Superordinate Schema as Mediators in Associative Learning[1]

Gordon H. Bower

Stanford University

Running Title: Schema in Learning

ABSTRACT

We tested a mathematical model of associative recall. Subjects studied 80 sets of four vaguely related words, then were tested for recall by cueing with one or two words of each quartet. The four words were jointly related to some single concept or schema. Half the subjects were told what the relating schema was as they studied each quartet; the remaining subjects had to discover it themselves. The groups did not differ in recall, so were pooled for model testing. The best fitting model states that the words become interconnected via a central memory schema, that each cue word has independent probability *a* of gaining access to the memory schema, and once assessed the schema generates each of the other items in the quartet with independent probability *r*. The model successfully fit the recall distributions obtained with one- and two-cue tests.

The traditional view of word associations is that they are direct links from the memory representation of one word to the representation of another. Thus, dog is directly linked to bark, and table to chair. However, in light of modern theories, it seems more likely that it is not words that are linked but rather concepts to which they refer. Moreover, it seems likely that the association between two concepts is often derivative in the sense that their worldly referents are related in some way. Thus, dogs do bark, and chairs appear with tables in multiple scenes. Often it is clear that words are not associated at all in isolation, but rather become so only as the person can retrieve a concept, schema, or "scripts" (see Schank and Abelson, 1977) within which two concepts of the words play some role. In such a case, the words refer to elements that are objects or attributes of larger schema. For example, consider the words grease, curtain, prop, program. Because words have multiple interpretations, the first two words suggest, say, a kitchen scene; but the next two words do not fit at all into kitchen. Rather they suggest theater play. Once that schema is aroused, grease (as in grease-paint) and curtain can be given a schema-related interpretation. Consider another example: apron, chair, brush, clip. Any two of the words suggest some ideosyncratic scenario, but all four words together can be seen to fit into a haircut schema. There is a sense in which it is incorrect to say that two words like apron and brush are associated in isolation; rather, they can be seen as connected within the larger context of getting a haircut.

An implication of the foregoing is that the learning of a cluster of four such words should be easy if the person "sees" the relation among the words--that is, if he activates a schema within which the items have their place. Activating a relevant schema is then equivalent to exciting pre-existing inter-relations among these words, so that at worst the person must only remember which items of the larger script were mentioned. Contrariwise, if the words do not arouse an appropriate schema, then they will continue to appear as unrelated words and should be difficult to learn as new clusters. In the vocabulary of verbal learning, the schema is a "mediator" for learning the words; it differs from the usual verbal-learning mediators insofar as the items subsumed under the script are already "inter-associated" prior to the subject's entry into the experiment.

In the experiment below, we tried to vary the likelihood that the subject would infer the appropriate script for the quartet of words. To do so, we deliberately selected quartets of words which were only vaguely and minimally related to a common script or schema. The words were often ambiguous, and no single word strongly suggested the underlying script. By choosing such minimally script-related words, we hoped to create a condition in which the script was not likely to be activated when the words were studied. Therefore, the word quartets alone were presented for learning to half the subjects. For the remaining subjects, as each word quartet was presented for study, the experimenter also read aloud the name of the script or schema (e.g., "haircut," "Theater play") within which the items were interconnected. It was expected that subjects given the script mediators would better interassociate the words within a quartet than would subjects not given the mediator.

A second, and the major, aim of this experiment was to collect recall results to test a particular mathematical model of associative recall. The basic idea of the model is that the items to-be-learned are not associated directly to one another, but rather become connected into a larger memory structure such as a script or a proposition. The basic diagram of the associative structure is shown in Figure 1.

Insert Figure 1 about here

The letters A, B, C, D denote the four words of a cluster that is to be learned, and the central hub represents the particular memory schema activated contiguously with the words and within which the words play some part. (There would be one such associative structure for each quartet the subject studies in a list.) The arrows refer to associative connections. These are drawn as one-step links but this is a short-hand for a chain of at least two links: the word is linked to a concept (possibly many, but only one is relevant here), and a token of the concept is linked into the particular schema or propositions used to inter-relate the items when they are studied. For details of this representation, see the theory of associative memory in the book by Anderson and Bower (1973).

Each item in the associative memory trace is characterized by two parameters, a and r. The parameter a (for access) is the likelihood that, at the time of retention testing, presentation of that word independently will allow access to the relevant memory schema. We will think of this as an all-or-none phenomenon: at any particular test occasion, the word cue either fully accesses and activates the memory schema or it fails completely. The parameter r (for response) is the likelihood that, during retention testing, the word will be independently recallable once the memory schema has been activated. Again, the linkage from the schema to output of a word may go through intermediate concepts, but we use r to summarize the probability of success of the entire schema-to-output chain.

In the experiment below, the four words were chosen to be equal, on average, in factors affecting their memory (e.g., concreteness, association to their schema, frequency). Also, we randomized over subjects which items of the cluster were used as cues and which as responses. By pooling over the four items and not distinguishing them, we may then treat the four words in each cluster as equivalent for purposes of cueing and recall. This means in particular that we may treat the four words as having the same parameters $\underline{a}$ and $\underline{r}$. This item-homogeneity assumption simplifies the theoretical equation considerably; one becomes concerned in a given cueing condition with only how many responses of the cluster are recalled, not with which responses are recalled.

One memory test we used consisted of giving one word from the quartet and asking the subject to recall as many of the other three words from that quartet as he could. Let $\underline{X}$ be a random variable denoting the number of the three remaining words recalled after a single-word cue. The probability distribution of X can be shown to be as follows:

$$\Pr\left\{\underline{X} = \underline{i}\right\} = \begin{cases} 1 - \underline{a} + \underline{a}\,(1-\underline{r})^3 & \text{for } \underline{i} = 0 \\ \underline{a}\,\binom{3}{\underline{i}}\,\underline{r}^{\underline{i}}\,(1-\underline{r})^{3-\underline{i}} & \text{for } \underline{i} \geq 1 . \end{cases} \qquad (1)$$

To understand some of the terms in Equation 1, we note that the cue word gains access to the schema with probability $\underline{a}$; once the schema is accessed, there is independent binomial probability, $\underline{r}$, that each of the three responses will be recalled. The probability that the cue leads to no recall ($\underline{i}$ = 0) is the likelihood that it fails to activate the schema, 1-$\underline{a}$, or that it activates the schema except none of the three correct responses can be output, $\underline{a}(1-\underline{r})^3$. It's clear that Equation 1 can be generalized to any size of memory cluster. For a cluster of size N cued by one of its elements, replace the 3 in Equation 1 by N-1 to obtain the appropriate equation.

A second memory test in the experiment used two words from a given quartet, and the subject was asked to recall whatever he could of the remaining two items. Let $\underline{Y}$ denote the number of such items recalled (0, 1 or 2) with two cues. The probability distribution of $\underline{Y}$ is as follows:

$$\Pr\{\underline{Y} = \underline{k}\} = \begin{cases} (1 - \underline{a})^2 + [\underline{a} + \underline{a}\ (1 - \underline{a})]\ (1 - \underline{r})^2 & \text{for } \underline{k} = 0 \\ [\underline{a} + \underline{a}\ (1 - \underline{a})]\binom{2}{\underline{k}}\underline{r}^{\underline{k}}\ (1 - \underline{r})^{2-\underline{k}} & \text{for } k \geq 1. \end{cases} \qquad (2)$$

The terms of Equation 2 are intuitively sensible. The term $\underline{a} + \underline{a}(1 - \underline{a})$ is the probability that one or the other cue-word accesses the relevant schema in memory. Once that schema is accessed, the to-be-recalled words are output with independent probability $\underline{r}$. This yields the binomial distribution shown, with an additive term at $\underline{k} = 0$ to reflect the possibility that the memory schema is not accessed at all. To generalize, for an $\underline{N}$-element cluster, and two cues, the corresponding equation would have $\underline{N} - 2$ replacing the 2 in Equation 2. In fact, we may generalize further and write an equation for use of $\underline{J}$ independent cues for recall of the remaining $\underline{N} - \underline{J}$ items from an $\underline{N}$-item cluster. The probability of recalling $\underline{k}$ elements is a product of (1) the probability that at least one cue activates the schema, which is $1 - (1 - \underline{a})^{\underline{J}}$, and (2) the binomial probability that $\underline{k}$ elements out of $\underline{N} - \underline{J}$ will be recalled, each with probability $\underline{r}$.

The attractive feature about Equations 1 and 2 is that two-cue results are relatable to one-cue results because the same parameters, $\underline{a}$ and $\underline{r}$, are involved in both equations. This arises because we have assumed that (1) the words act independently as accessing cues, and (2) once the schema is activated, the subject outputs each word independently with the same probabilities. The equal-probability assumption can be more or less guaranteed by randomization and then pooling of cues and responses. However, the independence assumption is a substantive assumption that is empirically testable. The experiment below permits a test of these independence assumptions by seeing whether Equations 1 and 2 will fit the one-cue and two-cue recall results.

METHOD

Design

Two groups of subject had one study-recall cycle on each of four lists of 20 quartets of words. For one group of subjects, the experimenter named the underlying script (e.g., "haircut") as the word quartet was shown during study; for the other group, he said nothing and they simply studied the quartet. After study of a list, recall of it occurred: half the quartets at random were cued for recall by presenting a single word and half were tested by presenting two words from the quartet. In each case, the person was to recall the remaining words of the cued quartet.

Materials

We constructed 80 quartets of words for this experiment. We began with some scenario (e.g., making a movie, surfing, riding a subway) or object (e.g., a can of beer, a clam, a business office) or event (e.g., July 4th Parade, Attack on Pearl Harbor). We then selected unique words that, by themselves, were only weakly related to the schema. Some examples will illustrate the quartets: princess, mouth, hold, dial for telephone; bag, fire, penknife, boots for camping; scratch, collar, friend, license for dog; driver, trap, rough, handicap for golf; and glove, ring, bruise, bell for boxing. Due to the unsystematic selection procedure, the words within a quartet could not be equated on concreteness or frequency of usage; indeed, many of the words used do not appear in the standard norms for association or concreteness. However, our selection procedure clearly resulted in quartets that were intuitively related, particularly once the schema was mentioned. A single word appeared in only one quartet.

The 80 quartets were divided randomly into four lists of 20 and each subject studied and was tested on all four lists in a random order. For each list, we constructed eight different recall test lists (one per subject per condition). The eight test lists cued recall of a given quartet in eight different ways--four times with each of the single-word cues, and four times with a differently selected pair of the four words as cues. Thus, across subjects within a condition, a given cluster memory was cued in a statistically balanced fashion. The one or two cue words were printed down the left side of a recall sheet given to the subject after he had studied each list. He wrote his recall of the other two or three words in the cued cluster on the page to the right of the cues.

Procedure

The subjects were fully instructed about the learning task, about the quartets to be learned and the cueing method of recall. In addition, the Informed subjects were told that the four words of a cluster were about some specific object, activity, or event, and that they would be told what that was as they saw the four words. They were to call to mind the object or scenario, to see how the words all fit together in that scene, and to memorize it. They were also instructed to call this scene back to mind from the cue word to do their recall. Control subjects were simply told to study the word quartets as a group to be remembered.

The words were presented on a screen in front of the subjects via an overhead projector. The quartets, were presented one at a time for ten sec. each. For Informed subjects, the name of the mediating schema was said aloud by the experimenter immediately as each quartet was shown. For recall, the subject had a booklet with four recall pages for the four lists, separated by blank cover sheets. After studying a quartet list, the subject then turned to the recall page, which he covered with a black cardboard mask. Upon signal from the experimenter, he slid the mask down the page one line at a time, exposing (in a 0.5 by 5.0 inch slit) the successive word cues and writing his recall words to the right of the cues. The signal to move the mask down was given every 15 seconds. The mask tended to isolate recall testing of each quartet since the subject could not see other cues or responses as he worked on recalling the cued quartet. A one-minute rest occurred after each test. The subjects knew that each quartet list would be studied and tested just once.

Subjects

The subjects were 16 Stanford undergraduates participating to fulfill a service requirement for an Introductory Psychology class. They were run in small groups of 1 to 4. The groups were assigned in semi-random order to the Informed and Control conditions, insuring that eight subjects ended up in each group.

RESULTS

Group Comparisons

The words recalled for each quartet were noted and counted as correct or incorrect accordingly as they duplicated the sense of the words originally presented. Lenient scoring was used, but the conclusions are not affected by stricter scoring. Each quartet was scored according to the number of responses recalled to the one or two cues. Table 1 below shows the complete frequency distributions of the 320 scores for the Informed and Control groups.

- - - - - - - - - - - - -

Insert Table 1 about here

- - - - - - - - - - - - -

The observed frequency distributions do not differ according to a chi-square text, $X^2(6) = 10.92$, $p > .05$. The means were also computed and compared. For the one-cue condition, the average number of words recalled (out of 3) was 1.26 for the Informed group and 1.02 for the Controls. These do not differ significantly, $t(14) = 1.30$, $p > .10$. For the two-cue test, the average number of words recalled (out of 2) was 1.15 for the Informed subjects and 1.03 for the Controls, which do not differ significantly. We may conclude that explicit naming of the schema during study had no differential impact.

In interviewing the control subjects, it became obvious why the groups did not differ. The four words of a quartet suggested a specific schema for the Control subjects; they all reported that nearly all quarets reminded them of some scene or activity as they read the words. Within a few "aha" experiences on the first study list, the Controls were aware of how the quartets were constructed, and so they searched for single concept mediators to memorize later quartets. The Control subjects recalled slightly worse on the first list before they had "caught on" to the nature of the quartets, but the difference was not statistically significant. To the extent that Control subjects are activating the same memory schema as are Informed subjects, the two groups become functionally similar, so they should have equivalent levels of recall.

One aspect of the data where the groups did differ is in the nature of their intrusion errors. While they gave similar totals of intrusions, the Informed subjects gave the actual schema label rather more frequently than did Control subjects, with an average of 5.8 versus 1.6 such intrusions per subject for Informed vs. Controls, respectively. (Total opportunities for such an intrusion was 80 per subject). The Informed subjects also gave more intrusions that were thematically related to the schema, with means of 14.0 per Informed subject vs. 10.5 for Control subjects. Fully 71% of the intrusions for the Informed subjects (vs. 47% for Controls) were either the schema-word or thematically related to the schema. Finally, the Control subjects gave more intrusions of unrelated words (8.0 per subject) than did Informed subjects (1.2 per subject). The conclusion from these intrusions is that Informed subjects were somewhat more likely to make errors based on the schema activated during learning.

Model Fitting

The most interesting results of this experiment concern the fit of the mathematical model. Equations 1 and 2 were first fit to the distributions of one-cue and two-cue recalls for each subject individually. The model fit the individual data exceedingly well, with only one subject of the 16 yielding a significant discrepancy. Summing the chi-square values for the individual fits (each with 3 degrees of freedom) yields a total chi-square of 41.00 on 48 df, which indicates a very good fit of the model. There were also significant individual differences among subjects within groups. Table 1 shows the fit of Equations 1 and 2 to the observed frequencies pooled over subjects within groups. The parameters, a and r, were estimated for each group. The parameters, a and r, were estimated for each group separately using a programmed minimum chi-square method. The parameters estimated are: for the Informed subjects, $a = .589$ and $r = .702$; for the Control subjects, $a = .514$ and $r = .668$. The fit of the model is excellent, for both single-cue and double-cue results. The minimum chi-square obtained was 0.83 for the Informed subjects, and 0.22 for the Control subjects. The degrees of freedom for each chi-square is 7-2-2 = 3, derived from seven frequencies, two sets of which add to 320, and two parameters are estimated. The chi-square values are far from significant, reflecting the closeness of fit. The two groups differ more in schema-access parameters (a) than they do in the response-output parameters (r). In fact, the goodness of the fits is not materially worse if a common response parameter is used for both groups.

The good fit of the model implies that the cues have independent access probabilities, and that, once schema access occurs, the responses are output independently with the same probability. The fit to the two-cue data from the one-cue parameters implies, for example, that on the average the two cues are adding independently in their probability of accessing the memory schema. This is not to say that there will never be "configural" information emerging from the two cues (see, e.g., Foss and Harwood, 1975); rather, the present data gives no evidence for it with these materials.

The power of the model's fit to the data can be strengthened by comparing it to the fit of plausible theoretical alternatives. One such alternative is the simple idea that each cue word can cue recall of any other word in the quartet with probability θ. Assuming the items are independent, this model predicts a simple binomial distribution for the number of items recalled. It differs from the schema model in that the schema model has more probability density at $X = 0$ and $Y = 0$ than the simple binomial expects, due to the fact that the cue word may fail to access the core schema. The schema theory reduces to the binomial model when schema-access is assured ($a = 1$). For comparative model testing, we pool the data from the two groups. The expected recall for the one-cue test for the binomial model is 3θ. Equating this to the observed mean recall of 1.14, we obtain the maximum-likelihood estimate, $\hat{\theta} = 0.38$. From this value we then calculated the predicted values in Table 2. Table 2 compares the pooled observed frequencies to those predicted by the strict binomial model and by the schema model. The fit of the strict binomial model is obviously inferior. It has a chi-square for goodness of fit of 596.0, which is a significantly poor fit, $p < .001$.

Another model that can be briefly mentioned is an all-or-none model which assumes that the subject stores either all items of a quartet or none of them, and that the mean recall is an average of complete recalls with complete failures. Such a model is readily rejected by noting the high likelihood of partial recalls (i.e., $\underline{X} = 1$ or 2, $\underline{Y} = 1$) in the data of Table 2.

Insert Table 2 about here

Finally, we consider a model, called the "memory fragment" theory, proposed by Jones (1977). According to this model, the outcome of attempting to store all four words will be some groups ("fragments") of inter-associated words. With quartets, the subgroups can be of size 2, or 3, or 4. The possible logical forms of the groupings are depicted in Figure 2, and their probabilities are to be estimated empirically. There is one fragment of size 4, and we let $\underline{c}$ denote its probability. There are four distinct fragments of size 3, and we let $\underline{d}$

Insert Figure 2 about here
----

denote the sum of these probabilities. There are six fragments of size 2, where the other two elements are not grouped together, and we let $\underline{e}$ denote the summed probability of these. There are also six fragments of size 2 but where the other two elements also are grouped together in memory; we let $\underline{f}$ denote the summed probabilities of these. Finally, we let $\underline{g} = 1-\underline{c}-\underline{d}-\underline{e}-\underline{f}$ denote the likelihood of all remaining states of the memory trace. It is assumed that, at the time of testing, a subject's memory for a given quartet will be in some one of these states. The response rule is that all items that are in the same memory fragment as a cue word will be recalled. On the assumption of homogeneous selection of the words as members of fragments and as cues, the derived probabilities for this model for one-cue tests are:

$$\Pr\{\underline{X} = 3\} = \underline{c}$$

$$\Pr\{\underline{X} = 2\} = .75\ \underline{d}$$

$$\Pr\{\underline{X} = 1\} = .50\ \underline{e} + \underline{f}$$

$$\Pr\{\underline{X} = 0\} = .50\ \underline{e} + .25\ \underline{d} + \underline{g}\ .$$

For two-cue tests the expressions are:

$$\Pr\{\underline{Y} = 2\} = \underline{c} + .50\ \underline{d} + .67\ \underline{f}$$

$$\Pr\{\underline{Y} = 1\} = .50\ \underline{d} + .67\ \underline{e}$$

$$\Pr\{\underline{Y} = 0\} = .33\ \underline{e} + .33\ \underline{f} + \underline{g}\ .$$

There are four parameters to be estimated from the five degrees of freedom in the data. A modified maximum likelihood method was used to fir the model to the frequency data of each subject. An exceedingly good fit was obtained and only one of the 16 subjects deviated significantly ($\underline{p} < .05$) from his predicted distribution. Summing across subjects, the chi-square statistic for goodness of fit was 15.73 on 16 degrees of freedom, which has $\underline{p} = .47$. As before, there were significant variations in parameters across subjects. Table 2 reports the predicted frequencies obtained by fitting the model to the pooled data. The parameter estimates are $\hat{c} = .20$, $\hat{d} = .33$, $\hat{e} = .27$, $\hat{f} = -.02$, and $\hat{g} = .22$. The fit to the overall distribution is marginally significant, with $\underline{X}^2(1) = 3.85$, $\underline{p} = .05$. The overall fit is somewhat better for the schema model. The fact that the fragment model fits the individual data much better than the pooled data in Table 2 simply indicates that there are rather large individual differences in parameter values of the model. Although the fragment model and schema model fit the individual data about equally well, the schema theory is somewhat preferable because it has only half the number of arbitrary parameters.

The outcome of our several theoretical comparisons is a qualified success for the schema model which nicely predicted the results. Other research in our laboratory by Arnold (Note 1) has also supported the schema model. Arnold applied the model to learning of unrelated word pairs and unrelated word triplets, gathering data from single-cue recall, double-cue recall, pair recognition (old vs. new pairs) and triplet recognition (old vs. several types of new triplets). In each experiment, the data from the various tests was well predicted quantitatively by the schema model. Thus, there is substantial evidence in its favor, at least for these study-test types of association experiments.

The schema model is interesting in that it is one mathematical representation of the propositional learning hypotheses proposed by Anderson and Bower (1973), whereby arbitrary groups of words to be learned are linked into a proposition stating some conceptual relation among the individual words. Similarly, the model is related to the current theories about schema (e.g., Anderson, 1977; Rumelhart and Ortony, 1977) and scripts or frames (Schank and Abelson, 1977). A schema is a large conceptual structure pre-existing in memory; it inter-relates a number of actors, objects, and properties; it gives general information about these categories along with the type of values which fill those variable slots; and it can be applied to new cases by instantiating the variables in terms of the value of particular objects or events in the case at hand. Of relevance to the present data are schema for activities (camping, cutting hair), or situations (chemistry lab, dentist's office). Thus, when the person reads _test tube_, _litmus paper_, _Bunsen burner_, and _acid_, by some intersection search in memory (cf. Anderson, 1976) the relevant "chemistry-lab" schema is activated. These items are already interassociated props or objects within the script, commonly appearing in high school and college chemistry lab classes. Learning consists, first, of strengthening each word-to-schema linkage, and, second, of tagging specifically these items of the schema (rather than _bench_, _pipette_, etc.) as having occurred in the list studied. These two functions of learning correspond to the access parameter and the response parameter in the mathematical model. If schema access succeeds but the subject fails to discriminate accurately between the mentioned vs. unmentioned props of the script, then he is likely to intrude a script-related word into recall. We saw earlier that such thematic intrusions were likely, particularly for the Informed subjects for whom schema-activation during learning was guaranteed.

All in all, then, there seems to be a close correspondence between the underlying schema theory for memory and the mathematical model which represents part of it and which fits the data exceedingly well.

References

Anderson, J. R. Language, memory and thought. Hillsdale, N. J.: Lawrence Erlbaum Assoc., 1976.

Anderson, J. R. & Bower, G. H. Human associative memory. Washington, D. C.: Winston, 1973.

Anderson, R. B. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, and W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N. J.: Lawrence Erlbaum Assoc., 1977.

Foss, D. J. & Harwood, D. A. Memory for sentences: Implications for human associative memory. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 1-16.

Jones, G. V. A fragmentation hypothesis of memory. Journal of Experimental Psychology: General. 1976, 105, 277-293.

Rumelhart, D. E., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N. J.: Lawrence Erlbaum Assoc., 1976.

Schank, R., & Abelson, R. Scripts, plans, goals, and understanding. Hillsdale, N. J.: Lawrence Erlbaum Assoc., 1977.

Reference Note

Arnold, P. On the nature of associations. Unpublished Ph.D. dissertation. Stanford University, Stanford, Calif. 1976.

Footnote

1. This research was supported by a research grant, MH-13905, from the National Institutes of Mental Health. Thanks are due to Dr. Michel Aickin of the Dept. of Mathematics of Arizona State University for helping fit the mathematical models to the individual data. Requests for reprints should be sent to Dr. Gordon H. Bower, Dept. of Psychology, Stanford University, Stanford, Calif. 94305

Table 1. Observed and predicted frequencies of responses recalled to one or two cues for Informed and Control subjects.

Cue	No. Recalled	Informed		Controls	
		Obsd.	Pred.	Obsd.	Pred.
Single Cue	3	70	65	49	49
	2	79	83	71	73
	1	37	35	36	36
	0	134	137	164	162
Double Cue	2	128	131	109	109
	1	112	111	111	108
	0	80	78	100	103

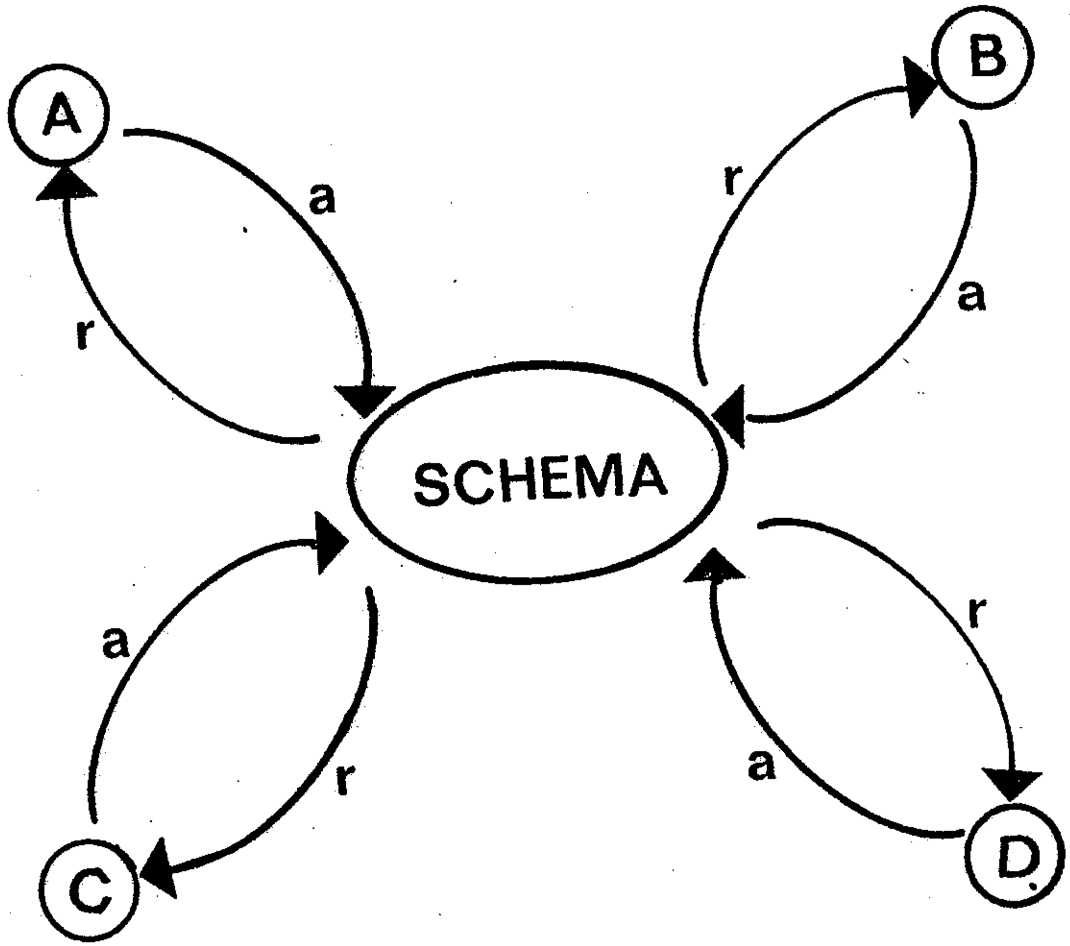
Table 2. Observed and predicted pooled frequencies of recall patterns.

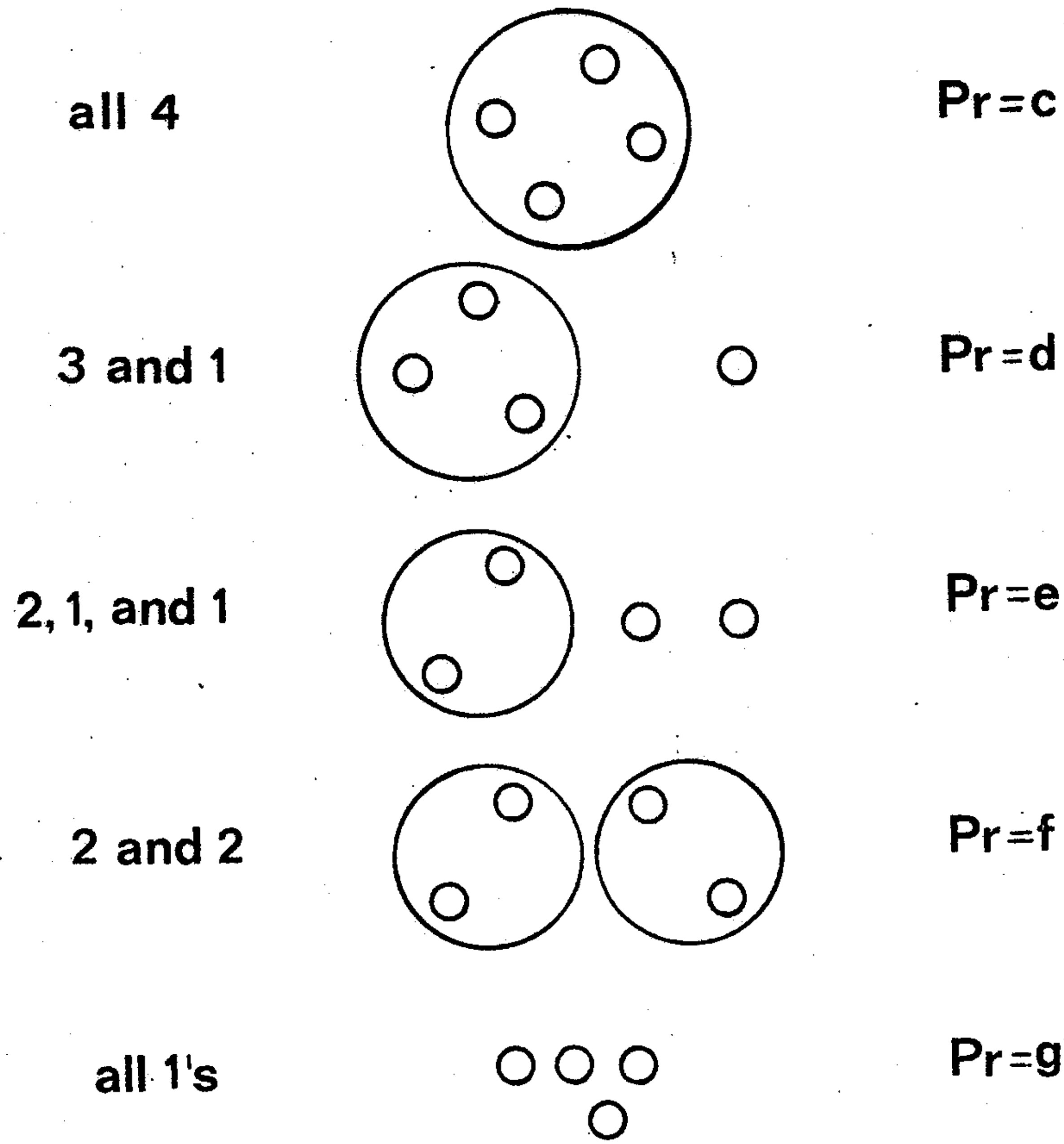
Testing Condition	No. Recall	Observed Frequencies	Schema Model	Binomial Model	Fragment Model
1-Cue	3	119	113	35	127
	2	150	159	172	160
	1	73	75	280	74
	0	298	293	153	279
2 -Cues	2	237	238	243	224
	1	223	224	303	224
	0	180	178	94	192

Figure Captions

Figure 1. Representation of the schema memory structure for a given word quartet.

Figure 2. The four logical groupings of four words for the fragment model. The probabilities of the types of groupings are written to the right.


A
B
C
D
SCHEMA
a
r
a
r
a
r
a
r


all 4
Pr=c
3 and 1
Pr=d
2, 1, and 1
Pr=e
2 and 2
Pr=f
all 1's
Pr=g