ways in which the subjective memorability of an event affects its recognition or recall. In a laboratory experiment, the event is typically the occurrence of a particular word in an input list. When retention is tested, one of three things may happen. First, the subject may access a memory code which uniquely defines the occurrence of the word in the list. Highly confident recognition or recall will then ensue. Second, he may access a partial code compatible with more than one possibility. For example, this might make him uncertain whether SHEEP or GOAT was an input item. Third, he may fail to access a relevant code at all. In these second two cases, before he accepts or rejects an item whether in recognition or recall, he is assumed to assess its memorability. If the memorability is high, he is likely to reject it since the absence of a positive memory will then constitute quite strong *evidence* against its acceptance. If memorability is low, the absence of a positive memory constitutes only weak evidence that the event did not occur. An important feature of this analysis is that judged memorability affects the amount of *evidence* for accepting the occurrence of an event whether or not it also has an effect on the *criterion* for acceptance. (The analysis holds for correct judgments of memorability: incorrect judgments will mislead the subject in his assessment of whether an event, for which he has no positive memory, is likely to have occurred.)

Several interesting predictions can be derived from this analysis. First, overt intrusions in recall are likely to consist of items of low subjective memorability, since intrusions of high memorability are likely to be edited out. Second, intrusions in recognition (false positives) are also more likely with items of low memorability. Put the other way round, this means that distractors of high memorability should be more often rejected and rejected more firmly than distractors of low memorability. This prediction runs directly counter to a simple strength theory of memory. On such a theory, a memorable item presumably has a high resting strength so that when a memorable item occurs as a distractor its strength is liable to be comparable to the strengths of targets.

The third and final prediction concerns rejection not of distractors but of targets. A target of high memorability is of course less likely to be forgotten and therefore less likely to be rejected. The prediction is that when a subjectively memorable target *is* rejected it is likely to be firmly rejected. This follows because the absence of a positive memory for a memorable item will be strong evidence against its acceptance. This prediction is also difficult to reconcile with a simple strength theory. It applies to both recognition and recall but is difficult to test in the case of recall.

The following experiments are concerned with recognition and were primarily designed to test the prediction concerning rejection of memorable distractors. One problem in testing the prediction is that the strength with which distractors are rejected must be assessed independently of the overall level of recognition. There is a danger that rating categories will be used differently for memorable and non-memorable items. For example, the degree of evidence required for placing an item in the top category may be higher for memorable items. However, this is likely to occur only when memorable and non-memorable items are tested separately. If a mixed-list design is used, in which the subject encounters both types of item in random order, it is reasonable to assume that the category boundaries will remain the same for both types of item. This issue will be further discussed.

A second problem in testing the prediction is to identify which items are memorable. Strictly, the prediction concerns idiosyncratic memorability, viz. the subjective memorability of a particular item to a particular subject in a particular experiment. If memorability is assessed independently of the individual subject, the observed effect of memorability will be attenuated. If memorability is assessed by the subject himself, his judgment is liable to be contaminated by whether he thinks he remembers the inclusion of the item in the test. In Experiment I, memory is assessed independently of the individual subject. In Experiments II, III(a) and III(b) memorability is assessed from objective personal data given by the subject in response to a questionnaire. This represents a step in the direction of identifying idiosyncratic memorability.

In addition to memorability, Experiments III(a) and III(b) incorporate word frequency as a variable. The incorporation of both variables in a single experiment helps to show whether they represent independent dimensions of the stimulus material.

Experiment I

In this experiment, the input list consisted of surnames drawn from three different pools. The subjects were students. The first pool consisted of the names of famous people, the second the names of fellow students and the third the names of past students. These pools were expected to involve names of high, medium and low memorability respectively. In the rating recognition test, input items and further items drawn from these three pools were presented in random order. The strength with which distractors from each of these pools was rejected was assessed by the proportion placed in the lowest rating category.

The prediction derived from the memorability analysis is that the proportion will be highest for famous names, next highest for peer names and lowest for control names.

Method

Subjects

The subjects were 18 second-year honours students of both sexes studying psychology at Bristol University. In this and the other experiments, subjects were paid a nominal sum for their participation.

Stimuli

Each of the three pools consisted of 24 names, divided at random into two sets of 12 items each and designated Set A and Set B. For half the subjects Set A were list items and Set B were distractors in the recognition test and for the other half the roles of the sets were reversed. Thus the original list consisted of 36 names (three sets of 12 mixed together) and the recognition test consisted of 72 names (three sets of 24 mixed together).

Famous names. The names were: Sawyer, Bach, Brunel, Olivier, Bogarde, Jagger, Newton, Gladstone, Renoir, Barnard, Dickens, Shelley, Wills, Woolworths, Best, Freud, Parkinson, Murdoch, Biggs, Tajfel, Brubeck, Moore, Gagarin, Eisenhower. They were chosen from widely differing fields so that they would not make each other less memorable by competition.

Peer names. These were fellow second-year psychology honours students. Of the 24 names, 18 were the names of students participating in the experiment.

Control names. These were the names of former students in the Department of Psychology at Bristol University.

Procedure

The subjects were split into two groups of nine subjects, with each group tested separately. Group 1 received the Set A as the original list and Group 2 received the Set B as the original list. The list was presented by holding up 36 cards, each bearing one word, at a rate of approximately 2 s per card. The whole list was presented twice with a 10 s break between presentations. There was an interval of one week before recognition was tested. The common recognition list of 72 words was on a duplicated sheet with instructions at the top. These asked the subject to write a number beside each word according to the probability it was included in the original list using the following scheme: 1 = highly probable, 2 = probable, 3 = improbable, 4 = highly improbable. The session in which this experiment was performed also contained other memory experiments using different materials.

Results and discussion

Overall recognition was evaluated by calculating the R index (Brown, 1965). This is equivalent to the area under the ROC curve and can be derived from 4-category rating with negligible bias (Brown, 1974). The estimated values were 0·85, 0·81 and 0·79 for famous, peer and control names respectively. This suggests that it is legitimate to treat famous names as more memorable than peer names and peer as more memorable than control names. However, the results for peer names (but not other names) were different for the two groups of subjects, both when they occurred as distractors and when they occurred as targets: significant interactions were obtained in analyses of variance between type of name and group with non-significant main effects. It was therefore decided to analyse the results for famous and control names only. This has some additional justification on the grounds that these are the extreme groups in terms of memorability.

The proportions of distractors placed in category 4 were 0·52 and 0·42 for famous and control names respectively. (The proportion for peer names was 0·43). Analysis of variance showed that this difference was significant ($F = 8{\cdot}85$, $df = 2{,}32$, $P < 0{\cdot}01$) and there was no significant interaction with groups. Subjects' own names, which can be assumed to be highly memorable, occurred as distractors in four cases: they were all placed in category 4. These results accord with the prediction that distractors of high memorability will be more confidently rejected than distractors of low memorability.

Experiment II

In the previous experiment, famous names as a class were assumed to be memorable, and the results were consistent with this assumption. However it is clear that strong idiosyncratic factors operate. A famous name is not necessarily known to all subjects and may not be equally memorable even to those familiar with it. The most striking result was that subjects always placed their own names in the lowest rating category when they occurred as distractors. In Experiment II, the memorability of particular names to particular subjects was assessed by administering a questionnaire at the end of the experiment. Instead of surnames, the names of English towns were used as stimulus material. In this experiment and in the subsequent experiments, strength of rejection of distractors was assessed on the basis of the mean rating.

Method

Subjects

The subjects were 15 psychology students at Bristol University and 14 psychology students at York University. None had participated in the previous experiment (run a year earlier).

Stimuli

A pool of 80 names of well-known English towns and cities was drawn up. From this pool, 20 names were randomly selected and mixed with 20 names of fruit and vegetables to form the input list. The fruit and vegetable names were included as filler items to increase the difficulty of subsequent recognition of town names. The recognition test consisted of all 80 items from the pool in random order on a recognition sheet. A questionnaire was constructed which asked the subject to write down (a) his home town, (b) any other towns in which he had lived for more than a year, (c) the names of universities to which he had applied, (d) the five towns of most personal significance, excluding Bristol, York, Bath and London. In the analysis of results, all the towns listed in the questionnaire were classed as cited items for that subject.

Procedure

The subjects were tested in groups. The input list was presented using a carousel projector at a rate of 2·5 s per slide/name. There was then a 15 min retention interval before recognition was tested. During this interval an experiment involving presentation, generation and recall of foreign town names was conducted followed by a cancellation task lasting 5 min. As in the previous experiment, the recognition sheet contained instructions at the top but the rating scheme now read: 1 = the English town was definitely on a slide, 2 = the English town was probably on a slide, 3 = the English town was probably not on a slide, 4 = the English town was definitely not on a slide. The test was untimed. As each subject completed the recognition test he was handed the questionnaire to complete.

Results and discussion

The initial analysis of the results was disappointing. The questionnaire was intended to elicit the names of towns which would be subjectively memorable. On the memorability hypothesis, these towns would be more confidently rejected by the subjects who cited them than not-cited towns. In fact the mean ratings for cited and not-cited towns were almost equal. Examination of the raw data suggested that this was due to the inclusion of town names cited in answer to the question (c). This question asked for the universities to which the subject had applied: it seems that applying to a university was too transient an event for the corresponding town name to become memorable. These names constituted about 40% of all cited names. When they were excluded from the analysis the mean ratings for cited and not-cited names became 3·30 and 3·00 respectively. (In this analysis, a name cited in answer to question (c) but also cited in answer to another question was *not* excluded). The difference is statistically significant on a Wilcoxon test ($P < 0{\cdot}02$, two-tail) and implies more confident rejection of cited names. The most striking result was for home towns. In all eight cases where the home town occurred as a distractor it was placed in category 4. Thus, as with own names in Experiment I, home towns were rejected both confidently and correctly.

Taken on its own this experiment would not be convincing. Post hoc elimination of a sub-set of cited names cannot be defended as a method of reaching firm conclusions. Nevertheless the results do provide some support for the memorability hypothesis. The need to exclude a sub-set of cited names emphasizes the importance of what questions are included in the questionnaire if memorable names are to be reliably identified.

Experiments III(a) and III(b)

The results of Experiments I and II have shown that memorable names are more confidently rejected in a recognition test than other names. Memorable names will often have a high personal frequency of usage. This is obviously true, for example, of the subject's own name and of the name of his home town. Word frequency is known to affect recognition. (For a review, see Gregg, 1976). Moreover, Glanzer and Bowles (1976) have recently shown that word frequency affects both the probability of accepting a target and the probability of rejecting a distractor. The question arises as to whether memorability and frequency are distinct variables capable of being separated experimentally. The two following experiments concern recognition of forenames with both memorability and word frequency as variables. Word frequency was assessed indirectly from the frequency with which different forenames were produced in a forename generation task. The experiments employ a mixed-list design. This is necessary to ensure that the category boundaries are the same for high-frequency and low-frequency names. (The overall effect of word frequency is unaffected by whether a mixed list or an unmixed-list design is used—see Gregg, 1976). In Experiment III(a), an unrelated task intervened before presentation of the recognition test. In Experiment III(b) the intervening task was to generate forenames and rate each forename for the probability that it was included in the list. This experiment was part of a separate study of the generation process. It provided an opportunity for the partial replication of Experiment III(a) but was not specifically designed to investigate the effects of the generation task on subsequent recognition. If a sufficient number of both high-frequency and low-frequency names are generated, such a task might be expected to influence the effect of the word frequency variable.

Method

Subjects

In Experiment III(a) there were 15 subjects and in Experiment III(b) there were 27 subjects. The subjects were all psychology honours students and were taken from different years for the two experiments.

Stimuli

Data were available from previous work in our laboratory on the frequency with which students generate different forenames when asked to write down as many forenames as possible within a limited time. The names generated were arbitrarily divided into a high-frequency of generation group and a low-frequency of generation group.

The high-frequency names were:

Alison, Andrew, Anne, Barbara, Bill, Brian, Catherine, Charles, Christopher, David, Diane, Donald, Edward, Elizabeth, Eric, Fiona, Frances, Fred, George, Gillian, Grace, Harold, Helen, Henry, Ian, Ivor, Ivy, James, Jane, John, Kate, Kathleen, Keith, Leonard, Linda, Lucy, Margaret, Mark, Mary, Nicola, Nigel, Nora, Pamela, Paul, Peter, Richard, Robert, Ruth, Sarah, Simon, Susan, Teresa, Timothy, Tina, Valerie, Victor, Victoria, Walter, Wendy, William.

The low-frequency names were:

Ailsa, Ambrose, Arabella, Beatrice, Benjamin, Bertrand, Clarissa, Clifford, Clodagh, Dermot, Desmond, Dorothea, Eloise, Emmeline, Errol, Fenella, Fletcher, Fritz, Gavin, Gudrun, Gwenda, Hazel, Hector, Hubert, Ingmar, Ingrid, Irvin, Joad, Jocelyn, Jonquil, Kelvin, Kirk, Kirsty, Leopold, Lilith, Loretta, Marlene, Marvin, Mervyn, Nanette, Nathaniel, Noel, Patten, Perdita, Petra, Rollo, Rosanne, Rowena, Shane, Sheena, Sybil, Tarik, Theodosia, Tobias, Vanessa, Vincent, Viviana, Warren, Wilfred, Winifred.

From the high-frequency group, a pool of 40 male forenames and a pool of 40 female forenames were selected. Two corresponding pools were selected from the low-frequency group. The input list consisted of 10 forenames selected at random from each of the four pools (making 40 forenames in all) plus the names of 20 English towns. The town names were included as fillers to lower the level of recognition of forenames. The order of all the names in the input list was randomized. The recognition test for forenames consisted of 20 rows of six names per row. All the six names in a given row began with the same letter: two were targets and four were distractors, three were male names and three were female names, three were high-frequency and three were low-frequency names.

A questionnaire was prepared which asked for the following information: (a) the subject's own forenames, (b) the five forenames of the highest personal significance, (c) the forenames of close family members, (d) the forenames of close friends, (e) the forenames of famous people occurring to the subject.

Procedure

Experiment III(a). Subjects were told that they would be shown 60 slides showing town names and forenames and that their memory for either or both sets of names would be tested later. The slides (one name per slide) were then presented at 2·5 s intervals. There followed a retention interval lasting about 10 min. During this interval, short humorous anecdotes were read to the subjects who were then asked to give written answers to questions about these anecdotes. Next, the recognition test was given together with written instructions to rate each word according to the scheme used in Experiment II, with "forename" substituted for "English town". As each subject completed the recognition test, he was handed the questionnaire to complete.

Experiment III(b). The procedure was identical to Experiment III(a) except for the retention interval task. For this task, each subject received a sheet containing an instruction to write down as many forenames as possible and to rate each forename using a 4-category rating scheme similar to the one used for the recognition test. The retention interval itself was again 10 min.

Results and discussion

At the outset, certain difficulties in the analysis of the results should be pointed out. The number of cited or not-cited names depends not on the design of the experiments but on the answers given in the questionnaires. They also vary from subject to subject. Among cited names, high-frequency and low-frequency names are not equally represented. In particular, some subjects do not cite both low frequency targets and distractors. Consequently, this rules out analysis of variance. Instead individual comparisons must be tested one at a time. Fortunately the important aspects of the results emerge clearly and appear to be reliable.

Memorability

The results of Experiments III(a) and III(b) were similar both to one another and to those obtained in Experiments I and II and accord with the prediction that memorable names are easier to reject when they occur as distractors. The following results are for Experiments III(a) and III(b) combined, unless otherwise stated.

The most striking results were for the subject's own name. In the 12 cases where it occurred as a distractor it was always placed in category 4. In order to examine the effect of memorability on other names, names cited and not cited in the questionnaire were then identified for each subject with his or her own name excluded. The mean ratings for distractors were 3·46 and 3·29 for cited (excluding own) and not-cited names respectively. This difference is statistically significant ($P < 0{\cdot}01$) on a two-tail Wilcoxon test and is in the direction predicted on the memorability hypothesis. The difference is also significant ($P < 0{\cdot}05$) if Experiment III(b) is taken on its own but is not significant in Experiment III(a) which employed fewer subjects. The difference between the differences (involving two independent groups of subjects) did not approach statistical significance on a Mann–Whitney test.

An interesting result for targets was that there were *more* cited targets firmly rejected than weakly rejected (31 placed in category 4 as compared with 19 in category 3), whereas the converse was true for not-cited targets (144 in category 4 as compared with 163 in category 3). Although this result is difficult to test statistically, it accords with the prediction derived from the memorability hypothesis that, *when* memorable targets are rejected, they are likely to be firmly rejected. (The numbers of targets placed in categories 1 and 2 were 213 and 13 respectively for cited targets as compared with 913 and 182 for not-cited targets).

Overall recognition was slightly better for cited names, excluding own names, than for not-cited names (the R values were 0·89 and 0·86 respectively).

Word frequency

As with memorability, the trends for word frequency were similar in Experiments III(a) and III(b). However there was some evidence to suggest that the forenames generation task in Experiment III(b) affected the relative magnitudes of the trends (see later). Accordingly the results of the two experiments will be presented separately. Table I gives the mean ratings from both experiments.

TABLE I

Mean ratings in Experiments III(a) (15 subjects) and III(b) (27 subjects)

	1. Overall		2. Excluding cited		3. Cited only	
	HF	LF	HF	LF	HF	LF
Targets III(a)	1·76	1·62	1·81	1·64	1·62	1·0
Targets III(b)	1·77	1·46	1·89	1·48	1·53	1·0
Distractors III(a)	3·07	3·32	3·02	3·32	3·34	3·64
Distractors III(b)	3·34	3·39	3·30	3·39	3·52	3·75

The ratings in the first pair of columns imply that low-frequency distractors were more confidently rejected than high-frequency distractors in both experiments. The difference between the mean ratings was statistically significant in Experiment III(a) on a Wilcoxon test ($P < 0{\cdot}02$) but not in Experiment III(b). Overall this implies that low-frequency distractors were more confidently rejected than high-frequency distractors. If memorability and word frequency were one and the same variable, the opposite result would have been expected: memorable names (especially the subject's own name) can be assumed to have a high personal word frequency so that high memorability and high word frequency should have affected recognition in a similar manner. The ratings for targets in the first pair of columns imply that low-frequency targets were more confidently accepted than high-frequency targets in both experiments. Here the difference between the mean ratings is statistically significant on a Wilcoxon test ($P < 0{\cdot}002$) in Experiment III(b) but is not significant in Experiment III(a). Thus low-frequency targets tended to be more confidently accepted than high-frequency targets. Again, the opposite result would have been expected if memorability and word frequency were essentially the same variable.

Since memorability and word frequency affect recognition in opposite ways, it should be possible to reveal the effect of word frequency more clearly by removing memorable names from both the high-frequency and the low-frequency names. The second pair of columns in Table I show the mean ratings with cited names removed. As expected, the contrast between the ratings for high-frequency and low-frequency distractors is increased. The difference between ratings now becomes more highly significant in Experiment III(a) ($P < 0.002$ instead of $P < 0.02$) although it remains non-significant in Experiment III(b). The contrast between the ratings for high-frequency and low-frequency targets is also increased.

The third pair of columns in Table I show the mean ratings for cited names only. Only seven subjects contributed to these means. (Note that a subject contributed only if he happened to cite target and distractor names in his questionnaire). It is of interest that high-frequency and low-frequency names still appear to behave differently although with the contrasts between them somewhat reduced.

The apparent differences between the results of the two experiments with respect to word frequency warrants some comment. The difference between the mean ratings for high-frequency and low-frequency distractors was significant in Experiment III(a) but not in Experiment III(b). Conversely, the difference for targets was significant in Experiment III(b) but not in Experiment III(a). For distractors, the difference between the differences obtained in the two experiments just reached significance on a Mann–Whitney test ($P < 0.05$) but the corresponding difference for targets was not significant. It is not clear therefore how much importance should be attached to the apparent differences in the results of the two experiments. Nevertheless, it is worth noting that the forenames generation task in Experiment III(b) may affect ratings in the recognition test. The proportions of high and low frequency targets generated were approximately 0·68 and 0·37 respectively: the corresponding proportions for distractors were 0·50 and 0·20 respectively.

Discussion

All the experiments provided evidence that distractors of high memorability are rejected more firmly than distractors of low memorability. The most striking result was that obtained when the name of the subject or of his or her home town happened to occur as a distractor in the recognition test. Such names can be presumed to be highly memorable: they were always placed in the lowest rating category (30 cases summed across experiments). The results therefore support the view that a negative recognition decision is influenced by judged memorability, as suggested independently by Brown (1976) and Groninger (1976). If the subject judges the item concerned to be of high memorability, the absence of a positive memory constitutes stronger evidence against that item than if he judges it to be of low memorability. This analysis is made more plausible by Groninger's finding (Groninger, 1976) that subjects are good at predicting which item in an input list they will subsequently recognize successfully.

In Experiments III(a) and III(b), there was suggestive evidence that memorable targets are rejected more firmly than non-memorable targets on the rare occasions when they are rejected. This finding needs confirmation. Because memorable targets are only rarely rejected only a large-scale experiment will generate enough data. If the finding is accepted at its face value, it is entirely in keeping with the view that the absence of a positive memory is taken as strong evidence against a memorable item. Indeed it would be paradoxical if, in the absence of a positive memory, memorable distractors were firmly rejected but not memorable targets.

If judged memorability plays a role in memory decisions, this emphasizes that human cognitive decisions are complexly determined. A simple strength analysis of recognition memory can be seriously misleading. This does not mean that it is inappropriate to use a signal detection analysis of recognition memory. However, the recognition decision is determined not by trace strength or familiarity but by the amount of evidence from all sources in favour of the item. This has been called the "plausibility of the item" (Brown and Routh, 1970). The underlying distributions for targets and distractors are distributions along a dimension of plausibility. (For other arguments against assuming that recognition is based on trace strength or familiarity see Anderson and Bower, 1972; and Brown, 1976.)

The overall results on word frequency (or, strictly, generation frequency) were as follows. Low-frequency targets were more confidently accepted than high-frequency targets and low-frequency distractors were more confidently rejected than high-frequency distractors. Thus the well-known superior recognition of low-frequency words arises because word frequency affects *both* acceptance of targets and rejection of distractors. This conclusion is supported by a recent study of Glanzer and Bowles (1976). In their recognition test, pairs of words were presented and subjects made a forced choice between the words of each pair. Most pairs consisted of a target and a distractor, i.e. standard 2-AFC. Unknown to the subjects, some pairs consisted of either two targets or of two distractors. When a pair consisted of one high-frequency target and one low-frequency target, the latter was chosen significantly more often. Conversely, when it consisted of one high-frequency distractor and one low-frequency distractor, the former was chosen significantly more often. Thus low-frequency favourably affected both the acceptance of targets and the rejection of distractors. These results, using a forced choice method, confirm those obtained by the rating method in the present study. Potentially the rating method is susceptible to shifts in the use of rating categories. It was argued that such shifts are unlikely in a mixed-list design in which the subject encounters high-frequency and low-frequency targets and distractors in random order. Moreover, Brown (1974) found that 4-category rating and multiple-response forced-choice recognition gave closely similar estimates of recognition efficiency in terms of the *R*-index (the area measure of signal detection theory). This helps to encourage the view that shifts in category boundaries are not a serious problem with the rating method.

An important issue is the relationship between memorability and word frequency. It is clear that the memorability of a word and its frequency of usage are at least partially independent dimensions. Memorability and frequency affect recognition differently. High memorability is associated with good recognition whereas high frequency is associated with poor recognition. A subject's own name has a high frequency of usage for him and yet is reliably recognized: this includes confident rejection if his or her name occurs as a distractor.

If memorability and word frequency are partially independent dimensions, do we need separate explanations for their effects on recognition and especially their effects on distractors? The answer seems to be "No". Word frequency helps to determine memorability. A low-frequency word tends to be more memorable because it tends to have more distinctive meanings and associations (cf. Gregg, 1976). A memorable name with a high frequency can also have distinctive associations and meanings; one's own name is a prime example. In order to state this analysis more fully, it is useful to draw a distinction between the public and private properties of a word. The public properties are those it has for all or most speakers of the language. The private ones are additional properties it has for a particular individual. The distinctiveness of both public and private properties may be either high or low. Low-frequency words tend to be memorable because of their distinctive public properties and sometimes because of their distinctive private properties as well. High-frequency words tend to be memorable only when they have distinctive private properties. However, it would be an over-simplification to imply that the memorability of a name is entirely a matter of the distinctiveness of its pre-existing properties. A memorable name—especially one's own—may receive preferential treatment at input: it may be rehearsed more and coded more elaborately.

This analysis of the relationship between word frequency and memorability helps to explain why word frequency appears to have appreciable effects even amongst memorable words, as shown by the third pair of columns in Table I. Memorable names were defined as names occurring in the questionnaire responses. Such names will not have been of equal memorability. Variations in the memorability of cited names may occur for a number of reasons, one of which is word frequency. It is plausible to suppose that if your sister has an unusual forename this will have higher subjective memorability than if she has a very common name which you constantly encounter in many different contexts.

One methodological implication of the relationship between word frequency and memorability is that the effect of word frequency on recognition can be studied more effectively if words which are memorable for reasons unconnected with word frequency are removed from the results. This was achieved in Experiment III by removing names cited in the questionnaire. This method could be made more effective by pre-testing questions before including them in the questionnaire. The method has two disadvantages. One is that it involves dealing with the data from each subject on an individual basis. The second is that it is apt to lead to difficulties in the statistical treatment of the results.

A number of interesting questions arise concerning memorability and the rejection of distractors. Is memorability always taken into account consciously? Or can it be a rapid covert process, like so many of our cognitive activities? Does it always influence recognition or only sometimes? Does the effect of memorability disappear when a recognition decision must be taken very quickly? Do young children rely less on memorability judgments than adults? Can subjects be misled about subjective memorability? These and other questions can only be answered by further experiments.

This research was part of an investigation of the relation of recall to recognition which was supported by the Social Science Research Council.

References

ANDERSON, J. R. and BOWER, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, **79**, 3, 97–123.

BROWN, J. (1965). Multiple response evaluation of discrimination. *British Journal of Mathematical and Statistical Psychology*, **18**, 125–37.

BROWN, J. (1974). Recognition assessed by rating and ranking. *British Journal of Psychology*, **65**, 1, 13–22.

BROWN, J. (1976). An analysis of recognition and recall and of problems in their comparison. In BROWN, J. (Ed.), *Recall and Recognition*. London and New York: John Wiley & Sons.

BROWN, J. and ROUTH, D. A. (1970). Recognition assessed by d' and by a non-parametric alternative (The A-Index) as a function of the number of choices. *Quarterly Journal of Experimental Psychology*, **22**, 707–19.

GLANZER, M. and BOWLES, N. (1976). Analysis of the word frequency effect in recognition memory. *Journal of Experimental Psychology*, (Human Learning and Thinking), **2**, 21–31.

GREGG, V. (1976). Word frequency, recognition and recall. In BROWN, J. (Ed.), *Recall and Recognition*. London and New York: John Wiley & Sons.

GRONINGER, L. D. (1976). Predicting recognition during storage: the capacity of the memory system to evaluate itself. *Bulletin Psychonomic Society*, **7**, 425–8.

Received 13 October 1976