

A METHODOLOGICAL STUDY ON THE EVALUATION OF LEARNING FROM STORY NARRATIVES

GAIL GLINER, SUSAN R. GOLDMAN, and LAWRENCE J. HUBERT
University of California at Santa Barbara

ABSTRACT

Exploratory multidimensional scaling and confirmatory nonparametric procedures (Hubert and Levin, 1976) were used to represent data from similarity rating and sorting tasks performed on nine animal names. Confirmatory procedures demonstrated that the organization of the data from the two tasks was similar. Analyses of data from sorting tasks performed after reading two stories with the nine animals as main characters (Bisanz, LaPorte, Vesonder, and Voss, 1978) suggested a change from pre-reading organization that was similar to the organization of the characters intended by the authors in one of the two stories. One of the two dimensions used to write the second story appeared not to be salient to the readers.

Recent work on the acquisition of knowledge emphasizes that learning involves an interaction between the learner and the material to be learned. This perspective suggests the importance of a number of factors in the learning process. For example, Bransford (1979) has specified four such factors: the nature of the materials to be learned, the characteristics of the learner, the nature of the learning activities, and the nature of the criterial tasks used to evaluate learning. The nature of the materials deals with characteristics of information presentation. Characteristics of the learner refers to the unique set of information a particular learner possesses prior to the learning situation. The learning activities factor refers to the different activities in which one might engage in order to learn, e.g., rehearsal and elaborative processes. Finally, criterial tasks refer to how learning is measured, including what aspects of learning are tested and how these are assessed.

The present research focuses primarily on methodological issues associated with characterizing the nature of the materials and the nature of the learner's understanding of information in the materials prior to instruction and after exposure to the mater-

Experiment 1 was conducted as part of the first author's doctoral dissertation. The first author acknowledges the contribution to this work of the other members of her committee: Drs. John Cotton, Richard Mayer and Melvyn I. Semmel.

ials. The framework used to characterize both the materials and the learner assumes that both information in text, one type of materials, and the learner's understanding of that information may be described as a pattern of interrelated ideas. This organization can be described in terms of concept nodes, relations, and linkages between the two (Norman and Rumelhart, 1975).

Previous research has examined such patterns of interrelated ideas in a variety of ways and for a variety of purposes. For example, one way in which semantic memory organization has been examined is in terms of patterns among lexical items (Anglin, 1970; Deese, 1962; Fillenbaum and Rapoport, 1971; Miller, 1969) or among concepts from a particular domain, such as animal terms (Henley, 1969; Rips, Shoben, and Smith, 1973). Other researchers have focused on characterizing patterns of ideas that a learner acquires in an instructional context (e.g. Shavelson, 1972). This work typically examines patterns acquired by the learner in more of an episodic than semantic (Tulving, 1972) context. That is, the pattern of concepts that the learner demonstrates after exposure to instruction is compared to the pattern in the instructional material, and the nature of semantic memory organization for this information prior to instruction is ignored. Empirical investigations of this type have been primarily in the field of physics teaching (P. Johnson, 1967; Preece, 1977; Shavelson, 1972). Finally, a third orientation, and the one that is the focus of this research, assesses the learner's representation of concepts described in a text, where that description differs from that in semantic memory (e.g. Bisanz, LaPorte, Vesonder, and Voss, 1978; LaPorte and Voss, 1979). Note that the pattern learned from the text is not assumed necessarily to affect patterns in semantic memory. This orientation focuses on the degree to which a new pattern of relations among concepts is acquired from the text.

In previous attempts to measure information patterns, or organization, the statistical methodology has been primarily exploratory. This has been true across a variety of criterial tasks, e.g., free recall tasks, similarity ratings, sorting tasks. These tasks all yield estimates of information organization based on some type of distance measure between all pairs of ideas. The present research replicates exploratory procedures, in particular, multidimensional scaling, but emphasizes newer confirmatory techniques in which hypotheses concerning particular information patterns can be tested (Hubert and Levin, 1976). This confirmatory orien-

tation provides a framework for developing a set of significance tests that can be applied to issues involving the correspondence between patterns the learner demonstrates and the pattern of concepts in the learning material. Thus, confirmatory procedures allow the statistical evaluation of the relationship between patterns acquired by the learner and those hypothesized to be in the material.

The present research specifically focuses on patterns acquired from narrative material and methodological issues involved in measuring organization that is acquired after reading a story. It is a replication and extension of the work of Bisanz et al. (1978). Their view of story comprehension does not involve analyzing the passage content, but considers how closely the reader organizes the story characters to represent the relationship among the characters in the text. Text organization is based on the pattern of relationships among the characters in terms of the degree to which they share certain salient personality attributes. Bisanz et al. (1978) wrote two stories based on the same nine animal characters. In writing the stories, they first assigned each animal a position along two dimensions, from helpful to not helpful and from leader to follower in one story, and old-young, cautious-not cautious, in the other story. They then wrote the stories based on these dimensions. If readers use these attributes in organizing the characters, then the reader's organization should be similar to the text organization of the characters. By using animal terms as the main characters, a story could be written that has different relationships among the animal terms from those in semantic memory (Henley, 1969; Rips et al., 1973). In summary, this approach focuses on three of the factors Bransford (1979) has emphasized in the learning process: knowledge of the learner before instruction (organization of the animals in semantic memory), organization of the learning material (text organization), and criterial tasks to assess learning (e.g., similarity rating tasks).

In the work of Bisanz et al. (1978), subjects first rated the similarity of pairs of nine animals before reading two stories. A multidimensional scaling solution for the similarity ratings was consistent with semantic memory representations (Henley, 1969; Rips et al., 1973). After reading each story, subjects rated the pairs of animals according to their similarity as described in the story. The scaling solutions reflected a different set of relationships for the characters in each, and both organizations appeared

to be different from the organization of the animals in semantic memory.

EXPERIMENT 1

The first study extended the work of Bisanz et al. (1978). While Bisanz et al. used only a similarity rating task, we used a similarity rating task and a sorting task before reading so that we could compare data from two tasks that measure organization. We also compared results from exploratory multidimensional scaling and newer confirmatory statistical techniques. This confirmatory orientation provided a way to address two primary questions in the first study: whether the readers appeared to change the organization of the animal names after reading each story and whether this change reflected the intended organization of the characters in each story. In addition, the first study attempted to identify and confirm the dimensions that readers were using to organize the animals before and after reading.

Method

Subjects. Subjects were forty undergraduate psychology students, 31 females and 9 males, who participated in the experiment as part of a class requirement.

Materials. The two stories were from Bisanz et al. (1978). The characters in each story were nine animals. One of the stories was based on the animals' deciding whether to help a lion out of a hunter's trap. The lion had been a good leader, but a tiger now wanted to be leader, but he was trying to convince the other animals not to help the lion and to follow him. The animals were described in terms of the attributes of helpfulness and leadership (HL story). The other story involved the animals' advising a fox whether or not to eat a piece of meat he had seen on the road. Some of the animals were cautious and felt it was a trap, and others advised him to eat the meat. The salient attributes were age and cautiousness (AC story). If the readers used these attributes in organizing the characters, then the readers' organization should be similar to the text organization of the characters.

Design. There were three between-subjects factors and two within-subjects factors. Of the three between-subjects factors, two were introduced by counterbalancing procedures and the third,

the factor of interest, was presence versus absence of the passage during the post-reading sorting tasks. A second between-subjects factor was the order in which the reading tasks were performed: for half of the subjects the similarity rating task was followed by the sorting task and this order was reversed for the other half. The third between-subjects factor was story presentation order, with half of the subjects reading the AC story first and half the HL story. Thus, eight groups of subjects resulted from the factorial combination of passage present versus absent, task order before reading, and story order.

The two within-subjects factors were type of pre-reading task and story. All subjects performed a similarity rating task and a sorting task prior to reading the passage. Subjects then read the appropriate first story and performed three sorting tasks. The first sort was based on "similarity in the story" and the other two were based on each of the two dimensions in the story. Subjects then read the second story and did three sorts for that story. Subjects in the passage present groups conducted these post-reading sorts with the text available to them while for the passage absent groups, the text was removed prior to the sorting tasks. The order in which the two sorting dimensions for each story were presented was controlled so that each dimension was the first sort dimension an equal number of times across subjects within the passage present and within the passage absent groups. In addition, after subjects were satisfied with their sorting groups, they assigned a rating of "high," "medium," or "low" to each group on each dimension.

Procedure. All groups first performed the similarity rating task and the sorting task on the nine animal names used in the stories. In the similarity rating task, subjects assigned each animal pair a number between 1 and 7, where 1 meant the animals were very similar and 7 meant they were very different. In the sorting task, they put animals into as many groups as they liked, based on similarity. Subjects then read the first story. After reading, they sorted the animals based on similarity in the story. They then sorted based on each of the dimensions in the first story, helpfulness and leadership in one story and age and cautiousness in the other. After sorting the animals based on each individual story dimension, each subject ordered his/her groups into "high," "medium," and "low" on the particular dimension. The data from this task were used to aid in interpreting the dimensions subjects used when sorting based on similarity in the story.

Subjects were then told to try to "wipe the story and tasks they had done from their mind." They then read the second story, sorted the animals based on similarity in the story, sorted based on each of the individual story dimensions, and ordered the groups in terms of "high," "medium," and "low" on each dimension. Subjects were run in eight groups of five subjects each.

RESULTS AND DISCUSSION

Multidimensional scaling analysis of prereading organization.

As a prime illustration of an exploratory strategy, multidimensional scaling procedures can be used to elicit the structure of data from the similarity rating task and the sorting task to get a picture of the animals as they may be perceived by the subjects. Results from the spatial representation can then be compared with multidimensional scaling solutions for the same animal terms in Bisanz et al. (1978), Henley (1969), and Rips et al. (1973).

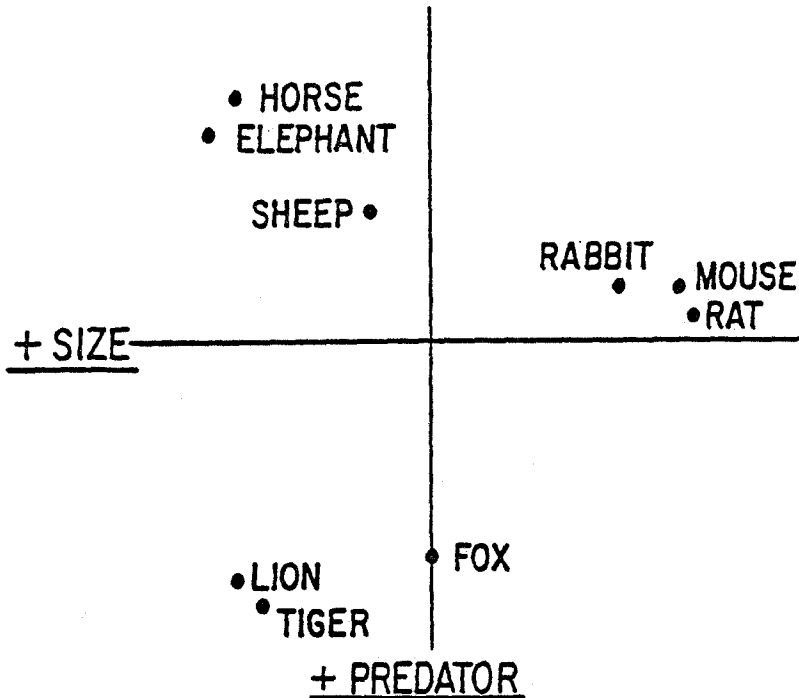


Figure 1: INDSCAL solution for similarity rating task.

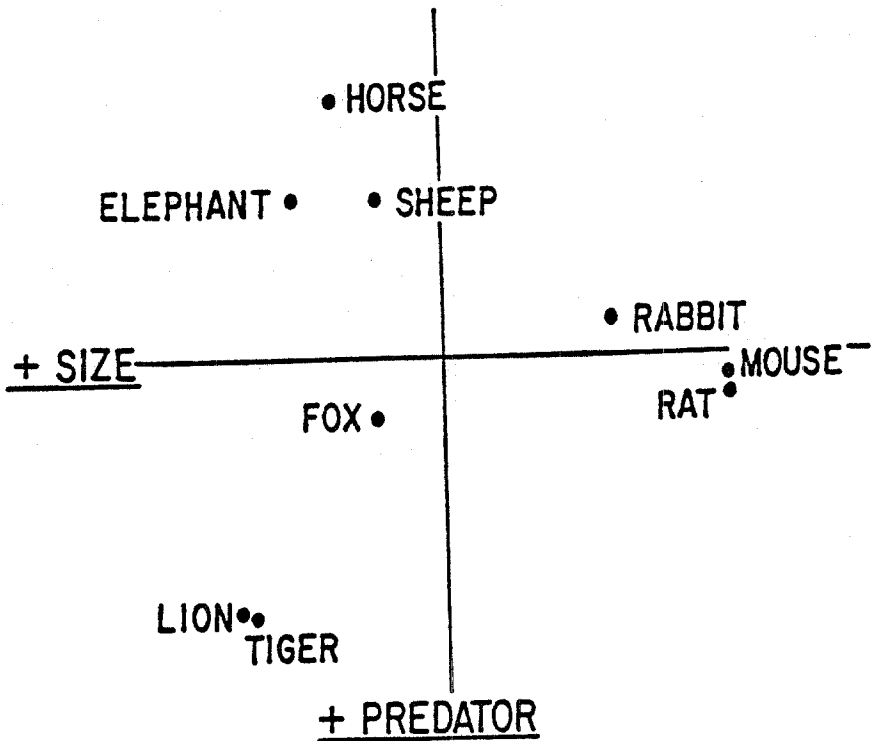


Figure 2: INDSCAL solution for sorting task.

Data from each subject for the similarity rating task were put in matrix form, called a "proximity" matrix, where each entry was a number from 1 to 7 representing the similarity rating for the animal pair. The proximity matrix for each subject for the sorting task consisted of ones and zeros, where one meant the animal pair was grouped together and zero otherwise. Multidimensional scaling involves generating the "best fitting" spatial representation from distances in a proximity matrix. In particular, individual differences multidimensional scaling allows for a representation of both an overall group space of the stimuli and a space for the weights given by each subject for each dimension of the group space. Based on the INDSCAL version of individual differences scaling (Carroll and Chang, 1970), a group space and a weight space for the pre-reading structure of the nine animals were generated using both the similarity rating and sorting matrices for each subject. Because of the previous work by Henley (1969) and Rips et al. (1973) and the small number of stimuli, the group

spaces for each analysis given in Figures 1 and 2 are limited to two dimensions.

The INDSCAL solutions for the similarity rating task and the sorting task appear similar to each other and to previous work of Bisanz et al. (1978), Henley (1969), and Rips et al. (1973). The dimensions in both solutions apparently represent size (dimension 1, horizontal dimension) and predacity (dimension 2, vertical dimension), as interpreted by Rips et al. The variance accounted for by each dimension (the goodness-of-fit measure generated by INDSCAL) in each of the two solutions is as follows: 1) sorting task: dimension 1 = 34.8%, dimension 2 = 22.3%, and 2) similarity rating task: dimension 1 = 51%, dimension 2 = 15.3%. Comparison of similarity rating task and sorting task using confirmatory methods.

While multidimensional scaling showed that data from the similarity rating task and sorting task appear similar to each other and to previous work, another orientation involves confirming the similarity in a hypothesis testing framework. To do this, the structures of the total proximity matrix obtained from the similarity rating task and the total matrix for the sorting task were compared using the quadratic assignment (QA) paradigm (Hubert and Levin, 1976) to confirm whether the patterns of entries in both matrices were similar. Based on the matrix of similarity ratings of each subject, the total group similarity rating matrix was formed by adding the ratings for each subject. Since entries in each individual matrix were numbers between 1 and 7, the entries in the group matrix ranged between 40 and 280, where numbers close to 40 would suggest the group, as a whole, saw the animal pair as very similar and larger numbers meant the group viewed the animals in the pair as very different. The entries for the group matrix in the sorting task represented how many subjects sorted the animals into the same group. The group matrices for the similarity rating task and sorting task are given in Figure 3.

The comparison of the two data matrices using the QA paradigm essentially involves testing for the significance of the correlation between the two matrices, in which the null hypothesis is that there is no similarity between the patterns of entries in both matrices. In the confirmatory mode, the group proximity matrix from the sorting task is compared to the group matrix from the similarity rating task to assess whether the pattern represented in one matrix is also present in the other matrix. A statistic, T ,

(a)

MOUSE	TIGER	RABBIT	HORSE	RAT	ELEPHANT	FOX	SHEEP	
258	49	240	198	261	198	168	209	LION
	261	135	257	53	262	218	224	MOUSE
		239	201	258	203	168	215	TIGER
			241	139	250	189	187	RABBIT
				259	163	211	157	HORSE
					266	207	233	RAT
						246	228	ELEPHANT
							203	FOX

(b)

MOUSE	TIGER	RABBIT	HORSE	RAT	ELEPHANT	FOX	SHEEP	
0	38	1	1	3	3	9	5	LION
	0	20	4	34	3	2	2	MOUSE
		0	0	4	3	8	4	TIGER
			5	16	4	7	11	RABBIT
				0	16	3	12	HORSE
					0	2	4	RAT
						3	2	ELEPHANT
							6	FOX

Figure 3: Totals (n=40) for similarity rating tasks (a) and for sorting tasks (b).

representing the sum of the products of the corresponding elements between the two matrices is computed. If the probability of the particular value occurring given the distribution of Γ based on all equally likely permutations of rows and corresponding columns of one of the matrices is sufficiently small, then the structure in the second matrix is considered mirrored in the first. For the matrices used in this study, the permutation distributions are large ($9!$ for the 9×9 matrices of animal comparisons). Therefore, the expectation and variance of Γ was found and a z statistic was obtained. The conservative Cantelli's inequality was used to assert significance for a p value of at most $1/(z^2 + 1)$ for a one-tailed test. Cantelli's inequality is sufficient in this case, but if more exact p values are needed, Monte Carlo significance tests can be used (Edgington, 1969). It should be noted that Γ is merely an "unnormalized" Pearson Product moment correlation coefficient, and therefore, comparing the index value for Γ to a sampling distribution is equivalent to performing the same procedure on the correlation coefficient for the entries in the two matrices. The only quantity in the Pearson correlation coefficient that varies, depending upon the order of the rows and columns, is the index Γ .

Using the QA procedure, the correlation between the similarity rating task matrix and the sorting task matrix was $-.91$ with a corresponding z value of -4.98 , which is significant ($p < .05$) using Cantelli's inequality. The correlation was negative since large values in the similarity rating meant the animal pair was perceived as dissimilar, and large values for the sorting task meant the pair was similar. Thus, while exploratory multidimensional scaling has suggested that data from the similarity rating task and sorting task appear similar, the use of the QA paradigm in the confirmatory mode has provided a statistical test of the significance of this apparent similarity in organization.

Evaluating organization after reading

Besides comparing the similarity rating task and the sorting task as measures of organization before reading, a primary purpose of this work was to measure the change in organization of the animal names after reading and to compare the readers' organization to that intended by the authors. Since Bisanz et al. (1978) had used a similarity rating task, we could compare results of post-reading analysis using a sorting task. Besides using multidimensional scaling, confirmatory procedures were used to test hypoth-

eses concerning the similarity of the patterns of the story characters perceived by the readers to the logical story structure of each story.

Passage present vs. passage absent condition. Bisanz et al. (1978) found that similarity ratings done by subjects with the story present differed from ratings done without such access, at least in terms of variance accounted for in the INDSCAL solutions. Since half the subjects had access to the story and half did not, this conjecture could be evaluated for the sorting task.

The INDSCAL solutions for the passage present and passage absent conditions in the present study show almost identical measures of variance accounted for, as opposed to the results of Bisanz et al. of less variance accounted for in the passage absent condition. Thus, the free sort task may be measuring something closer to overall acquired organization of the characters rather than ability to find the correct responses in the story passage when it is present.

Besides looking at the variance accounted for using INDSCAL for the free sort data in the present work, homogeneity within the passage present group was investigated using a variation of the QA paradigm. A 40x40 correlation matrix was formed for each story, where each entry, r_{ij} , represented the correlation between the free sort matrices for subject i and subject j . The first twenty subjects represented the passage present condition, and the second twenty represented the passage absent condition. To test the homogeneity of the group of correlations for the subjects under the passage present condition, an index representing only the sum of the within-group proximities for this group was used. This process generated a z value of $-.26$ (for the HL story) and a z value of $.49$ (for the AC story), suggesting that there may be no greater homogeneity among the passage present subjects. Again, contrary to Bisanz's results with similarity ratings, results from both multi-dimensional scaling and the QA paradigm suggest that the free sort task did not produce more homogeneous responses if subjects had access to the text when they performed the task.

Evaluating a change in organization. A salient question in the present research was whether a change between pre-reading and post-reading organization of the animal names, based on subjects' sorting of the animals into groups, could be detected. If there is a change in organization, then the structure of the group matrices for the sorting task based on each of the stories should not

be similar to the structure of the sorting matrix done before subjects read either story. The group matrix from each task was compared to the group matrix for the sorting task done before reading the two stories. The correlation between the pre-reading group task matrix and the sorting for the HL story was $-.27$ ($z = -1.46$); for the AC story the correlation with the pre-reading sorting matrix was $-.15$ ($z = -.82$). Thus, there is no evidence that the matrices for the sorting of the animals after the story are similar to the pre-reading structure. A similar comparison was done to assess any similarity between the AC ratings and the HL ratings. The correlation between these group matrices for each story was $-.04$ ($z = -.24$), suggesting that the structure of the group task matrices for each story was not the same. Again, the use of the QA paradigm has answered an important question in the study. We have shown, using a hypothesis testing procedure, that the readers appear to have changed the organization of the animal characters after reading each story.

In Figures 4 and 5 it can be noted that the group spaces from

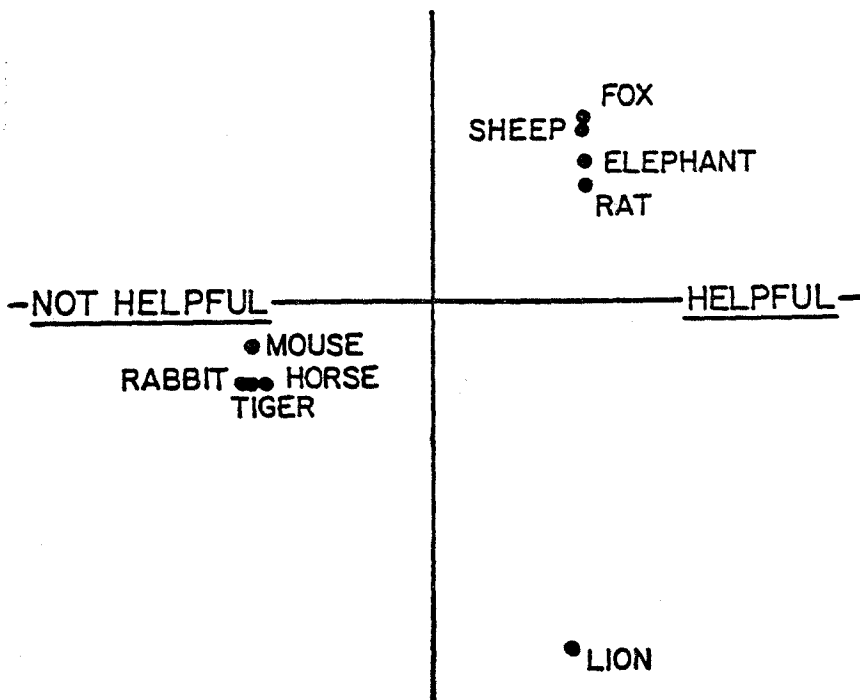


Figure 4: INDSCAL solution for sorting task for Helpful-Leader story.

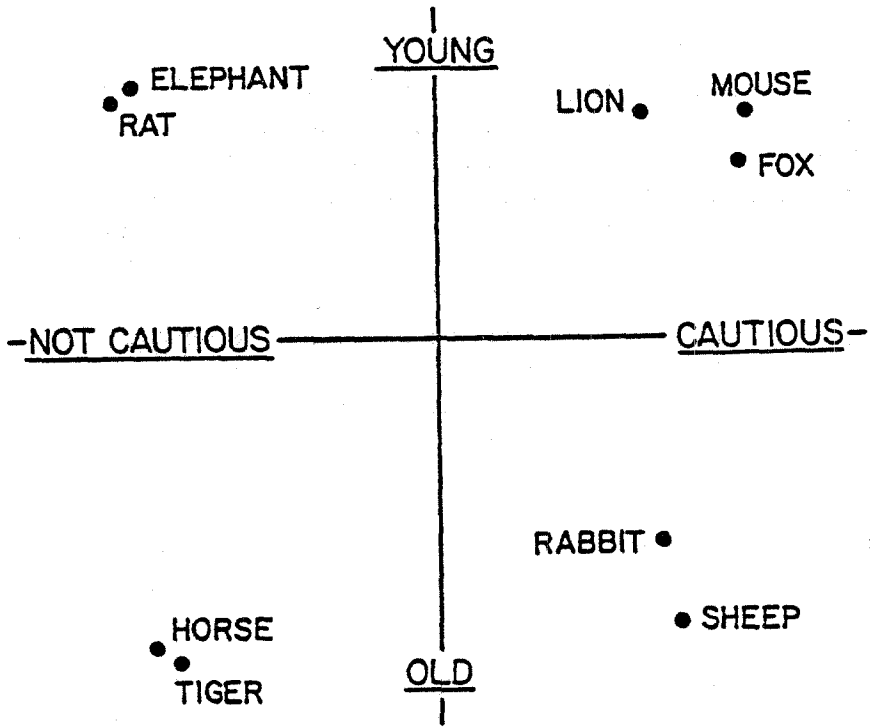


Figure 5: INDSCAL solution for sorting task for Age-Caution story.

INDSCAL are different from each other and from the pre-reading group space for the animals in the sorting task (Figure 2). The group space for the HL story is similar to the INDSCAL solution from Bisanz et al. (1978), except that the mouse in their solution was at the opposite end of the vertical dimension. The AC story is similar to Bisanz et al. along the first dimension, although they could not as easily interpret the second dimension as "age," which can be done in the present study.

In the present study, the variance accounted for in the INDSCAL solution for the HL story was 62% for the first dimension and 11% for the second dimension. For the AC story, the variance accounted for was 39% and 18% for the first and second dimension, respectively. Thus, since the group spaces for the before-reading and after-reading sorting "look" different, and the QA analysis confirmed that the group task matrices are not similar in structure, there is corroborating evidence that the readers have changed their organization of the animals, as measured by the sorting task.

*Comparison of the structure of the sorting data to a
preposited logical structure*

We wished to compare the subjects' sorting of the characters for each story against the logical story structure to assess whether they had "acquired" the intended structure. Therefore, an analysis of the structure of each story was carried out. Since Bisanz et al. (1978) had first decided on a logical structure for each story and then wrote each story around these dimensions, the current analysis was carried out to determine if the logical structure could be recovered. This was done by an analysis of what information in each story could contribute to rating the animals along each dimension used to write the story. A count was made of the number of times each animal was described in each story in terms of the particular dimensions of the story characters. For example, in the AC story, three references were made to the sheep as "old" and one of these references was as "very old." The rabbit was described once as "old" and once as "older"; therefore the rabbit would be lower on the "age" scale than the sheep.

The logical structure generated in this way for both stories is represented in Figure 6. The logical structure for the AC story was similar to Bisanz et al. (1978) except they had meant the rabbit and tiger to be neutral on the age scale. References to the rabbit as "older" and the tiger as "in his prime" would seem to indicate that they belong a little "higher" on the age dimensions.

These logical structures were then compared to logical structures derived from the sorting task data to determine how the subjects viewed the animals along these dimensions. A proximity matrix for each logical story structure was derived and compared with the subjects' proximity matrix for each story by obtaining Euclidean distances between the animals from the logical structure. A proximity matrix from Bisanz's structure was also formed in the same way. Based on the QA paradigm, each logical structure was compared to the group sorting matrix. A comparison of the group sorting matrix to the logical structure from Bisanz et al. yielded a correlation of $-.87$ with a z value of -5.07 . The correlation between the group task matrix and the logical structure generated from references in the AC story was slightly higher: $r = -.90$, $z = -5.00$. In short, there was evidence that the data matrix generated from the sorting task for the AC story was similar to the logical story structure posited by Bisanz et al. and to the logical structure that resulted from the current analysis.

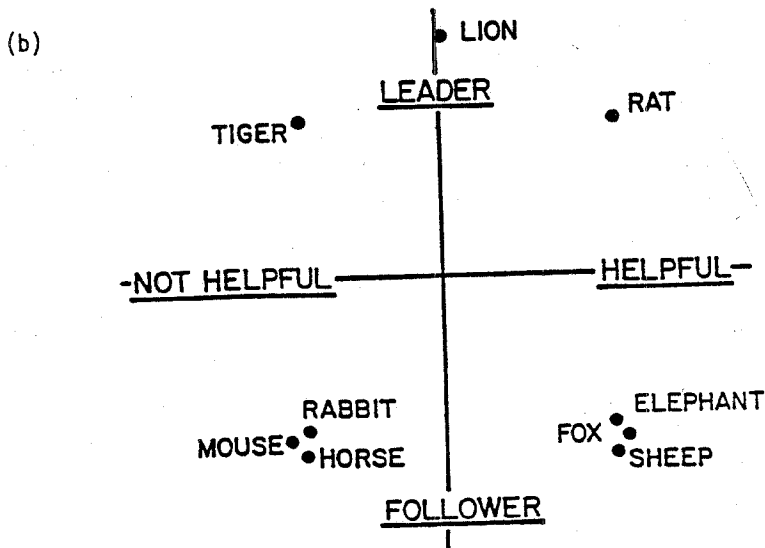
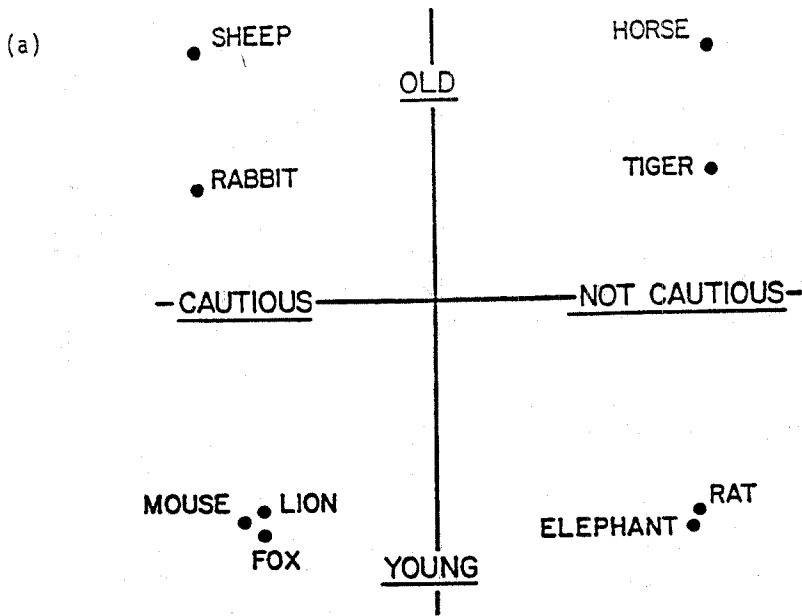


Figure 6: Logical structure for Age-Caution story (a) and the Helpful-Leader story (b).

The correlation between the data matrix for the HL story and the structure matrix from the logical structure was somewhat disappointing: $r = -.61$ ($z = -3.49$). Although this correlation is marginally significant ($p < .07$) using the very conservative Cantelli's inequality, its square suggests that only 36% of the variance in the data from the HL story is accounted for by a linear relationship to the logical structure. In summary, there is evidence that readers organized the characters in the AC story as intended by the authors, but this may not be as true for the HL story.

To gain further information about whether readers "acquired" the organization intended in each story, a variation of INDSCAL was used in a confirmatory mode to compare how closely the individual ratings fit a preposited logical structure. The "zero iterations option" from the INDSCAL program (Kruskal and Wish, 1978) allows a generation of the subject weight space directly from an a priori group structure instead of iteratively computing the "best fitting" group space from the individual data. Since the a priori logical structure was represented by the coordinates of the animal characters in the logical structure for each story, the measure of accounted-for-variance from the INDSCAL program for each dimension should be high if the subjects have used the dimensions from the logical story structure. The weight space in each solution using the zero iterations option gives a picture of how well each subject's ratings fit the group space dimensions, which in this case is the logical story structure. These weight spaces are given in Figure 7. Each point represents a single individual; the further a point is away from the origin, the better the explanation of that particular subject's data by the *two* given dimensions. Also, the closer an individual's point is to a given axis, the more the individual is weighting that dimension.

The weight space for the HL story using the zero iterations option shows the subjects rating much more on the first dimension (Figure 7a). Note that this was also evident in the weight space in which the zero iterations option was not used. The variance accounted for on the first dimension, using the zero iterations option for the HL story, was 60.8% but on the second dimension was only 3.2%. For the AC story, the variance accounted for on the first dimension was 37.8% and 15.9% on the second dimension (Figure 7b). Using the logical structure from Bisanz et al. for the AC story, the variance accounted for was 37.8% and 10.0% for

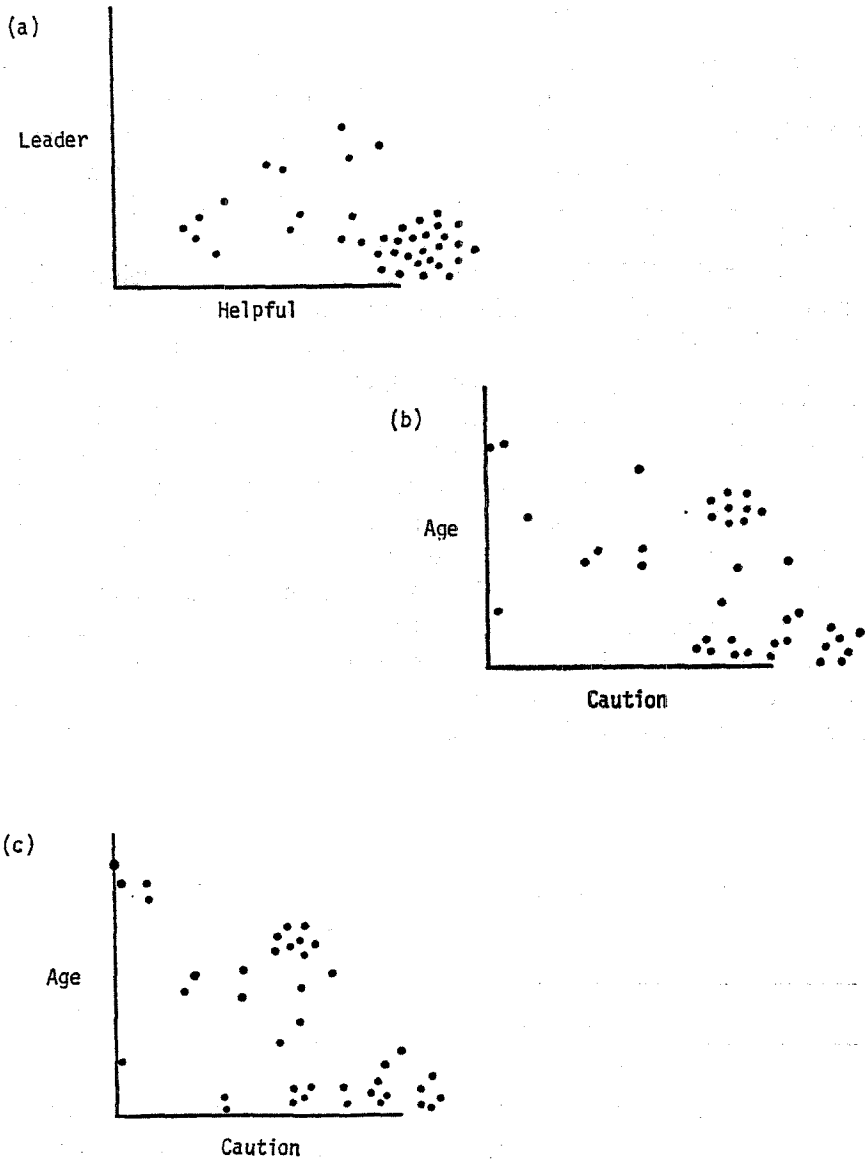


Figure 7: Subject weight spaces using zero iterations option for Helpful-Leader logical structure (a); for Age-Caution logical structure (b); and for Age-Caution logical structure from Bisanz et al. (1978) (c).

the first and second dimensions, respectively (Figure 7c). Thus, the second dimension from the HL story appears not as salient to the readers.

Interpretation of dimensions

The next question involved interpreting the dimensions subjects used in sorting the animals and whether these dimensions were the same as those used to write the stories. This is a crucial question, since one premise of Bisanz et al. (1978) was that thematic organization could be interpreted in terms of the relationship of attributes of the main characters.

Although INDSICAL provided a picture of the relationship among the animals in multidimensional space, the interpretation of the dimensions from the group space until this point has been largely subjective. However, more rigorous methods are available. Specifically, multiple linear regression can provide verification for a conjectured interpretation. As a final task in Experiment 1, subjects were asked to group the animals into three groups for each dimension. They were to sort the animals into high, medium, and low groups on age and cautiousness for the AC story and helpfulness and leadership for the HL story. Using these groupings, a mean rating, given in Table 1, was found for each animal on each dimension, where high, medium, and low ratings were coded as 3, 2, and 1, respectively. Using each of these mean ratings as the dependent variable and the coordinates of the animals from the

Table 1
Mean values for ratings of animals and coordinates of
each animal in INDSICAL group space

	Age	Caution	Helpful	Leader	AC story		HL story	
					X1	X2	X1	X2
lion	1.43	2.23	2.35	2.78	.22	.30	.30	-.69
mouse	1.40	2.45	1.30	1.05	.34	.29	-.37	-.08
tiger	2.08	1.33	1.20	2.95	.33	.45	-.36	-.17
rabbit	2.28	2.45	1.15	1.65	.29	.23	-.38	-.18
horse	2.85	1.13	1.08	1.43	.37	.44	-.38	-.18
rat	1.20	1.43	3.00	2.98	.40	.30	-.30	.23
elephant	1.18	1.20	2.30	1.50	.39	.32	.30	.31
fox	1.20	2.78	2.43	1.93	.34	.22	.30	.38
sheep	2.93	2.90	2.73	2.38	.30	.34	.30	.37

group space from each INDSCAL solution (Figures 5 and 6) as the independent variable, four multiple regression analyses were performed. If subjects rated the animals consistently along a continuum on each dimension, then these ratings should also be reflected in the orderings of the coordinates in the group space. The consistency is reflected in the angle that the resulting regression line makes with the coordinate axis assumed to represent the particular dimension and the size of the variance accounted for in the regression framework.

The result of the regression analyses are presented in Table 2. A large direction cosine (close to 1.0) for the regression line in relation to a coordinate axis implies that the line makes a small angle with the coordinate axis, and thus the dependent variable for the line helps to interpret the dimension represented by the axis. The variables age, cautiousness, and helpfulness all had large values for the variance explained by regression and fit very closely to the coordinate axes interpreted to represent these dimensions in Figure 4 and 5. Bisanz et al. (1978) called the second dimension in the AC story "eagerness to give advice," but the regression analysis suggests that this dimension represents "age" to the readers, which was one of the dimensions used to write the story.

The ratings on "leadership" for the HL story do not help to

Table 2

Multiple regression of dimensions from logical story structure against coordinates of the animals in the INDSCAL solutions for sorting data

	Direction cosines		Multiple correlation	
	Caution	Age		
Age	.17	-.99	.93	$p < .01$
Caution	.99	-.02	.97	$p < .01$
	Helpful	Leader		
Helpfulness	.99	.12	.96	$p < .01$
Leadership	.84	.55	.52	$p < .39$

interpret the second dimension from the INDSCAL solution for the HL story. This was also seen earlier when comparing subjects' sortings of the characters based on similarity in the story to the a priori logical structure. The correlation between the sorting task matrix and the matrix from the logical structure was only .61. A further comparison of the INDSCAL solution from the sorting task (Figure 4) and the logical structure from the text (Figure 6b) reveals that the subjects have grouped the animals consistently on the "helpfulness" dimension, as confirmed by the regression analysis. The tiger, rabbit, mouse, and horse were seen as not helpful, the lion was seen as neutral or slightly towards the positive end of the helpful scale, and the rat, fox, elephant, and sheep were seen as helpful. Not only is there no consistency on the second dimension, but the group space itself is made up of "clusters," rather than objects along a scale. Furthermore, for all the points, the distances between animals in different clusters are larger than distances within animals in the same cluster. This type of clustered solution is analogous to a "degenerate" solution in the context of nonmetric scaling (Kruskal and Wish, 1978). In this case, there appear to be three clusters, one with the helpful animals, one with the animals who were not helpful, and the lion, who does not fit into either cluster. Thus, the readers appear to have organized the characters in the HL story in terms of "helpfulness," but they seem to have "missed" the leadership theme in the story.

EXPERIMENT 2

Experiment 1 suggests that the readers' organization of the animals was different from the pre-reading organization of the animal names. However, the stories were based on dimensions related to personality attributes, and the pre-reading dimensions appeared to be dimensions that are specific to animals—size and predacity. Therefore, before we can affirm that the readers have changed their organization of the animal names after reading stories based on personality dimensions, we must be sure that the change is measured from a pre-reading structure based upon how readers view animals on personality attributes. The purpose of Experiment 2 was to measure change in organization of the animal names, in terms of personality traits, and to determine what information in each story may contribute toward making the story dimensions salient.

Method

Subjects. Subjects were forty undergraduate psychology students, 30 females and 10 males, who participated in the experiment as part of a class requirement.

Materials. Subjects read the same stories used in Experiment 1.

Design and Procedure. Before reading each story, subjects sorted the animal names into groups based on personality traits. They were told that people often think of animals as having certain personality traits, e.g., a donkey is often thought of as "stubborn" and a hare as "quick." They were then asked to group the nine animal names in terms of similar personality traits. Half the subjects were also asked to label the criteria they used for putting the animal names into each group.

The remainder of the study was similar to Experiment 1. Subjects read and sorted the animals based on the stories. Half read and grouped the animals based on the helpful-leader (HL) story first and half the students read and sorted on the age-caution (AC) story first. As a final task, subjects rated statements in the last story read on a 0-3 scale in terms of how relevant the information in each statement was in deciding whether a character possessed a particular personality trait—helpfulness, leadership, cautiousness or age. Subjects were again run in eight groups of five.

Results and Discussion

The crucial question for Experiment 2 was whether subjects grouped the animal names differently before reading when asked to do so based on personality traits rather than just based on similarity. To answer this question, the QA paradigm was used to compare the group matrix from the pre-reading sorting task from Experiment 1 with the group matrix from the pre-reading sorting task for Experiment 2. The correlation between these two matrices was .84 ($z = 4.56$). This result suggests that, in general, the group structures of the data from Experiments 1 and 2 are similar, although there may be some small differences attributable to the different instructions. Thus, semantic memory dimensions of size and predacity (Henley, 1969; Rips, et al., 1973) may be correlated with personality trait names.

The next step was to examine the INDSCAL solution from Experiment 2 for the pre-reading sorting task in an attempt to

discern any differences in the dimensions subjects may have used when grouping animals in terms of personality traits. This solution is shown in Figure 8.

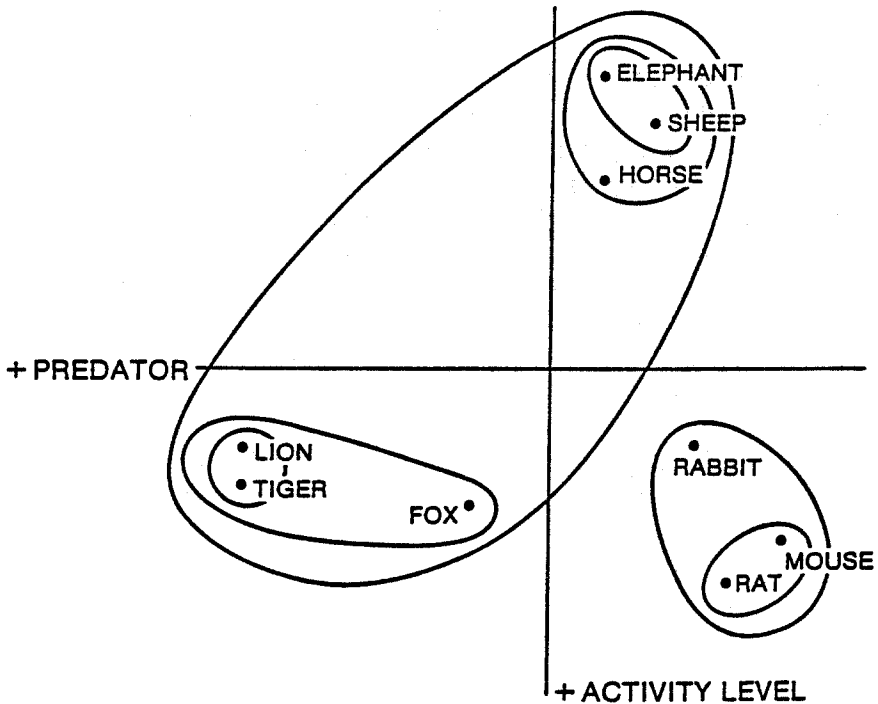


Figure 8: INDSCAL solution for sorting task on personality traits.

To interpret the dimensions, the criteria that subjects listed when labeling the groups were considered, along with the position of the animals in the INDSCAL solutions and hierarchical clustering solutions generated from the group matrix. Hierarchical clustering (S. Johnson, 1967) produces a sequence of partitions of a set of objects in which the partition at each level is formed by joining one or more of the sets based on a distance criterion from the immediately preceding partition. If a multidimensional scaling solution is a valid representation of the distances from a proximity matrix, then a hierarchical clustering solution should produce separate, nonoverlapping clusters that assist in interpreting the dimensions.

The clustering procedure, along with the labeling of the

groups, help to explain the groupings for the sorting task in Experiment 2. The subjects in Experiment 2 appear to be grouping the animals according to predacity and activity level. For example, the lion, tiger, and fox form a cluster, and subjects labeled them as predators and "mentally" quick. The elephant, horse, and sheep formed another group and were described as "gentle" and "un-intelligent" in terms of personality, explaining why they appear lower on both the predator and activity dimensions.

In summary, the structures of the group matrices from Experiments 1 and 2 based on sorting on similarity and personality traits respectively appear the same. The INDSCAL solutions from the pre-reading sorting tasks from both studies also appear similar, although there seem to be some small differences in the dimensions subjects used when sorting the animals that may be attributed to the different instructions for the studies.

Another purpose of Experiment 2 was to measure a change in organization based on personality. The correlation between the pre-reading group task matrix and the pre-reading sorting for the HL story in Experiment 2 was $-.09$ ($z = -.46$); for the AC story the correlation with the prereading sorting matrix was $-.20$ ($z = -1.08$). Therefore, there is no apparent similarity between the patterns of entries in the matrices for sorting based on personality before reading and after reading each story. The correlation between the group matrices for the HL story and AC story was $-.07$ ($z = -.37$), suggesting, as in Experiment 1, that subjects formed a different set of relationships among the characters in each story. In summary, since the pre-reading organization in Experiment 2 was based on personality traits of the animal names, we have evidence for a change in organization of the animal names in relation to both personality traits from one story (AC story) and one of the two traits from the other story (HL story).

A final purpose of Experiment 2 was to try to determine what information in each story may contribute to making the two intended story dimensions salient. Each story was broken into "idea units," statements that express a complete idea. There were 99 statements in the AC story and 85 in the HL story. Subjects rated each statement on a 0-3 scale in terms of how relevant the statement was in determining whether a character possessed one of the two traits in the story, where "3" meant the subject felt the statement was very relevant and "0" meant it had no relevance. For example, the statement "Why should we risk our lives helping

you?" was rated highly relevant to deciding whether a character was helpful, while "Once upon a time" was given a mean rating of "0." Mean ratings for each statement were subsequently calculated.

We found higher mean ratings for statements related to the more salient dimension in each story from the multidimensional scaling solution—cautiousness in the AC story and helpfulness in the HL story. Thus, one reason that each of these dimensions is more salient in each story may be that readers perceive more statements as highly relevant in determining whether the character possesses the trait. It may also be that, particularly for the leadership trait, there is not enough relevant information in the story and the reader is required to make inferences about what it may mean to be a leader. Readers may not all make the same inferences, and looking at the subjects as a whole, they appear to "miss" the leadership theme.

Data from the last task also provides information about the relationship between age and cautiousness, and helpfulness and leadership. The correlation between mean ratings for statements relevant to age or cautiousness was .03, suggesting age and cautiousness may represent independent dimensions. However, the correlation between the ratings for helpfulness and leadership was .58. It may be that readers do not totally distinguish between the latter two dimensions, and possibly characteristics necessary for being a leader may also involve some of the same characteristics required for deciding whether a character is helpful.

GENERAL DISCUSSION

In general, after reading the two stories, subjects' representation of the structure of the characters was different from that produced by the pre-reading sorting task in which subjects grouped the animals based on similarity (Experiment 1) or based on personality traits (Experiment 2). Both multidimensional scaling and the newer QA paradigm have provided a comprehensive representation of the organization of the characters acquired from reading a story narrative.

Comparison of similarity rating and sorting tasks

The analyses of the pre-reading organization of the animal names suggest that, in general, the group structure of the similar-

ity rating task and sorting task data are consistent. Moreover, the techniques of multidimensional scaling and QA procedures provide complementary information about the representation of memory organization measured by these two tasks. Previous use of multidimensional scaling for animal names involved the similarity rating task (Bisanz et al., 1978; Henley, 1969; Rips et al., 1973), and the present results are consistent in terms of the group spaces generated and the interpretation of the dimensions as "size" and "predacity." However, the INDSCAL weight space solution and the information about accounted-for-variance suggest that subjects may be weighting more equally on both dimensions when sorting into groups. In short, while size is the more salient dimension when rating similarity between pairs of animals on a 1-7 scale, both size and predacity appear to be considered by subjects when sorting the animals into groups.

One possible explanation for the difference in dimensional salience for both tasks lies in the decision processes involved. When rating pairs on a 1-7 scale, each pair is judged separately and only one dimension may be considered before moving to the next pair. However, when sorting animals into groups, the groups can be readjusted to allow for "second thoughts" and consideration of more than one dimension.

In the INDSCAL analysis, the salience of "size" rather than "predacity" for both tasks may be due to the depth of processing needed when rating or grouping the animals. Judging animals based on "size" requires only a quick comparison of the general appearance of the animals, whereas deciding on whether animals are predators requires a deeper search of long-term memory. When comparing 36 animal pairs, one pair at a time, the subject would probably use the "easier" dimension on which to judge. When separating all the animals at once as in the sorting task, however, it may be just as efficient to use more than one dimension, either by forming homogeneous groups on each dimension or by having animals in one group that are alike on both dimensions.

Subjects also appeared to group animals in a similar manner whether they grouped based on general similarity (Experiment 1) or on personality traits (Experiment 2). This suggests that perhaps personality traits, at least among animal names, may be correlated with size and predacity.

Passage present versus passage absent

The analysis of readers' story organization based on task performance under the passage present and passage absent conditions from Experiment 1 also provide information comparing the similarity rating task and sorting tasks as measures of memory organization. Apparently, the similarity rating task produces more varied responses than the sorting task when done "from memory" (passage absent condition). The fact that the ratings must be done separately for each pair may make it more difficult for the subjects to "search" their memory for the correct response. Since the sorting task can be done while looking at the names of all nine animals at once, it may be easier for the subjects to reconstruct the correct representation in a more consistent manner. We noted previously that when performing the sorting task in the pre-reading condition, subjects seemed to use both dimensions from the pre-reading structure equally, and they may also be able to recall the dimensions from the story structure more easily when sorting the animals into groups. If we are interested in how the reader has represented the story in memory and not in how well he/she can find the correct responses when looking back over the story text, then the sorting task may provide more consistent data.

Measuring organization of story characters

The INDSCAL version of individual differences scaling and the QA paradigm provided evidence of a change in organization of the animal names after reading each story. Furthermore, this change was not caused merely by the fact that something had been read since the acquired organization was different for the two stories. The INDSCAL solutions are consistent with Bisanz et al. (1978), except that for the AC story both cautiousness *and* age appeared as dimensions in the subjects' acquired cognitive structure. Bisanz et al. interpreted the second dimension from the INDSCAL solution generated from the similarity rating task for the AC story as "eagerness to give advice," not as the "age" dimension from the logical story structure. Although they suggest that the primary theme may have been more dominant, it may also result from the use of the similarity rating task. In the present study, when sorting the animals into groups, the subjects seem to be able to use both themes intended by the authors in relating the characters.

With the HL story, we, like Bisanz et al., found results suggesting that the subjects acquired only the more salient "helpfulness" dimension from the story. Since the story action revolved around whether to help the lion, the readers may simply not have processed the second leadership theme during comprehension. Based on the similarity of the multidimensional scaling solutions for the passage present and passage absent conditions and the failure to find greater homogeneity for the passage present condition, it is clear that the subjects did not just "forget" the leadership theme. They simply were not aware of it in the story. Deciding whether a character is a leader may also not be as readily apparent in the story action and may require more of an inference by the reader than is required for the other story dimension. The inference process may be even more difficult since part of leadership "qualities" may include helpfulness, and the animals are explicitly described as acting in a helpful way or not. Since helpfulness and leadership qualities may reflect partially overlapping sets, readers may see the animals as helpful or not and never complete the inference process in terms of deciding about leadership.

REFERENCES

- Anglin, J. M. *The growth of word meaning*. Cambridge, Mass.: MIT Press, 1970.
- Bisanz, G., LaPorte, R., Vesonder, G., & Voss, J. On the representation of prose: New dimensions. *Journal of Verbal Learning & Verbal Behavior*, 1978, 17, 337-357.
- Bransford, J. D. *Human cognition: Learning, understanding, and remembering*. Belmont, California: Wadsworth, Inc., 1979.
- Carroll, J. D., & Chang, J. J. Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckhart-Young decomposition. *Psychometrika*, 1970, 35, 283-319.
- Desse, J. On the structure of associative meaning. *Psychological Review*, 1962, 69, 161-175.
- Edgington, E. S. *Statistical inference: The distribution-free approach*. New York: McGraw-Hill, 1969.
- Fillenbaum, S., & Rapoport, A. *Structures in the subjective lexicon*. New York: Academic Press, 1971.
- Henley, M. A psychological study of the semantics of animal terms. *Journal of Verbal Learning & Verbal Behavior*, 1969, 8, 176-184.
- Hubert, L. J., & Levin, J. Evaluating object set partitions; free-sort analysis and some generalizations. *Journal of Verbal Learning & Verbal Behavior*, 1976, 15, 459-470.
- Johnson, P. E. Some psychological aspects of subject-matter structure. *Journal of Educational Psychology*, 1967, 58, 75-83.
- Johnson, S. C. Hierarchical clustering schemes. *Psychometrika*, 1967, 32, 241-254.
- Kruskal, J., & Wish, M. *Multidimensional scaling*. Beverly Hills, California: Sage Publications, 1978.

Gail Gliner, Susan R. Goldman and Lawrence J. Hubert

- LaPorte, R., & Voss, J. Prose representation: A multidimensional scaling approach. *Multivariate Behavioral Research*, 1979, 14, 39-56.
- Miller, G. A. A psychological method to investigate verbal concepts. *Journal of Mathematical Psychology*, 1969, 6, 169-191.
- Norman, D. A., & Rumelhart, D. E. *Explorations in cognition*. San Francisco: W. H. Freeman, 1975.
- Preece, R. F. W. Developmental trends in the continued word associations of physics students. *Journal of Research in Science Teaching*, 1977, 14, 235-239.
- Rips, L. J., Shoben, E. J., & Smith, E. E. Semantic distance and the verification of semantic relations. *Journal of Verbal Learning & Verbal Behavior*, 1973, 12, 1-20.
- Shavelson, R. J. Some aspects of the correspondence between content structure and cognitive structure in physics instruction. *Journal of Educational Psychology*, 1972, 63, 225-234.
- Tulving, E. Episodic and semantic memory. In E. Tulving and W. Donaldson (Eds.), *Organization of Memory*. New York: Academic Press, 1972.