Journal of Experimental Psychology:
Learning, Memory, and Cognition

VOL. 11, NO. 4 OCTOBER 1985

Ideals, Central Tendency, and Frequency of Instantiation as Determinants of Graded Structure in Categories

Lawrence W. Barsalou
Emory University

Three possible determinants of graded structure (typicality) were observed in *common taxonomic* categories and *goal-derived* categories: (1) an exemplar's similarity to *ideals* associated with goals its category serves; (2) an exemplar's similarity to the *central tendency* of its category (family resemblance); and (3) an exemplar's *frequency of instantiation* (people's subjective estimates of how often it is encountered as a category member). Experiment 1 found that central tendency did not predict graded structure in goal-derived categories, although it did predict graded structure in common taxonomic categories. Ideals and frequency of instantiation predicted graded structure in both category types to sizeable and equal extents. A fourth possible determinant—*familiarity*—did not predict typicality in either common taxonomic or goal-derived categories. Experiment 2 demonstrated that both central tendency and ideals causally determine graded structure, and work showing that frequency causally determines graded structure is discussed. Experiment 2 also demonstrated that the determinants of a particular category's graded structure can change with context. Whereas ideals may determine a category's graded structure in one context, central tendency may determine a different graded structure in another. It is proposed that graded structures do not reflect invariant structures associated with categories but instead reflect people's dynamic ability to construct concepts.

A central theme in categorization research for the last decade has been that categories possess graded structure. Instead of being equivalent, the members of a category vary in how good an example (or how typical) they are of their category (Rips, Shoben, & Smith, 1973; Rosch, 1973; 1975; Smith, Shoben, & Rips, 1974). In *birds*, for example, American college students agree that *robin* is very typical, *pigeon* is moderately typical, and *ostrich* is atypical. In addition, nonmembers of a category vary in how good a nonmember they are of the category (Barsalou, 1983). For example, *chair* is a better nonmember of *birds* than is *butterfly*. *Graded structure* refers to this continuum of category representativeness, beginning with the most representative members of

The research in this article was supported by Grant MH 13905 from the National Institute of Mental Health to Gordon Bower, a National Science Foundation Graduate Fellowship to the author, and Experimental Training Grant MH 15757-02 from the National Institute of Mental Health to Stanford University. Parts of this article were drawn from a doctoral dissertation submitted to Stanford University (Barsalou, 1981) and an Emory Cognition Project Technical Report (Barsalou, 1984).

I am grateful to Gordon Bower for supporting this project and to Gordon Bower, Herbert Clark, Ellen Markman, and David Rumelhart for direction during its initial development. I am also grateful to Douglas Medin for his comments on every written form of this research and for his guidance in developing it. In addition, this article has benefited from the comments of Jeffrey Farrar, James Hampton, Barbara Malt, Gregory Murphy, Brian Ross, Daniel Sewell, and several anonymous reviewers.

Request for reprints should be sent to Lawrence W. Barsalou, Department of Psychology, Emory University, Atlanta, Georgia 30322.

a category and continuing through its atypical members to those nonmembers least similar to category members. No other variable is as important as graded structure in predicting performance on a wide range of categorization tasks (e.g., category acquisition, exemplar production, category verification). In addition, graded structure occurs in a diverse range of categories, suggesting that it may be a universal property of categories. The large body of work addressing graded structure is reviewed in Mervis and Rosch (1981), Smith and Medin (1981), and Medin and Smith (1984).

This article addresses the issue of what determines graded structure. Why are some exemplars of a category more typical than others? Two experiments examine three possible determinants of graded structure: *central tendency, ideals,* and *frequency of instantiation.* A fourth possible determinant—*familiarity*—is also briefly considered.

Central Tendency

Following the work of Rosch and Mervis (1975), there has been widespread acceptance that an exemplar's typicality depends on its *family resemblance,* where family resemblance is defined as an exemplar's average similarity to other category members and its average dissimilarity to members of contrast categories. The more similar an exemplar is to other category members and the less similar it is to members of contrast categories, the higher its family resemblance, and the more typical it is of its category. *Dog,* for example, is very similar to other members of *mammals* and not very similar to members of contrast categories (e.g., *fish, birds*). In contrast, *whale* is not as similar to other *mammals* and is highly similar to the members of a contrast category (i.e., *fish*). Consequently *dog* is more typical of *mammals* than is *whale.*

Another way to view an exemplar's family resemblance is as its similarity to central tendency (Hampton, 1979; Smith et al., 1974), where central tendency refers to any kind of central tendency information about a category's exemplars (e.g., average, median, or modal values on dimensions, highly probable properties, etc.). As just discussed, an exemplar's family resemblance is defined in part as its average similarity to other category members. However its average similarity to other category members must be at least roughly the same as its similarity to their central tendency (Barsalou, 1983). This is analogous to the average difference between a number and several other numbers being the same as the difference between the first number and the average of the others. In a related manner, an exemplar's average dissimilarity to the members of contrast categories must be at least roughly the same as its dissimilarity to their central tendencies. Consequently an exemplar's family resemblance can be specified either as its average similarity and dissimilarity to category members and nonmembers, or as its similarity and dissimilarity to their central tendencies.

Although people could determine family resemblance in either of these two ways, determining family resemblance through comparisons to central tendencies may be more psychologically plausible—comparing an exemplar to central tendencies requires much fewer comparisons than comparing an exemplar to members and nonmembers. Regardless of how people actually derive family resemblance, however, similarity to central tendencies and similarity to members and nonmembers are functionally equivalent at the level of predicting typicality. Because this article primarily addresses functional relations between typicality and other variables, family resemblance and similarity to central tendencies will be assumed to be equivalent.

Ideals

Ideals, which provide another possible determinant of graded structure, are characteristics that exemplars should have if they are to best serve a goal associated with their category. For example, an ideal for *foods to eat on a diet* is *zero calories.* The fewer calories an exemplar has, the better it serves the goal associated with its category, namely, *lose weight.* This ideal appears to determine graded structure in that exemplars with decreasing numbers of calories become increasingly good exemplars of the category. Similarly for *things to take from one's home during a fire,* finding exemplars near the ideal of *highest possible value* is relevant to the goal of *minimizing loss;* therefore this property appears to determine the category's graded structure.

Most categories probably have more than one ideal. For example, *possible restaurants to eat at* may have the ideals of *lowest possible cost, highest possible quality,* and *closest possible proximity.* The most important ideal(s) on a given occasion may depend on the goal a person is pursuing. If the goal is to have a memorable experience, then *high quality* may be most important. But if the goal is to have a quick meal, then *high quality* may succumb to *close proximity* and *low cost.*

Ideals differ from central tendency in at least two ways. First, ideals generally do not appear to be the central tendencies of their categories (although they may occasionally be). *Zero calories,* for example, is certainly not the central tendency with respect to *calories* for *things to eat on a diet;* nor is *closest possible proximity* the central tendency with respect to *distance* for *possible places to eat at.* Ideals tend to be extreme values that are either true of only a few category members or true of none at all. Instead of lying at the center of categories (as does central tendency), they generally lie at the periphery.[1]

Central tendency and ideals also differ in origin. Central tendency depends directly on the exemplars of a category, and more specifically, on the particular exemplars a person has experienced. Although people may form impressions of a category's central tendency through hearsay, they may generally acquire such information through experience with exemplars. In contrast, ideals may often be determined independently of exemplars, being acquired through the process of planning how to achieve goals before exemplars are ever encountered.

Frequency of Instantiation and Familiarity

Rosch, Simpson, and Miller (1976) and Mervis, Catlin, and Rosch (1976) argued that frequency does not determine graded structure, although their tests of frequency were not very sensitive. Rosch, Simpson, et al. (1976) pitted family resemblance against frequency and found that only family resemblance predicted typicality. However, their design was not capable of detecting simultaneous effects of family resemblance and frequency. Consequently frequency could have had an effect, but was not detected because it was the weaker of the two factors. Mervis et al. (1976) found that an exemplar's word frequency in Kučera and Francis's (1967) analysis did not predict typicality. However it is by no means clear that word frequency is a good measure of how often people encounter exemplars in their everyday routines. Other measures of frequency may be better predictors of typicality.

More recent work has contradicted these initial reports, finding that familiar exemplars are perceived as more typical than unfamiliar exemplars (Ashcraft, 1978; Glass & Meany, 1978; Hampton & Gardiner, 1983; Malt & Smith, 1982).[2] Familiarity can be defined as someone's subjective estimate of how often they have experienced an entity across *all* contexts. However an alternative form of frequency that could determine graded structure is frequency of instantiation, which can be defined as someone's subjective estimate of how often they have experienced an entity as a member of a *particular category.* Whereas familiarity is a *category-independent* measure of frequency, frequency of instantiation is a *category-specific* measure of frequency. For example, people generally appear more familiar with *chair* than with *log,* having experienced *chair* more often across all contexts. However people have probably experienced *log* more often as an instantiation of *firewood.* Increases in familiarity and frequency of instantiation could both be associated with increasing typicality. Although both possible determinants receive attention here, the focus will be on frequency of instantiation, because initial inspection of categories suggested it as the more important factor.

In summary, a number of factors could determine graded structure, including central tendency, ideals, frequency of instantiation, and familiarity. Because previous work has not observed ideals and frequency of instantiation, and because previous work has generally not

[1] Ideals are not always the most extreme values possible on a dimension. Exemplars of *clothes to wear in the snow,* for example, vary along the dimension of *how warm they keep people,* with the ideal not being *as warm as possible* (which could be fatal) but being instead *as much warmth as is necessary for survival and comfort.* This ideal, however, is probably not the central tendency of the category.

[2] McCloskey (1980) also reported effects of familiarity on conceptual processing, but his work primarily addressed the role of a category term's familiarity instead of an exemplar term's familiarity, which is of interest here.

performed comprehensive tests of possible determinants, one of the purposes of this project was to observe all four of these possible determinants simultaneously. The focus on these factors is not meant to imply that they are the only possible determinants of graded structure. Instead it is highly likely that other factors also determine typicality. For example, Hampton and Gardiner (1983) review findings that address whether the number of properties associated with an exemplar determines its typicality. In addition, Lakoff (in press) presents a number of other possible determinants, which in at least some cases appear to be composites of the factors examined here.

Common Taxonomic and Goal-Derived Categories

Previous work showing that central tendency and familiarity determine graded structure has focused on typicality in one particular kind of category, namely, *common taxonomic* categories (e.g., *birds, furniture, fruit*). However, because Hampton (1981) found that graded structure is not well-predicted by central tendency in some abstract categories, there is reason to believe that the generality of the previous studies is limited. The factors that determine graded structure may vary widely across categories. Consequently a second purpose of this project was to observe typicality in another kind of category, what will be referred to as *goal-derived* categories (e.g., *things not to eat on a diet, things to take from one's home during a fire, birthday presents*).

It should be noted that Barsalou (1983) found graded structure in *ad hoc* categories, which are those goal-derived categories that have been constructed to achieve a novel goal and that therefore are not well-established in memory. Once an ad hoc category is frequently used and becomes well-established in memory, however, it is no longer ad hoc by this definition (see Barsalou, 1983, pp. 224–225). Consequently goal-derived categories include both ad hoc categories and better established categories that were once ad hoc.

Of course it would be ideal to distinguish common taxonomic and goal-derived categories in terms of simple definitions. Unfortunately such definitions have not as yet been forthcoming, although it is as least possible to provide characteristic properties for each category type. One way common taxonomic and goal-derived categories generally appear to differ has to do with the "correlational structure of the environment." Correlational structure refers to the fact that properties in the physical environment are not independent; that is, a given property generally co-occurs with certain other properties but not with others. *Feathers,* for example, typically co-occurs with *wings* and *beak,* but not with *tires* and *engine.* As discussed by Rosch and Mervis (1975) and Rosch, Mervis, Gray, Johnson, and Boyes–Braem (1976), common taxonomic categories appear to circumscribe sets of things in the environment that share these clusters of co-occurring properties. Consequently these categories reflect the correlational structure of the environment. Many exemplars of *birds,* for example, share co-occurring properties that rarely occur outside the category, thereby making exemplars of this category very similar to each other and very dissimilar to nonmembers.

In contrast, goal-derived categories generally appear to violate the correlational structure of the environment. Many goal-derived categories include some members from each of several common taxonomic categories, but never all the members from a given one. *Things to take on a camping trip,* for example, includes members of *food, clothing, tools,* and so on, but it does not include all members. Because the members of these goal-derived categories are often quite dissimilar to each other and very similar to many nonmembers, they do not maximize the correlational structure of the environment. Other goal-derived categories contain subsets of one particular common taxonomic category. For example, someone with a back problem might be interested in *chairs that provide good back support.* In these cases, goal-derived categories do not maximize correlational structure because many noncategory members are highly similar to category members (e.g., *chairs that provide good back support* are very similar to *chairs that do not*). In general, because goal-derived categories do not maximize the correlational structure of the environment, they are not very salient and do not stand out as natural groups. Instead they appear to only become salient when relevant to currently pursued goals.

Another way in which common taxonomic and goal-derived categories generally seem to differ has to do with category use. Common taxonomic categories are often used for classification, whereas goal-derived categories are often used for instantiation. When classifying entities in the environment, people primarily appear to use common taxonomic categories. More specifically, people generally prefer basic level categories, which are a subset of common taxonomic categories (Jolicoeur, Gluck, & Kosslyn, 1984; Murphy & Smith, 1982; Rosch, Mervis, et al., 1976; B. Tversky & Hemenway, 1984). In contrast, people primarily appear to use goal-derived categories for instantiating schema variables while achieving goals. To achieve the goal of *taking a vacation,* for example, a planner has to instantiate variables in schematic knowledge about *vacations,* such as *where to go, who to go with, how to get there, what to take,* and so on. The goal-derived categories of *places to go, people to go with, types of transportation,* and *things to pack in a suitcase* facilitate locating and selecting instantiations for these variables. In general, successfully achieving a goal requires that people bind schema variables with instantiations appropriate in the current setting. Goal-derived categories provide pools of instantiations from which instantiations can be chosen.

Although common taxonomic and goal-derived categories generally seem to differ in the extent to which they reflect correlational structure and in the way they are used, these distinctions are by no means clear-cut or defining. For example, some common taxonomic categories such as *vehicles, clothing,* and *furniture* are highly related to people's goals and may often be used for instantiation. Conversely goal-derived categories may at times be used for classification. Although these distinctions are not defining, they provide characteristic properties of common taxonomic and goal-derived categories.

Determinants of Graded Structure in Common Taxonomic and Goal-Derived Categories

These general differences between common taxonomic and goal-derived categories suggest that different factors may determine their graded structures. To begin with, central tendency may be highly salient in people's representations of common taxonomic categories and thereby become the standard by which typicality is judged. Central tendency may be salient in common taxonomic categories for the following two reasons. First, because these categories generally reflect correlational structure, people may use them as a means of representing the structure of the environment. If so, then acquiring central tendency information for these categories provides *representative information* about the kinds of entities the environment contains. Central tendency information is clearly more representative than ideal information, because the former has a much higher likelihood of occurring for a category's exemplars than the latter.

A second reason central tendency information may be salient in common taxonomic categories has to do with their use. Because these categories are often used for classification, their representations may be designed to maximize classification performance. It is well-known in the category verification literature that classifying an entity proceeds faster to the extent the entity is similar to the category standard (Smith, 1978). Basing classification standards on central tendency information minimizes the average distance of category members to category standards (e.g., it is a statistical fact that the average absolute distance from all points in a set to one particular point is minimized when that point is the median; Hayes, 1973, p. 223). Therefore the average difficulty of performing classifications is minimized when the category standard contains central tendency information (as opposed to ideals). Because a primary use of common taxonomic categories is to serve classification, it would not be surprising if central tendency information were central to their representations.

In contrast, ideals may become highly salient in people's representations of goal-derived categories and thereby become the standards by which typicality is judged. Because goal-derived categories generally serve goals, their representations may contain ideals in order to maximize goal achievement. As people consider possible instantiations of a category, they can compare them to the category's ideals and thereby pick the exemplar or exemplars that will result in maximal goal satisfaction. For example, someone on a diet might compare selections on a menu to the ideals for *things to eat on a diet* in order to pick instantiations that will maximize the goal of *losing weight.*

Although central tendency may be the only determinant of graded structure for common taxonomic categories, and although ideals may be the only determinant for goal-derived categories, another possibility is that both factors determine graded structure in both category types. For example, the typicality of a particular *fruit* may be determined, not only by its similarity to central tendency information, but also by its similarity to revelant ideals (e.g., having to do with *taste* and *nutrition*). To the extent common taxonomic categories serve goals, ideals should also determine their graded structures. Analogously, to the extent central tendency information is important for the use of goal-derived categories, it should also determine their graded structures. On the basis of findings reported by Rosch and Mervis (1975) and Rosch, Simpson, et al. (1976), one might predict that central tendency determines graded structure in all categories.

Finally, initial inspections of categories suggested that frequency of instantiation determines graded structure to some extent in goal-derived categories but not in common taxonomic categories. Alternatively, the work of Rosch, Simpson, et al. (1976) and Mervis et al. (1976) suggests that frequency of instantiation should not determine graded structure in either.

Experiment 1

This first study examined whether central tendency, ideals, and frequency of instantiation predict graded structure in goal-derived and common taxonomic categories. For each category observed, these three variables were measured for every exemplar and were then correlated with typicality. This study also observed whether central tendency, ideals, and frequency of instantiation predict how often exemplars are generated during exemplar production.

Method

Nine goal-derived categories were selected that intuitively appeared to originate during goal-directed behavior as opposed to originating in the correlational structure of the environment. The common taxonomic categories were 9 of those studied by Rosch (1975). These 18 categories are shown in Table 3.

Thirty-eight subjects generated exemplars of these 18 categories after generating exemplars to each of three practice categories. Subjects received the 18 critical category names in one of two random orders. After a tape recording finished stating a name, subjects wrote down as many exemplars as they could think of in the subsequent 15-s interval. The recording then asked subjects to turn to the next page of their booklet and prepare for the next category.

The exemplars generated for each category were pooled across subjects to construct dominance orders of exemplars (as in Battig & Montague, 1969). Generated items were considered to be the same exemplar only if they were orthographically identical or differed by a minor inflection (e.g., *shirt* and *shirts; walk* and *walking*). All exemplars generated by only one subject were not used in the remainder of the experiment. However when two such exemplars were members of some superordinate not mentioned by two or more subjects, the superordinate was included (e.g., *Kung Fu* and *Karate* were combined to form *martial arts*). This occurred rarely. The number of exemplars generated by 2 or more subjects per category ranged from 9 to 24, the median being 19.83. The median for the goal-derived categories was 20.25 and for the common taxonomic categories was 19.25.

To obtain exemplar goodness judgments, subjects received the 348 exemplars blocked by category. In each of two versions, the category blocks were randomly ordered (one per page) as were the exemplars within each category. At the top of each page appeared the corresponding category name. To the right of each exemplar appeared a 9-point scale on which 1 was labeled *poor example* and 9 was labeled *excellent example.* Ten subjects circled 1 scale number for each exemplar to rate how good an example it was of its category. These instructions did not ask subjects to judge "how typical" each exemplar was because it was thought this might bias subjects towards using frequency of instantiation. "How good an example" seemed more open ended and less demanding. Although "typicality" will often be used to refer to these data out of convenience, they were collected using "goodness-of-example" instructions.

To obtain frequency of instantiation judgments, subjects received the same materials as just mentioned, except the endpoints for the scale were labeled 1 for *not frequently at all* and labeled 9 for *very frequently.* Ten subjects rated each member for how frequently they thought it subjectively occurred as a category instantiation. Subjects were explicitly asked not to judge how familiar each item was but instead how frequently they thought it occurred as a member of the category.

Except for the following two changes, the materials used to collect judgments about ideals were the same as those used to collect judgments of typicality and frequency. First, the endpoints for the scale was labeled 1 for *very low amount* and labeled 9 for *very high amount.* Second, the name of a dimension occurred at the top of the page. These dimensions—which will be referred to as *ideal dimensions*—are shown in Table 3. They were picked intuitively and seemed to contain ideals that exemplars should optimally have with respect to a goal served by the respective category.

The ideal for each dimension appeared to lie toward its upper end. Ten subjects rated each exemplar for its amount on its ideal dimension.[3]

It should be noted that a category could have more than one ideal or have a more important ideal than the one observed here. The goal of this study, however, was not to find the upper limit on how well ideals predict graded structure. Instead the goal was simply to determine whether ideals predict graded structure at all. Consequently the attempt to locate ideals was not exhaustive or oriented towards finding the most important ideal.

To assess the role of central tendency, a family resemblance score was obtained for each exemplar. Because an exemplar's average similarity to other exemplars is at least roughly the same as its similarity to their central tendency (as discussed earlier), these scores at least approximate how similar an exemplar is to its category's central tendency. All possible pairs of exemplars were formed for each category, with the two exemplars in each pair being randomly ordered. Two versions were formed in which pairs were blocked by category. In each, the category blocks were randomly ordered as were the pairs within each category. At the top of each page of pairs appeared the corresponding category name. To the right of each pair appeared the numbers from 1 to 9, with 1 labeled as *not similar at all* and 9 labeled as *very similar.* Six subjects circled 1 scale number to rate the similarity of each of the 3,319 pairs. Subjects were asked to think of the referent of each word in a pair and then rate the referents' similarity on the 9-point scale.

A subject's similarity judgments for pairs having the same exemplar were averaged to form its family resemblance score. All 6 subjects' family resemblance scores for the exemplar were then averaged to form its overall family resemblance score. Because Rosch and Mervis (1975) and A. Tversky (1977) report that similarity ratings and overlap in feature listings correlate around .90, this rating-based measure of family resemblance should be very close to the feature-based measure of Rosch and Mervis (1975).[4]

Subjects were 74 Stanford University students participating to earn either course credit or pay. A given subject provided data for only one of the five measures.

Results

Raw correlations. An item's exemplar goodness, frequency of instantiation (referred to as *frequency* in the next three sections), and ideal dimension scores were simply averages across the 10 subjects who produced the respective data. An exemplar's family resemblance score was computed as just discussed. An exemplar's output dominance score was the number of subjects, out of 38, who had generated the exemplar. The values of these variables for each exemplar in each category are shown in the appendix.

For each of the five measures, the mean and standard deviation of the exemplar means were computed for each category. Goal-derived and common taxonomic categories did not differ on the mean or standard deviation of any measure. No differences in standard deviations indicates that differences in range between category types will not be a factor in the correlations to be reported shortly. Of more theoretical interest is that goal-derived and common taxonomic categories did not differ in the standard deviations of their family resemblance scores (the mean values were .65 and .70, respectively). This suggests that the exemplars of goal-derived categories vary as much in their similarity to one another as do the exemplars of common taxonomic categories.[5]

For each category, the 10 possible correlations between measures were computed across exemplar averages. The average values for goal-derived and common taxonomic categories are shown in Table 1. The variance of each correlation type was computed across the nine categories within each category type. These 20 variances were averaged to form a pooled estimate of the variance, and the resulting standard deviation with 160 degrees of freedom was .24. This value was used in t tests to determine whether average correlations differed from zero. Correlations in Table 1 whose absolute value is equal to or greater than .16 are significant at the .05 level. For the .01 and .001 levels, the corresponding values are .21 and .36, respectively. The pooled estimate of the variance was also used to test differences between means in Table 1. Differences equal to or greater than .22 are significant at the .05

---

[3] Some of these ideal dimensions could actually be composites of several ideal dimensions. For example, the ideal dimension for *birds* (i.e., *how much people like it*) could depend on the more specific dimensions of *how colorful it is, how melodic its song is,* and so on.

[4] It should be noted that this measure of family resemblance does not reflect an exemplar's dissimilarity to members of contrast categories. However Rosch and Mervis (1975) did not include dissimilarity to nonmembers in some of their studies and yet found that similarity to members alone often did an excellent job of predicting typicality (i.e., correlations around .90). So although the measure used here does not reflect dissimilarity to nonmembers, it should indicate at least to some extent whether family resemblance predicts graded structure.

[5] An alternative explanation is that subjects maintain a constant distribution of similarity ratings for each category such that the distribution of family resemblance scores also remains constant.

Table 1
Average Raw Correlations From Experiment 1

Correlation	EG	CT	FOI	I
Goal-derived categories				
OD	.39	.10	.31	.22
EG		.38	.72	.70
CT			.30	.36
FOI				.56
Common taxonomic categories				
OD	.55	.24	.59	.45
EG		.63	.47	.46
CT			.10	.03
FOI				.49

Note. The correlations are averages across categories. OD is output dominance, EG is exemplar goodness, CT is central tendency, FOI is frequency of instantiation, and I is ideals.

level. For the .01 and .001 levels, the corresponding values are .28 and .36, respectively.

Most of the average correlations were significant for the goal-derived categories. Because the highest correlations were between ideals and exemplar goodness and between frequency and exemplar goodness, it appears that ideals and frequency are the determinants most central to the structure of goal-derived categories. Although central tendency (as measured by family resemblance) correlated significantly with typicality, it predicted typicality significantly less than did ideals and frequency.

Most of the average correlations were also significant for the common taxonomic categories. As predicted, central tendency predicted exemplar goodness. However ideals and frequency predicted exemplar goodness just as well for these categories (i.e., there were no significant differences between these three correlations).

As found by Barsalou (1983), typicality and production frequency correlated less for goal-derived than for common taxonomic categories. Better established representations in memory for common taxonomic categories may be the source of this difference. Highly familiar categories may have strong associations in memory from their category concepts to their typical examplars such that these exemplars are usually the first ones generated. Less established associations for goal-derived categories may result in generation being more random such that output dominance and typicality are not as well correlated. Consistent with this explanation is the additional finding that subjects generated exemplars at a faster rate for common taxonomic than for goal-derived categories. The average number of exemplars produced by a subject per category during the 15 s generation period was 4.28 for common taxonomic categories and 3.58 for goal-derived categories, $t(16) = 2.86$, $SE = .25$, $p < .02$.

Partial correlations. The following analysis assumed that there are three predictor variables—central tendency, ideals, and frequency—and that there are two criterion variables—exemplar goodness and output dominance. This analysis examines the possibility that some of the significant correlations discussed in the last section primarily resulted from variance shared between predictors. That is, a predictor could have correlated significantly with a criterion because it shared substantial variance with another predictor that was more strongly correlated with the criterion. For example, central tendency might have correlated with typicality in goal-derived categories because it was highly correlated with two stronger predictors, namely, ideals and frequency.

To assess the *unique* predictive power of each predictor, partial correlations were computed to remove shared variance. Of interest was how well a given predictor correlated with each criterion variable after the other two predictors had been partialed out. For each category, therefore, the second-order partial correlation was computed for each predictor, this being done separately for exemplar goodness and output dominance. The averages across categories are shown in Table 2.

The variance for each of the six second-order partial correlations was computed across the nine categories within each category type. These 12 variances were averaged to form a pooled estimate of the variance, and the resulting standard deviation with 96 degrees of freedom was .25. This value was used in t tests to determine which overall correlations differed from zero. Second-order partial correlations in Table 2 whose absolute value is equal to or greater than .16 are significant at the .05 level. For the .01 and .001 levels, the corresponding values are .21 and .27, respectively.

Table 2
Average Second-Order Partial Correlations From Experiment 1

Correlation	Original correlation	Second-order partial correlation
Goal-derived categories		
EG–CT	.38	.05
EG–FOI	.72	.51
EG–I	.70	.44
OD–CT	.10	−.05
OD–FOI	.31	.25
OD–I	.22	.04
Common taxonomic categories		
EG–CT	.63	.71
EG–FOI	.47	.36
EG–I	.46	.45
OD–CT	.24	.24
OD–FOI	.59	.43
OD–I	.45	.26

Note. OD is output dominance, EG is exemplar goodness, CT is central tendency, FOI is frequency of instantiation, and I is ideals.

The pooled estimate of the variance was also used to test differences between means. Differences equal to or greater than .24 are significant at the .05 level. For the .01 and .001 level, the corresponding values are .32 and .41, respectively.

For the goal-derived categories, the original relation between central tendency and typicality completely disappeared when ideals and frequency were partialed out. Central tendency accounted for no unique variance in the graded structures of goal-derived categories. Instead its apparent predictive power resulted from variance it shared with the other two predictors. Although partialing out the other predictors resulted in a loss of predicted power for ideals and frequency, each accounted for unique variance in the graded structures of goal-derived categories. Because frequency still accounted for unique variance after central tendency had been partialed out, frequency is not an artifact of central tendency as suggested by Rosch (1974).

For the common taxonomic categories, central tendency became an even better predictor of typicality after ideals and frequency were partialed out (both were suppressor variables; e.g., see Allen & Yen, 1979). In addition, central tendency was a much better predictor of typicality in common taxonomic than in goal-derived categories. It should be pointed out that this difference was not due to smaller ranges for these variables in goal-derived categories, because these ranges did not differ (as reported earlier).

Although central tendency accounted for a large amount of unique typicality variance in common taxonomic categories, ideals and frequency also accounted for significant amounts of unique typicality variance in these categories. In fact, these two second-order partials did not differ significantly from the corresponding ones for the goal-derived categories. Although ideals and frequency accounted for significant amounts of unique typicality variance in common taxonomic categories, they were significantly less predictive than central tendency.

Table 3 shows the second-order partial correlations for individual categories. Although there was much consistency within category types, categories varied substantially in their best predictors. In general, the goal-derived categories were most often structured by both ideals and frequency. However *camping equipment* and *picnic activities* were structured primarily by frequency, and *snow clothes* and *weekend entertainment* were structured primarily by ideals. Central tendency structured every common taxonomic category. However these categories differed from each other in the extent to which ideals and frequency were important. Both were important for *vehicles, birds,* and *weapons.* Frequency alone was important for *fruit,* and ideals alone was important for *clothing* and *sports.* It should be noted that failure of an ideal to predict typicality does not mean its category has no ideals, because the important ideals may not have been observed. Moreover some categories may have several ideals, all of which must be observed to estimate the role of ideals completely.

Turning to the second-order partial correlations in Table 2 for output dominance, the most important predictor for both category types was frequency. The importance of frequency for predicting output dominance is consistent with one of the oldest assumptions of memory theory: The more often two things co-occur—in this case a category concept and an exemplar concept—the more likely one is

Table 3
Second-Order Partial Correlations by Categories From Experiment 1

Correlation	EG–CT	EG–I	EG–FOI
Goal-derived categories			
Birthday presents (how happy people are to receive it)	.42	.53	.80
Camping equipment (importance to survival)	.15	−.12	.66
Transportation for getting from San Francisco to New York (how fast it gets people there)	−.51	.56	.40
Personality characteristics in people that prevent someone from being friends with them (how much people dislike it)	−.06	.78	.45
Things to do for weekend entertainment (how much people enjoy doing it)	.34	.43	.08
Foods not to eat on a diet (how many calories it has)	.31	.53	.62
Clothes to wear in the snow (how warm it keeps people)	−.22	.64	.34
Picnic activities (how much fun people think it is)	−.28	.17	.83
Things to take from one's home during a fire (how valuable people think it is)	.29	.47	.41
Common taxonomic categories			
Vehicles (how efficient a type of transportation it is)	.86	.63	.53
Clothing (how necessary it is to wear it)	.71	.81	−.10
Birds (how much people like it)	.75	.42	.78
Weapons (how effective it is)	.59	.91	.68
Vegetables (how much people like it)	.69	−.02	.29
Sports (how much people enjoy it)	.74	.53	.11
Fruit (how much people like it)	.71	.34	.49
Furniture (how necessary it is to have)	.84	.03	.14
Tools (how important it is to have)	.49	.37	.29

Note. Ideal dimensions are in parentheses. EG is exemplar goodness, CT is central tendency, FOI is frequency of instantiation, and I is ideals.

to elicit the other in a production task. The pattern of prediction varied for the two category types with regard to the other two predictors. Neither central tendency or ideals uniquely predicted output dominance for goal-derived categories, whereas both did for common taxonomic categories.[6]

The relation between typicality and the central tendencies of ideal dimensions. If central tendency is the sole determinant of typicality, then a category's central tendency on an ideal dimension should be a better predictor of typicality than its ideal value on that dimension. More specifically, the distance of an exemplar from the central tendency of an ideal dimension should be a better predictor of typicality than its distance from the ideal value of that

[6] In further analyses, central tendency, ideals, frequency of instantiation, and familiarity (to be discussed in a moment) were regressed upon typicality and output dominance for each category type in stepwise multiple regressions. For typicality in common taxonomic categories, significant factors entered in the order of central tendency, ideals, and frequency of instantiation to account for 64% of the variance. For typicality in goal-derived categories, significant factors entered in the order of frequency of instantiation, ideals, and familiarity to account for 69% of the variance (central tendency never entered). For output

dimension. In *things not to eat on a diet,* for example, exemplars should become more typical as they approach the central tendency for the ideal dimension of *calories* than as they approach the ideal value of *indefinitely many.*

To test this possibility, the difference was computed between each exemplar's average rating on the ideal dimension and its category's average rating on that dimension (i.e., the average rating across all exemplars in the category). The absolute values of these differences, which reflect distance from central tendency, were then correlated with typicality. If exemplars become more typical as their value on an ideal dimension approximates the central tendency of the dimension—that is, as exemplars possess increasing family resemblance on that dimension—then large negative correlations between absolute differences and typicality should occur. The average correlation, however, was −.19 for the common taxonomic categories and −.01 for the goal-derived categories. The proximity of exemplars to the central tendencies of ideal dimensions does not predict typicality.

In contrast, the proximity of exemplars to the ideal values of ideal dimensions does predict typicality. Because the ideal values of these dimensions were all at their upper end, positive correlations between typicality and amount on ideal dimensions indicate that typicality is increasing as exemplars approximate ideal values. As reported earlier, the average correlation between typicality and amount on ideal dimensions was .70 for goal-derived categories and .46 for common taxonomic categories. Even after partialing out the other two predictors, increasing values on ideal dimensions still correlated substantially with typicality (.44 and .45, respectively). Ideal values on these dimensions are clearly more important to typicality than central tendencies.

Familiarity versus frequency of instantiation. As noted earlier, Ashcraft (1978), Glass and Meany (1978), Hampton and Gardiner (1983), and Malt and Smith (1982) all found that typicality increased as exemplars became more familiar.[7] It is therefore important to determine if familiarity entered into subjects' frequency of instantiation judgments in the current experiment. To assess the relation between familiarity and frequency of instantiation, 10 additional subjects rated each exemplar of each category on a 9-point scale for how familiar they were with that kind of thing. Subjects received the exemplar names in one of two random orders. In contrast to the procedure for rating frequency of instantiation, the exemplars were not blocked by category, and nothing was said about the categories underlying the list. The average ratings for exemplars (which are shown in the appendix) were then correlated with other measures of interest for each category, and these correlations were averaged across categories within the two category types.

The average correlation between familiarity and frequency of instantiation was greater for common taxonomic categories, .57, than for goal-derived categories, .21, $t(16) = 3.60$, $SE = .10$, $p < .01$. But both values are much less than one would expect if familiarity and frequency of instantiation measure the same thing. According to reliability theory (see Guilford & Fruchter, 1973, pp. 263–264; Barsalou & Sewell, 1984), these correlations should approximate their group reliabilities if the two measures are identical. Given that the mean group reliabilities were .89 for familiarity and .78 for frequency of instantiation, these two measures are not the same.

Regarding the prediction of typicality, familiarity was a much poorer predictor than frequency of instantiation. The average correlation between familiarity and typicality was only .03 for the goal-derived categories and only .19 for the common taxonomic categories. These correlations are much smaller than those

dominance in common taxonomic categories, significant factors entered in the order of frequency of instantiation, ideals, and central tendency to account for 37% of the variance. For output dominance in goal-derived categories, the only significant factor was frequency of instantiation, which accounted for 5% of the variance.

[7] It should be noted that Hampton and Gardiner's (1983) operationalization of familiarity is perhaps closer to what is meant by "frequency of instantiation" here. More specifically, Hampton and Gardiner had subjects judge the familiarity of exemplars while they were blocked together in the context of their categories. As will be seen in this next analysis, making frequency judgments about exemplars varies systematically as a function of whether these judgments are made in the context of categories (frequency of instantiation) or not in the context of categories (familiarity).

just reported between frequency of instantiation and typicality (.72 for goal-derived categories and .47 for common taxonomic categories). Moreover when frequency of instantiation was partialed out of the correlations between familiarity and typicality, they became slightly negative, indicating that familiarity did not account for any unique variance (−.16 for goal-derived categories and −.11 for common taxonomic categories). In contrast, when familiarity was partialed out of the correlations between typicality and frequency of instantiation, the original correlations were unaffected (from .74 to .72 for goal-derived categories, and from .45 to .47 for common taxonomic categories).

Contrary to what earlier studies suggest, familiarity per se does not appear important to graded structure. Instead its relation with typicality appears to reflect variance it shares with frequency of instantiation.

Discussion

Although the work of Rosch and Mervis (1975) and Rosch, Simpson, et al. (1976) shows that central tendency can be an important determinant of graded structure, it clearly does not determine the graded structure of every category. Although central tendency (as measured by family resemblance) had a raw correlation of .38 with typicality for the goal-derived categories, it dropped to .05 when the other two predictors were partialed out. Central tendency accounted for no unique variance of its own, but only correlated with typicality initially because of variance it shared with two stronger predictors. These other predictors—ideals and frequency of instantiation—each accounted for substantial amounts of unique typicality variance in goal-derived categories.

Barsalou (1981, Experiment 1) provides a replication of this finding and also shows that it holds within specific subjects. After generating exemplars for 12 goal-derived categories, each subject provided typicality, frequency of instantiation, ideal dimension, and family resemblance judgments for the exemplars he or she generated. Correlations between the three predictors and typicality were computed individually for each subject within each category. Across subjects and categories, the average raw correlation between central tendency (as measured by family resemblance) and typicality was −.15. The highest value that the average correlation across categories ever reached for a given subject was .19, whereas the lowest value was −.58. Similar to the current experiment, ideals and frequency of instantiation correlated highly with typicality (average raw correlations across subjects of .51 and .60, respectively).

The role of central tendency clearly varies between common taxonomic and goal-derived categories. Although central tendency plays no role in the graded structures of goal-derived categories, it is clearly the most important determinant of graded structure in common taxonomic categories, correlating .71 with typicality after ideals and frequency of instantiation were partialed out. However factors besides central tendency also play roles in the graded structures of common taxonomic categories. Contrary to Mervis et al. (1976) and Rosch, Simpson, et al. (1976), frequency is important to graded structure. Frequency of instantiation accounted for as much unique variance in common taxonomic categories as in goal-derived categories. Moreover, because frequency of instantiation accounted for typicality variance even after central tendency had been partialed out, frequency is not an artifact of central tendency, as suggested by Rosch (1974). In addition, familiarity exhibited no significant relation with typicality for either the common taxonomic or goal-derived categories. Instead the important form of frequency appears to be frequency of instantiation. The more often exemplars are subjectively perceived in the context of their category, the more typical they become of it.

Ideals also accounted for significant amounts of unique variance in common taxonomic categories. In fact, ideals were as important in common taxonomic categories as they were in goal-derived categories. Although central tendency is the most important determinant of graded structure in common taxonomic categories, the structure of these categories also seems to depend on goals these categories serve. Because the ideal dimension of *how effective it is* had a second-order partial of .91 with typicality in *weapons*, goal-related information is clearly central to this category's graded structure. Similarly, the concept for *ve-*

hicles appears to contain goal-related information regarding *efficiency,* and the concepts for *sports, birds,* and *fruit* appear to contain goal-related information regarding *enjoyment.*

These data indicate that graded structure is not a fixed product of central tendency. Ideals and frequency of instantiation also play significant roles in determining the graded structures of categories. Consequently the origins of graded structure are much more complex than has so far been credited by previous work. This flexibility is underlined by the fact that individual categories vary substantially in the combination of factors that determines their graded structure (see Table 3).

In a related study, Barsalou and Sewell (1985) found a similar pattern of results for category verification. Analogous to the current study, central tendency predicted verification time in common taxonomic categories but not in goal-derived categories, whereas frequency of instantiation was the best predictor in goal-derived categories. Although ideals and familiarity correlated with verification time in both category types, they did so primarily through variance shared with central tendency and frequency of instantiation. Because this pattern of results for category verification is similar to the one just reported for typicality, it appears that similar processes may underlie these two categorization tasks.

Finally, it should be noted that the experiment by Barsalou and Sewell (1985), along with the one just reported, both suffer from the following limitation. Because only one ideal was observed for each category, and because each ideal was picked intuitively, these experiments only show that ideals are related to typicality; they do not provide accurate estimates of the strength of this relation (although they do provide minimum estimates). It is likely that more careful and exhaustive sampling would provide ideals that predict typicality to a higher extent than has been observed so far.

Experiment 2

Because the previous study was correlational, referring to ideals, central tendency, and frequency of instantiation as determinants of graded structure has been somewhat unjustified. If these factors are indeed determinants of graded structure, they must be shown to be causes of it.

Actually, Rosch and Mervis (1975) and Rosch, Simpson, et al. (1976) found in experimental studies with artificial categories that central tendency (as measured by family resemblance) causally determined graded structure. Although Rosch, Simpson, et al. (1976) also reported that frequency did not causally determine graded structure, their experimental design was not capable of detecting simultaneous effects of central tendency and frequency. However when central tendency and frequency are orthogonally manipulated in an experimental setting such that they can be detected simultaneously, both causally determine typicality (Barsalou, 1981, Experiment 3, which is the same as Experiment 2a in Barsalou, 1984). So far, no experiments have shown that ideals are causal determinants of graded structure. One purpose of this next experiment, therefore, was to examine whether ideals also causally determine typicality.

A second purpose of this experiment was to examine whether the determinants of a category's graded structure depend on the context in which the category is processed. Instead of there being a fixed determinant responsible for a category's graded structure on all occasions, different contexts may engender the use of different determinants such that the category's graded structure changes. For example, some contexts may engender the use of central tendency, whereas others may engender the use of ideals.

Subjects acquired two artificial categories whose exemplars each contained a person's last name (e.g., *Davis*) associated with five things they like to do in their spare time (e.g., *go horseback riding, jog daily, collect antiques, cook Chinese food, meditate*). Two variables structured each category. First, each category varied along a *defining dimension.* All members of one category *jogged,* and all members of the other category *read the newspaper.* Within each category, exemplars varied in the extent to which they jogged or read the newspaper (i.e., their *amount* on the defining dimension). Exemplars either jogged (read the newspaper) *daily, weekly,* or *monthly.* The second way in which exemplars varied was in how similar they were to the central tendency of their respective category, having either a high, medium, or low number of their category's characteristic activities. Categories were constructed such that amount on the defining dimension and similarity to central tendency were orthogonally manipulated.

The central manipulation in this experiment was how subjects were induced to perceive the defining dimensions. Subjects in the *related dimension* condition were told that exemplars who *jogged* belonged to the category of *physical education teachers* and that exemplars who *read the newspaper* belonged to the category of *current events teachers.* Because most people believe that physical education teachers ideally should be physically fit and that current events teachers ideally should be well-read, it was expected that these beliefs about ideals would determine graded structure. The more often a physical education teacher jogged (e.g., daily versus weekly versus monthly), the more typical that exemplar should be of its category; and the more often a current events teacher read the newspaper, the more typical that exemplar should be of its category. For each category, people should use an ideal obtained from stereotypes to determine graded structure. Because subjects may focus on how well exemplars approximate these ideals, they may not abstract each category's central tendency (although ideals and central tendency could simultaneously determine graded structure because they were manipulated orthogonally).

Subjects in the *unrelated dimension* condition were told that the category defined by one dimension contained *Q programmers* (i.e., people who program in the Q programming language) and that the category defined by the other dimension contained *Z programmers.* Because subjects probably do not have any beliefs about what constitutes ideal Q and Z programmers, it was unlikely that ideals would determine typicality for these subjects. Increasing (or decreasing) values on the defining dimensions should not determine typicality because subjects have no reason to assume that one value on a defining dimension is any more ideal than the others (e.g., there is no obvious ideal value for how often a Q programmer should jog). In addition, the spare time activities that constituted each exemplar were chosen so as not to be meaningfully related to computer programming in general. Consequently exemplars also did not vary in how ideal they were of all computer programmers. Because subjects could not use ideals to structure the categories, it was expected that they would instead abstract the central tendencies of these categories and use them to determine graded structure.

One other between subjects manipulation was included. Half the subjects performed a *relevant processing* task, and half performed an *irrelevant processing* task. Relevant processing required that subjects learn to discriminate members of the two categories from one another, where the defining dimensions and central tendencies of the category provided information relevant to performing discrimination. Irrelevant processing required that subjects decide which of two members of the *same* category were better suited to achieve some peripheral goal. The defining dimensions and central tendencies were irrelevant to these decisions, which were always based on other activities of the exemplars. It was expected that irrelevant processing would attenuate the extent to which ideals determined typicality in the related dimension condition and the extent to which central tendency determined typicality in the unrelated dimension condition. Because subjects need not focus on defining dimensions and central tendencies to perform the irrelevant processing task, they should not abstract this information.

Method

Materials. Two categories of nine exemplars were formed. Each exemplar had a common surname (e.g., *Davis, Wilson, Adams*) associated with five activities they do in their spare time (e.g., *dance, renovate houses, write poetry, go to movies, read the newspaper daily*). Each category had a defining dimension: All members of one category *jogged,* and all members of the other category *read the newspaper.* Three members of each category performed their defining activity *daily,* three performed it *weekly,* and three performed it *monthly.* Each category possessed three characteristic activities, each of which occurred for six category members and for no nonmembers. These activities were not correlated with one another or with values on the defining dimensions. Exemplars varied in how similar they were to the central tendency of their category, with an exemplar's similarity to its category's central tendency being the number of characteristic activities it possessed. Three exemplars in each category possessed all three characteristic activities, three possessed two, and three possessed one. For each level of similarity to central tendency, one exemplar had a *daily* value on the defining dimension, one had a *weekly* value, and the other had a *monthly* value (i.e., similarity to central tendency and amount on the defining dimension were orthogonally manipulated). Except for the defining and characteristic activities, no activity occurred more than once in a category. Finally, one activity for each exemplar was an *irrelevant processing* activity (to be described in a moment). There were nine such activities, each occurring twice, once for one exemplar in each category. Only these activities occurred in both categories. Two versions of the materials were formed in which surnames and spare time activities were randomly assigned to the stimulus structure, within the constraints of the design.

In the related dimension condition, the category defined by *jogging* was called *physical education teachers*, and the category defined by *reading the newspaper* was called *current events teachers*. In the unrelated dimension condition, the categories were called *Q programmers* and *Z programmers*.

Acquisition procedure. Subjects were asked to imagine they were being trained at a personnel agency to find teachers for high schools and were told that people's spare time activities predict the courses they are good at teaching. Subjects in the relevant processing condition were told they would learn to discriminate one category from the other on the basis of their spare time activities. Subjects in the irrelevant processing condition were told their personnel agency already had pools for each kind of teacher and that their job was to find someone from a particular pool who would be good at teaching a *special interest course*.

On each acquisition trial, subjects in the relevant processing condition received one description from each category, followed by a request to choose the one that belonged to a particular category (e.g., "Choose the better teacher of programming language Q"). Subjects in the irrelevant processing condition received two descriptions from the *same* category, followed by a request to choose the one better able to teach a special interest course (e.g., "Choose the Q programmer better able to teach *how to cook Indian food*"). The irrelevant processing activity of the correct exemplar was always related to the topic of the special interest course (e.g., *invests in real estate* was the irrelevant processing activity of an exemplar who would be good at teaching a special interest course on *how to invest in gold*). As described earlier, the irrelevant processing activities were never defining or characteristic of their category. Because an exemplar was the correct choice on more than one trial, it was always chosen to teach a different special interest course on each occasion. So if *invests in real estate* were an exemplar's irrelevant processing activity, that exemplar might be chosen to teach *how to invest in stocks* on one trial, *how to invest in gold* on a second trial, and *how to invest in bonds* on a third trial. Each of the 18 exemplars served as the target exemplar on three trials in all conditions.

Subjects performed the experiment in individual booths. On the table before them was a stack of 54 cards (one for each trial) and a response sheet. A tape recording paced all aspects of the experiment. At the start of each trial, the recording stated the trial number. Subjects read the request on the top card, decided which exemplar was correct, and wrote that person's initial on the response sheet next to the trial number (i.e., the name for each exemplar had a unique initial). Twelve seconds after the trial number was stated, the recording provided the correct name. For the next 7 s, subjects studied the correct exemplar's spare time activities. At the end of this period, the recording instructed subjects to turn the card over and, 3 s later, stated a spare time activity. Subjects had 3 s to decide if this was one possessed by the correct exemplar and circled *yes* or *no* on their response sheet. The choice of the probe was random with the constraints that: (a) on true trials, the probe was only true of the target, and (b) on false trials, it was not true of either the target or the contrast stimulus. The probe task served to insure that subjects encoded every activity for each exemplar.

For relevant processing, each category was correct on 50% of the trials, and for irrelevant processing, each category was used on 50% of the trials. Orthogonally varied with the target category were whether the target was on the left or the right, and whether the answer to the probe task was true or false. The contrast stimuli (i.e., the nontarget stimuli for each trial) were randomly chosen with the constraints that each exemplar occurred as a constrast three times, and no exemplar was the constrast for a given target more than once. The 54 trials were randomly ordered for each of two versions with the constraint that an exemplar was never the target on two or more consecutive trials. Prior to the 54 trials, all subjects briefly studied a page showing all 18 exemplars blocked by category to develop an initial impression of the categories.

Test procedure. Following acquisition, subjects performed three additional tasks. First, they provided paced typicality ratings. Subjects were asked to imagine they were teaching another employee about the two populations of teachers. As they considered each exemplar, they were to tell the other person how good an example it was of its category on the basis of experience acquired during the learning phase of the experiment. Each subject received the nine exemplars of each category blocked and in a different random order. Every 12 s, a recording instructed subjects to read a new card and rate the exemplar on it for "how good an example it was of its category." The card for each exemplar contained both its surname and its five activities. Subjects used a scale from 1 to 7, where 1 meant the exemplar was a *poor example*, and 7 meant it was *excellent*. Only one exemplar was ever visible at a time to a subject.

Subjects' second task was to *rank* exemplars by typicality. Subjects were able to observe all 18 exemplars at once and had as much time as they needed.

Subjects' third task was to fill out a questionnaire. They described their strategies for the typicality tasks and the activities they thought were characteristic of each category.

Subjects. Fifty-four Stanford University students participated for either credit plus pay or for pay only. Subjects participated in groups of one to four, and each session lasted about 1 hr and 15 min. Six subjects' data were discarded because of equipment failure. Six subjects were randomly assigned to each of the eight between-subjects conditions created by crossing the dimension relatedness, processing type, and version factors.

Results

Typicality. Because the same pattern of significant effects occurred for the ratings and rankings, only the rating data are reported. An analysis of variance (ANOVA) was performed on the ratings in which the factors were ideals (i.e., amount on the defining dimensions), central tendency (i.e., number of characteristic properties), dimension relatedness, processing

Table 4
Average Exemplar Goodness Ratings From Experiment 2

Condition	Central tendency Low	Central tendency Medium	Central tendency High	Ideals Low	Ideals Medium	Ideals High
Unrelated dimension						
Relevant processing	4.24	4.99	5.44	4.96	4.79	4.92
Irrelevant processing	4.44	4.71	5.11	4.88	4.51	4.88
Average	4.34	4.85	5.28	4.92	4.65	4.90
Related dimension						
Relevant processing	4.58	4.51	4.72	3.89	4.47	5.46
Irrelevant processing	4.06	4.08	4.76	3.81	4.03	5.07
Average	4.32	4.30	4.74	3.85	4.25	5.26

Note. The scale is from 1 to 7, where 1 is for *poor examples*, and 7 is for *excellent examples*.

type, and version. The relevant means are shown in Table 4.

Overall, ideals and central tendency both affected typicality, $F(2, 80) = 20.58$, $MS_e = 1.03$, $p < .001$ and $F(2, 80) = 20.16$, $MS_e = .85$, $p < .001$, respectively. However each factor interacted with dimension relatedness. Ideals interacted with dimension relatedness, $F(2, 80) = 18.01$, $MS_e = 1.03$, $p < .001$, having a significant effect in the related dimension condition, $F(1, 80) = 69.49$, $MS_e = 1.03$, $p < .001$, but having no effect in the unrelated dimension condition, $F(1, 80) = .01$, $MS_e = 1.03$, $p > .25$. This indicates that ideals determined typicality only when the defining dimensions were related to subjects' stereotypes for the categories. Central tendency interacted with dimension relatedness in the converse manner, $F(2, 80) = 3.83$, $MS_e = .85$, $p < .025$, having a larger effect in the unrelated dimension condition, $F(1, 80) = 37.42$, $MS_e = .85$, $p < .001$, than in the related dimension condition, $F(1, 80) = 7.47$, $MS_e = .85$, $p < .01$. This indicates that similarity to central tendency played a stronger role in determining typicality when subjects did not have ideals for the categories. The significant effect of central tendency in the related dimension condition was entirely attributable to the irrelevant processing condition, $F(1, 80) = 10.38$, $MS_e = .85$, $p < .01$. The fact that there was no effect of central tendency in the relevant processing condition, $F(1, 80) = .42$, $MS_e = .85$, $p > .25$, demonstrates that central tendency may play no role in determining graded structure under some conditions.

Taken together, these results demonstrate that (a) ideals, as well as central tendency, can causally determine graded structure, (b) the determinants of a particular category's graded structure can vary with context, and (c) multiple determinants can simultaneously determine graded structure in a particular category (e.g., for irrelevant processing subjects in the related dimension condition).

Irrelevant processing did not attenuate the effects of ideals and central tendency as expected. It it not clear why having subjects focus on irrelevant information did not reduce the effects of these variables on typicality. Irrelevant processing subjects may have had sufficient exposure to exemplars during acquisition to abstract central tendencies; and they may have later used ideals at testing when they appeared relevant to the categories.

There were unpredicted interactions between central tendency and ideals, $F(4, 160) = 6.44$, $MS_e = .44$, $p < .001$, and between processing type, dimension relatedness, and central tendency, $F(2, 80) = 3.27$, $MS_e = .85$, $p < .05$. These may have resulted in some way from two further significant interactions involving version.

Acquisition performance. Performance on both the choice and probe tasks was excellent. For each, an analysis of variance (ANOVA) was performed on subjects' error percentages, transformed as suggested by Winer (1971). The average error rate for the choice task was .038. The only effect was a marginally significant interaction between central tendency and dimension relatedness, $F(2, 80) = 2.81$, $MS_e =$

.01, .10 > p > .05. Subjects in the unrelated dimension condition made fewer errors as similarity to central tendency increased, but subjects in the related dimension condition did not. This further supports the conclusion that central tendency was more important for subjects in the unrelated dimension condition than for subjects in the related dimension condition. For the probe task, the average error rate was .064, and there were no significant effects.

Questionnaire. Perhaps because of the ease with which people ascribe traits to behaviors, 36 subjects mentioned assimilating the categories to personality stereotypes. When asked to describe their strategies or the characteristic spare-time activities of each category, these subjects described a category in the context of a stereotype. In general, the stereotypes that subjects adopted stemmed from the defining dimensions. Categories defined by *jogs* were assimilated to physically oriented stereotypes, whereas categories defined by *reads the newspaper* were assimilated to culturally and socially oriented stereotypes.

There was a substantial difference between the reports of subjects in the related and unrelated dimension conditions. Of the 24 subjects in the unrelated dimension condition, 14 reported using central tendency when judging typicality, and 1 reported using ideals. Of the 24 subjects in the related dimension condition, 2 reported using central tendency, and 19 reported using ideals. Subjects were not probed directly about specific strategies but volunteered them in response to an open-ended question about the strategies they used. These strategies were distributed evenly over the relevant and irrelevant processing conditions and are clearly consistent with the rating data in Table 4.

Discussion

These results demonstrate that ideals as well as central tendency can causally determine graded structure. When subjects were aware of ideals for the categories, an exemplar's similarity to the relevant ideal determined its typicality. These results even go so far as to suggest that ideals may be a more important determinant of graded structure than is central tendency. When subjects were aware of ideals for the categories, central tendency either had no effect (in the relevant processing condition) or had a smaller effect (in the irrelevant processing condition).

These results also show that the determinants of a particular category's graded structure depend on the context in which the category is perceived. In some contexts, ideals may determine a category's graded structure; whereas in others, central tendency may determine a different graded structure. In addition, ideals and central tendency can simultaneously determine a category's graded structure, as shown by the results for irrelevant processing subjects in the related dimension condition. This is consistent with the results from Experiment 1 in which ideals, central tendency, and frequency of instantiation simultaneously determined graded structure in some categories.

General Discussion

These studies indicate that graded structure is a complex and dynamic phenomenon. It is not the case that a single determinant, such as central tendency, is responsible for the graded structure of all categories. Instead at least two other factors—ideals and frequency of instantiation—also play major roles in determining graded structure. It is also not the case that the graded structure of a category remains constant across contexts. As shown by Experiment 2, ideals may determine a category's graded structure in one context, and central tendency may determine a different graded structure in another.

Findings further demonstrating such flexibility have been reported by Roth and Shoben (1983) and Barsalou and Sewell (1984). Roth and Shoben (1983) varied a category term's linguistic context and found that its typical referents varied across contexts. *Animals,* for example, had different graded structures when it appeared in "Stacy volunteered to milk the *animal* whenever she visited the farm" and "Fran pleaded with her father to let her ride the *animal.*" Whereas *cow* and *goat* were typical in the first context, *horse* and *mule* were typical in the second. Barsalou and Sewell (1984) found that the graded structures of both common taxonomic and goal-derived categories shifted substantially when people took various points of view while judging typicality.

For example, American undergraduates believe that *robin* and *eagle* are typical exemplars of *birds* from the point of view of the average American citizen, but believe that *swan* and *peacock* are typical from the point of view of the average Chinese citizen. Although both studies demonstrate flexibility in graded structure, neither examines the determinants responsible for such shifts. Consequently it is not clear whether such flexibility results from changes in ideals, central tendency, frequency of instantiation, or other determinants.

Most important, such flexibility indicates that graded structure does not reflect some invariant property of categories—there do not appear to be invariant structures that underlie categories. Instead such flexibility suggests that people's perception and structuring of categories is a highly dynamic and context-dependent process.

One way to explain dynamic graded structures begins by assuming that people have a highly creative ability to construct concepts, where concept refers to the information that represents a category (e.g., a prototype). As suggested by Barsalou (in press), people may not retrieve the same concept from long-term memory everytime they deal with a particular category. Instead they may construct a diverse variety of concepts in working memory to represent a particular category across different situations such that the concept used to represent a category is rarely, if ever, the same. According to this view, long-term memory does not contain invariant concepts. Instead it contains generic and episodic information from which concepts are constructed. Although some information about a category may always be incorporated into a concept (context-independent information), other information may only be incorporated in relevant contexts (context-dependent information). As suggested by Barsalou (1982), context-independent and context-dependent information provide stability and instability, respectively, for the concepts that represent a particular category. Barsalou (in press) outlines a general theory of how concepts are constructed in working memory from information in long-term memory.

Dynamic graded structures may simply reflect this highly creative ability to construct concepts. Assuming that an exemplar's typicality increases as it becomes more similar to the concept for its category (Barsalou, 1983; Hampton, 1979; Smith et al., 1974), graded structure will change with different concepts. As the information comprising a concept changes, different exemplars will be highly similar to the concept such that the ordering of exemplars by typicality changes. For example, as a category is perceived in the context of different goals, different ideals may be incorporated into its concept. If different exemplars best approximate these ideals, then different exemplars will be typical in the respective contexts. Similarly the central tendency of a category may vary across contexts, with different exemplars being highly similar to these different central tendencies. For example, the central tendencies of *animals* on the dimensions of *size* and *ferocity* may be higher when viewing the category from a forest ranger's point of view than from a pet store owner's point of view. As a result, different exemplars will be typical from these two perspectives. In general, graded structure may largely reflect people's current concept of a category, and to the extent this concept changes, their graded structure will change.

It appears, however, that another factor is also important for graded structure. As found in Experiment 1, frequency of instantiation is an important predictor of typicality. One explanation of this finding is that exemplars seem to occur frequently in a category because they possess properties that occur often across the category's exemplars (Rosch, 1974). This property frequency explanation of graded structure, however, is equivalent to the central tendency explanation; both assume typicality increases as exemplars become more similar to the characteristic properties of their category. But if frequency of instantiation ratings were based on central tendency, then they should not predict typicality after central tendency is partialed out. As found in Experiment 1, however, partialing out central tendency only slightly diminished the substantial predictive power of frequency of instantiation.

Because frequency of instantiation's role in graded structure does not seem to depend on central tendency information in the concept for a category, another explanation is that it depends on frequency information stored with exemplars. More specifically, information may be stored with exemplars that in some way reflects how often they have been encountered in the context of their category. When people estimate an exemplar's typicality, such information may be accessed and be incorporated into judgments of typicality. To the extent an exemplar has a high frequency of instantiation in a category, it is perceived as typical.

This use of frequency further complicates our understanding of people's sensitivity to frequency. It is well known that people acquire information without much effort that allows them to estimate how often they have encountered particular items (see the reviews by Hasher & Zacks, 1979; Hintzman, 1976). This kind of frequency is analogous to familiarity in Experiment 1, which reflects an exemplar's category-independent frequency. However people also appear sensitive to other kinds of frequency. First, Alba, Chromiak, Hasher, and Attig (1980) and Barsalou and Ross (in press) have shown that people acquire information without much effort that allows them to estimate how often a superordinate category has been instantiated by different exemplars (superordinate frequency). Second, Hintzman and Block (1971) have shown that people acquire information without much effort that allows them to estimate how often an item has occurred in each of several lists. This latter kind of frequency is analogous to frequency of instantiation, which is how often an exemplar has occurred in each of several categories. Similar to how familiarity, superordinate frequency, and list frequency are acquired, frequency of instantiation may also be acquired in the course of everyday experience without much effort. Such information may later become available when exemplars are accessed and may affect processing during tasks such as typicality judgments.

In general, it appears that there are many different kinds of information available for people to use when judging typicality. People can incorporate various ideals and central tendencies into category concepts when making typicality judgments, and they can also incorporate information reflecting frequency of instantiation that may be stored with exemplars. Given that more than one of these factors often predicted graded structure in Experiment 1 (see Table 3), it appears that people may often simultaneously incorporate several kinds of information into their judgments of typicality. And given that these factors varied widely in the extent to which they predicted graded structure, it also appears that people have the ability to differentially weight these sources of information.

But what determines the information used to structure a particular category on a particular occasion? What conditions determine whether ideals, central tendency, frequency of instantiation, or other factors are used to generate graded structure? And what determines the relative weighting of the information chosen?

As suggested earlier, ideals may become important when categories are used in the context of achieving goals. Little if anything, however, is known about how goals and ideals are actually related, and much remains to be learned about how goals generate ideals. Ideals may also be important in other ways. For example, Lakoff (in press) suggests that people may at times use "paragons" to represent categories (e.g., *Babe Ruth* with respect to *baseball players*). Certain exemplars may become paragons because they closely approximate ideals associated with their respective categories. In general, little if any attention has been given to the role of ideals in human knowledge. Instead, because most theories of knowledge are normative in spirit, the representations they use only include central tendency and frequency information (e.g., as in prototype, exemplar, and schema theories). Given the importance of goals in human behavior, however, ideals may also be central to knowledge. If so, then future theories of knowledge should incorporate them.

Although central tendency has received more attention than ideals, much remains to be learned about it as well. For example, it is not clear why central tendency determines graded structure in common taxonomic categories but not in goal-derived categories (as found in Experiment 1). Two possibilities suggested earlier were as follows. First, central tendency may be important for common taxonomic categories because these categories provide information about the structure of the environment. By including central tendency information in the concepts for these categories, people establish representative information about the kinds of entities in the environment. Second, central tendency may be important for common taxonomic categories because it maximally facilitates classification. By representing these categories with their central tendencies, people minimize the average distance from exemplars to classification standards. Central tendency may not be important for goal-derived categories because these categories are not used to provide information about the structure of the environment and because these categories are not regularly used for classification.

Two other factors may also result in central tendency becoming important to graded structure. First, members within common taxonomic categories generally bear a strong perceptual similarity to one another. As noted by Rosch, Mervis, et al. (1976), members within basic level categories share a common shape; and as noted by B. Tversky and Hemenway (1984), members within basic level categories share the same parts in the same configuration. Perceptual similarity also appears to occur in subordinate and superordinate common taxonomic categories, even though it does not seem to be as compelling as in basic level categories. In contrast, goal-derived categories often possess little if any perceptual similarity (e.g., *things to take on a camping trip*). For some reason, people may be more apt to acquire the central tendencies of categories that contain high perceptual similarity.

Another reason central tendency may only be acquired for common taxonomic categories is because people may be much more familiar with them than with goal-derived categories. If much exposure to a category is necessary for its central tendency to be abstracted, then central tendency may only be acquired for common taxonomic categories. If this is correct, then sufficient experience with a goal-derived category should result in the abstraction and use of its central tendency.[8]

In conclusion, it appears that the human conceptual ability is extremely dynamic. As shown by Experiment 1, people incorporate various kinds of information into concepts (e.g., central tendency and ideals), such that different kinds of information determine typicality in different categories. Moreover people construct different concepts for the same category in different contexts, tailoring concepts to represent the demands of current situations (Experiment 2; Barsalou & Sewell, 1984; Roth & Shoben, 1983). Not only do people represent well-established categories in a dynamic and context-dependent manner, they also construct new concepts for new categories that serve new goals (i.e., ad hoc categories; Barsalou, 1983). As further discussed in Barsalou (in press), these observations suggest that a fundamental characteristic of the human cognitive system is its ability to construct context-dependent representations in working memory from a large knowledge base in long-term memory to meet the constraints of specific situations.

---

[8] I am grateful to Edward J. Shoben for suggesting this possibility.

References

Alba, J. W., Chromiak, W., Hasher, L., & Attig, M. S. (1980). Automatic encoding of category size information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 6,* 370–378.

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory.* Monterey, CA: Brooks/Cole.

Ashcraft, M. H. (1978). Property norms for typical and atypical items from 17 categories: A description and discussion. *Memory & Cognition, 6,* 227–232.

Barsalou, L. W. (1981). *Determinants of graded structure in categories.* Unpublished doctoral dissertation, Stanford University.

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10,* 82–93.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11,* 211–227.

Barsalou, L. W. (1984). *Determinants of graded structure in categories* (Tech. Rep. No. 4). Atlanta, GA: Emory University, Emory Cognition Project.

Barsalou, L. W. (in press). The instability of graded structure: Implications for the nature of concepts. To appear in U. Neisser (Ed.), *Concepts reconsidered: The ecological and intellectual bases of categories.* Cambridge: Cambridge University Press.

Barsalou, L. W., & Ross, B. H. (in press). The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Barsalou, L. W., & Sewell, D. R. (1984). *Constructing representations of categories from different points of view* (Tech. Rep. No. 2). Atlanta, GA: Emory University, Emory Cognition Project.

Barsalou, L. W., & Sewell, D. R. (1985). *The relation between determinants of category verification and determinants of graded structure in common taxonomic and goal-derived categories.* Manuscript in preparation.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monographs, 80,* 3, Pt 2.

Glass, A. L., & Meany, P. J. (1978). Evidence for two kinds of low-typical instances in a categorization task. *Memory & Cognition, 6,* 622–628.

Guilford, J. P., & Fruchter, B. (1973). *Fundamental statistics in psychology and education.* New York: McGraw-Hill.

Hampton, J. A. (1979). Polymorphous concepts in semantic memory. *Journal of Verbal Learning and Verbal Behavior, 18,* 441–461.

Hampton, J. A. (1981). An investigation of the nature of abstract concepts. *Memory & Cognition, 9,* 149–156.

Hampton, J. A., & Gardiner, M. M. (1983). Measures of internal category structure: A correlational analysis of normative data. *British Journal of Psychology, 74,* 491–516.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Hayes, W. L. (1973). *Statistics for the social sciences.* New York: Holt, Rinehart, and Winston.

Hintzman, D. L. (1976). Repetition and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 10). New York: Academic Press.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88,* 297–306.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: Making the connection. *Cognitive Psychology, 16,* 243–275.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Lakoff, G. (in press). *Women, fire and dangerous things: What categories tell us about the nature of thought.* Chicago: University of Chicago Press.

Malt, B. C., & Smith, E. E. (1982). The role of familiarity in determining typicality. *Memory & Cognition, 10,* 69–75.

McCloskey, M. (1980). The stimulus familiarity problem in semantic memory research. *Journal of Verbal Learning and Verbal Behavior, 19,* 485–502.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology, 35,* 113–138.

Mervis, C. B., Catlin, J., & Rosch, E. (1976). Relationships among goodness-of-example, category norms, and word frequency. *Bulletin of the Psychonomic Society, 7,* 283–294.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32,* 89–115.

Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21,* 1–20.

Rips, L. J., Shoben, E. J., & Smith, E. E. (1973). Semantic distance and the verification of semantic relations. *Journal of Verbal Learning and Verbal Behavior, 12,* 1–20.

Rosch, E. H. (1973). On the internal structure of perceptual and semantic categories. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language.* New York: Academic Press.

Rosch, E. H. (1974). Universals and cultural specifics in human categorization. In R. Breslin, W. Lonner, & S. Bochner (Eds.), *Cross-cultural perspectives on learning.* London: Sage Press.

Rosch, E. H. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104,* 192–233.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Rosch, E. H., Simpson, C., & Miller, R. S. (1976). Structural bases of typicality effects. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 491–502.

Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology, 15,* 346–378.

Smith, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 6). Potomac, MD: Erlbaum.

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review, 81,* 214–241.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113,* 169–193.

Winer, B. J. (1971). *Statistical principles in experimental design.* New York: McGraw-Hill.

(*Appendix follows on next page*)

Appendix

Output dominance (OD) and the average scores for exemplar goodness (EG), central tendency (CT), frequency of instantiation (FOI), ideals (I), and familiarity (FAM) are shown below for the exemplars of the nine goal-derived and the nine common taxonomic categories in Experiment 1. The measures are defined in the method of Experiment 1, and the exemplars are ordered within each category by their values for exemplar goodness.

Birthday Presents

Exemplar	OD	EG	CT	FOI	I	FAM
Clothing	17	7.9	3.029	7.7	6.9	8.2
Party	2	7.3	3.616	7.2	6.4	7.4
Jewelry	10	7.2	4.522	5.3	5.8	7.7
Dinner	2	6.9	3.254	6.5	5.6	8.6
Watch	4	6.6	4.246	4.6	6.2	7.2
Cake	4	6.3	2.928	7.0	4.5	7.8
Card	5	6.3	3.268	7.5	4.7	6.3
Flowers	3	6.3	3.464	5.4	5.9	6.7
Perfume	4	6.0	3.565	5.0	5.7	6.4
Money	12	5.9	3.464	7.1	7.1	7.9
Record	7	5.8	3.565	5.4	5.2	7.6
Camera	2	5.6	3.819	4.2	6.1	5.6
Book	6	5.5	3.247	6.4	5.4	9.0
Gold	2	5.3	3.304	3.6	7.4	7.3
Sports equipment	2	5.3	3.420	6.0	6.7	4.5
Booze	2	5.1	3.290	4.5	5.8	5.9
Tie	2	4.9	3.247	3.0	3.8	4.0
Toy	5	4.9	3.341	6.1	4.6	5.0
Game	3	4.6	3.899	5.6	4.5	6.2
Nic nac	2	4.2	3.587	5.3	3.6	3.8
Car	5	4.0	2.638	2.6	8.8	8.6
Appliance	2	3.8	3.087	3.0	6.8	5.8
Candy	2	3.1	3.514	3.7	3.4	8.3
Art	3	2.8	3.660	2.4	4.8	6.0

Camping Equipment

Exemplar	OD	EG	CT	FOI	I	FAM
Sleeping bag	26	8.8	3.140	8.5	7.1	4.7
Tent	34	7.7	3.737	5.8	4.3	3.4
Flashlight	4	7.5	4.009	7.0	6.9	5.1
Matches	5	7.4	3.263	7.9	7.6	6.0
Backpack	10	7.0	3.553	7.4	5.3	3.9
Food	3	6.9	3.237	8.5	8.6	9.0
Lantern	7	6.7	4.158	6.0	5.1	2.5
Knife	2	6.2	3.719	7.8	8.3	7.0
Canteen	4	6.1	4.272	7.0	6.6	3.5
Camping stove	14	6.0	4.974	6.5	3.1	2.2
Ground cover	2	5.8	3.105	5.3	4.2	3.5
Utensils	2	5.6	4.947	6.0	3.4	6.9

Camping Equipment (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Pots	3	5.2	4.421	4.7	2.7	6.7
Fuel	2	4.8	3.167	5.8	4.5	5.7
Can opener	2	4.4	4.106	4.8	4.3	5.6
Pans	2	4.3	4.263	5.3	2.8	6.6
Camping vehicle	2	4.1	3.404	4.6	1.9	2.8
Camping shoes	3	3.5	2.316	4.9	5.3	3.4
Dried food	2	3.5	2.561	4.1	5.1	4.5
Equipment to keep flies away	2	2.6	2.526	3.4	1.8	3.1

Transportation for Getting From San Francisco to New York

Exemplar	OD	EG	CT	FOI	I	FAM
Airplane	36	9.0	3.500	8.4	9.0	7.6
Car	31	6.6	4.792	7.5	7.1	8.6
Train	26	6.2	3.958	4.7	6.0	4.2
Bus	21	4.3	4.479	4.4	5.5	5.4
Motorcycle	3	3.3	4.250	4.0	5.4	3.5
Bike	9	2.5	3.833	2.1	2.8	5.1
Hitch hike	8	2.1	3.875	3.2	3.3	2.0
Boat	12	2.0	2.688	2.0	2.3	5.1
Walking	3	1.2	3.333	1.2	1.0	7.4

Things to Do for Weekend Entertainment

Exemplar	OD	EG	CT	FOI	I	FAM
Parties	19	8.3	4.333	7.4	7.6	7.7
Skiing	2	8.0	3.720	5.1	7.1	3.8
Drinking	3	7.1	3.894	5.7	7.6	6.3
Go the beach	12	7.1	4.576	7.8	8.1	5.5
Movies	19	6.9	3.901	7.4	6.0	6.9
Go to the city	2	6.7	4.076	5.5	6.0	6.2
Play sports	5	6.5	4.500	8.2	7.3	5.9
Camping	2	6.4	4.015	4.5	6.9	3.8
Concerts	3	6.4	4.060	6.0	6.6	6.9
Sports events	2	6.2	4.205	7.1	7.4	5.8
Dinner at restaurants	5	5.8	4.098	5.5	6.4	7.0
Sex	4	5.7	3.667	5.5	7.8	7.5
Dancing	7	5.4	4.273	6.2	6.6	6.5
Picnics	5	4.9	4.621	4.7	7.1	5.7
Softball	4	4.7	3.712	5.6	5.9	4.3
Eating	3	4.5	4.273	6.4	5.8	8.9
Swimming	2	4.5	3.826	6.5	5.3	5.9
Walks	3	4.1	3.523	4.6	4.4	7.4
Sleep	3	4.0	2.697	7.1	6.8	9.0
Drives	2	3.8	3.371	4.6	6.0	7.1
Tennis	3	3.8	3.705	5.1	6.2	6.3

Things to Do for Weekend Entertainment (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Watch TV	7	3.3	2.993	6.7	5.7	8.0
Work	2	2.0	1.795	3.9	1.9	8.7

Picnic Activities

Exemplar	OD	EG	CT	FOI	I	FAM
Bar-be-queing	2	8.5	3.563	8.0	7.0	4.8
Eating	22	8.2	3.917	8.4	7.4	8.9
Laying in the grass	2	7.1	4.281	7.7	6.1	5.3
Softball	9	7.1	4.531	7.0	6.7	4.3
Talking	2	7.1	4.250	8.3	7.3	8.5
Frisbee	20	7.0	4.406	7.1	7.4	3.5
Drinking	5	6.6	3.594	7.0	7.1	6.3
Volleyball	9	6.5	4.281	4.8	6.1	5.3
Play games	3	6.4	5.240	7.5	6.1	6.5
Hiking	2	6.2	3.573	6.3	7.3	3.8
Football	6	5.8	4.261	6.2	6.2	5.8
Dealing with insects	3	5.7	2.750	4.2	2.2	4.2
Suntanning	3	5.7	4.250	7.0	6.8	6.6
Swimming	9	5.5	4.240	6.9	6.0	5.9
Baseball	4	5.3	4.313	5.7	5.2	4.7
Sex	3	3.9	3.875	2.3	7.6	7.5
Sleeping	2	2.5	3.010	2.5	4.0	9.0

Personality Characteristics in Others That Prevent You From Being Friends With Them

Exemplar	OD	EG	CT	FOI	I	FAM
Asshole	2	8.2	5.739	8.2	8.1	6.3
Phony	2	7.9	4.333	7.1	6.5	4.9
Obnoxious	2	7.8	5.804	5.8	7.0	6.0
Bitchy	3	7.4	5.580	6.7	7.7	7.1
Unfriendly	4	6.7	5.218	6.6	7.7	4.9
Hostile	2	6.5	5.529	5.8	7.0	4.7
Untruthful	2	6.5	3.290	6.3	7.1	5.3
Mean	5	6.4	5.102	6.2	6.7	4.4
Egotistical	2	6.3	5.486	5.7	7.4	6.0
Snobish	5	6.1	5.297	7.6	6.0	5.4
Selfish	4	6.0	5.406	5.6	6.0	5.3
Complaining	2	5.9	4.833	6.7	5.8	7.5
Cocky	2	5.7	4.848	5.1	5.9	5.4
Humorless	3	5.7	4.181	5.7	5.3	2.3
Bad temper	2	5.6	4.971	6.3	5.5	5.5
Foul-mouthed	2	5.4	4.406	5.3	5.2	5.3
Jealous	3	5.4	4.297	4.9	4.7	5.9
Mentally ill	3	5.3	2.326	3.7	5.7	5.1
Showy	2	4.8	4.536	4.4	5.1	4.2
Narrow-minded	3	4.7	4.355	6.1	5.2	5.8
Aggressive	5	4.4	5.087	5.0	4.1	6.6
Loud	3	4.4	5.051	4.8	3.9	7.0

Personality Characteristics in Others That Prevent You From Being Friends With Them (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Sarcastic	2	4.3	5.094	3.7	4.7	6.5
Too quiet	3	3.5	2.275	3.7	2.6	4.2

Foods Not to Eat on a Diet

Exemplar	OD	EG	CT	FOI	I	FAM
Chocolate	5	8.2	4.444	7.0	6.5	7.6
Ice cream sundaes	3	8.2	4.659	7.2	8.1	7.1
Candy	16	8.0	4.476	6.7	7.3	8.3
Pie	2	8.0	4.738	7.0	7.7	5.3
Sugar	2	8.0	4.540	7.0	7.0	6.4
Cookies	4	7.8	5.079	6.9	6.8	7.1
Cake	13	7.6	5.135	6.9	7.3	7.8
Pastry	2	7.6	5.468	6.8	7.5	5.7
Ice cream	19	6.9	4.683	6.4	7.4	8.3
French fries	4	6.4	4.294	6.4	7.4	6.6
Pasta	2	6.1	4.849	5.9	5.9	5.5
Potato chips	2	5.9	4.326	6.7	6.6	6.3
Spaghetti	2	5.9	4.222	6.1	5.3	6.7
Pizza	3	5.7	4.318	6.5	5.5	7.1
Butter	2	5.4	3.897	6.4	6.3	6.6
Starches	2	5.4	5.691	6.7	6.8	4.0
Bread	8	4.9	4.913	3.8	5.5	6.7
Hamburgers	3	4.5	3.539	6.4	4.8	6.5
Potatoes	10	3.8	4.397	4.9	6.1	5.9
Salt	3	3.3	2.675	5.8	4.5	6.5
Cheese	2	2.9	3.429	3.9	5.4	7.4
Meat	3	2.2	3.389	5.2	4.4	7.6

Things to Take From One's Home During a Fire

Exemplar	OD	EG	CT	FOI	I	FAM
Children	2	9.0	3.140	8.9	9.0	7.2
Other people	5	9.0	3.184	8.2	9.0	8.6
Family	7	8.6	3.184	8.9	8.5	8.1
Important documents	8	7.5	3.921	6.0	6.3	5.2
Pets	10	7.4	3.439	6.0	6.3	6.6
Prized personal possessions	3	7.2	4.895	6.0	7.0	7.8
Money	20	7.0	3.974	5.9	6.2	7.9
Valuables	3	7.0	4.877	6.4	7.0	6.1
Dogs	2	6.9	3.351	7.1	6.6	7.0
Cats	2	6.7	3.211	4.3	5.7	5.7
Family records	4	5.3	3.921	4.7	5.2	4.2
Jewelry	8	5.2	4.228	4.5	5.8	7.7
Pictures	9	4.5	4.035	3.8	4.7	8.3
Camera	2	4.1	4.018	3.7	2.3	5.6
Memorabilia	3	4.1	3.886	4.4	5.0	5.7
Clothes	9	4.0	3.088	3.7	3.4	8.5

(*Appendix Continues*)

Things to Take From One's Home During a Fire (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Stereo	2	3.4	3.641	3.2	4.1	7.6
Blankets	3	2.7	2.649	3.2	3.3	7.1
TV	2	1.9	3.807	2.7	3.0	8.2
Food	2	1.2	2.289	3.0	2.2	9.0

Clothes to Wear in the Snow

Exemplar	OD	EG	CT	FOI	I	FAM
Down jacket	8	8.2	5.178	6.4	7.7	6.7
Mittens	8	7.2	3.792	6.6	6.1	5.0
Thermal underwear	11	7.1	4.126	5.8	8.0	5.1
Boots	22	6.7	3.386	6.4	6.4	6.7
Sweater	11	6.7	4.813	6.4	5.3	7.2
Coat	4	6.6	5.146	8.0	7.0	7.1
Ski jacket	3	6.6	5.438	6.6	7.7	5.0
Socks	5	6.6	3.917	7.2	6.5	6.6
Parka	3	6.5	5.104	6.0	7.3	3.8
Snow pants	4	6.5	4.271	5.6	6.4	3.8
Jacket	12	6.4	5.167	8.1	6.8	7.6
Cap	4	6.2	3.802	6.3	5.6	3.8
Hat	17	6.2	3.761	6.5	4.4	4.6
Scarf	9	5.7	4.011	4.2	5.2	4.5
Ski boots	2	5.3	3.698	6.0	5.9	3.9
Ski suit	3	4.6	4.802	4.7	7.4	3.9
Jeans	2	2.9	3.823	6.4	2.9	7.9

Birds

Exemplar	OD	EG	CT	FOI	I	FAM
Robin	16	8.4	5.413	6.4	5.9	2.7
Bluejay	10	7.9	5.102	6.8	6.3	3.2
Sparrow	14	7.8	4.935	7.2	4.2	3.1
Blackbird	2	7.6	5.174	5.9	3.3	2.2
Bluebird	3	7.6	5.268	4.5	5.8	3.0
Parakeet	6	7.2	4.674	5.8	5.5	1.8
Parrot	6	7.2	4.261	4.0	6.5	2.3
Seagull	3	7.2	3.812	7.7	5.3	2.8
Canary	5	7.1	4.739	4.4	5.8	2.2
Eagle	14	7.1	3.913	4.2	8.1	2.9
Hummingbird	4	6.7	3.667	6.5	7.0	2.1
Pigeon	3	6.3	4.681	7.4	3.9	3.2
Hawk	8	6.2	3.812	5.7	7.2	2.2
Cardinal	2	5.8	5.094	3.1	5.3	3.4
Dove	5	5.3	4.696	5.8	7.9	1.6
Oriole	4	5.1	5.145	2.2	5.7	1.6
Falcon	3	4.6	4.174	3.2	7.2	2.7
Condor	2	4.4	3.406	2.7	6.4	2.2
Finch	5	3.8	4.920	2.7	5.2	1.5
Chicken	2	3.7	2.776	6.2	3.3	8.2
Pelican	2	3.7	2.638	4.3	4.8	2.5

Birds (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Thrush	2	3.4	4.971	1.5	4.4	2.1
Ostrich	3	2.4	2.174	2.2	4.5	1.9
Penguin	2	2.4	2.051	2.6	6.5	2.2

Vegetables

Exemplar	OD	EG	CT	FOI	I	FAM
Green beans	3	7.6	4.358	6.9	6.2	5.1
Spinach	8	7.4	4.367	5.1	3.5	5.2
Corn	6	7.3	3.667	7.0	8.0	7.0
Carrot	18	7.2	3.900	6.4	7.2	5.3
Zucchini	6	7.2	4.708	4.6	4.8	4.9
Peas	7	7.1	3.975	5.3	5.8	4.6
Broccoli	10	6.7	4.492	4.7	3.7	6.0
Lettuce	17	6.6	4.192	7.4	7.4	6.4
Squash	3	6.6	4.667	4.9	3.5	3.6
Asparagus	4	6.5	4.050	5.0	4.0	3.6
Cauliflower	4	5.8	3.875	3.7	3.6	5.3
Cucumber	6	5.7	4.192	3.8	5.1	5.1
Celery	9	5.5	4.267	6.2	6.3	5.6
Cabbage	2	5.4	4.108	3.1	3.0	4.8
Tomato	14	5.1	2.992	7.1	7.1	6.2
Artichoke	3	4.6	3.650	3.3	4.8	2.6
Potato	5	4.5	3.433	7.5	8.0	6.6
Avocado	4	3.8	3.158	4.6	5.8	2.6
Beans	2	3.7	3.892	6.0	5.6	4.6
Onions	2	3.4	2.983	5.3	5.0	4.9
Sprouts	2	3.4	3.725	3.9	5.4	4.8

Sports

Exemplar	OD	EG	CT	FOI	I	FAM
Football	33	8.9	4.755	7.7	6.8	5.8
Baseball	23	7.9	4.382	6.1	6.5	4.7
Basketball	21	7.9	4.118	6.5	7.1	4.4
Soccer	14	7.9	4.402	4.7	6.5	3.8
Tennis	17	7.1	4.108	5.7	6.6	6.3
Hockey	4	6.8	4.313	3.5	4.2	2.4
Skiing	4	6.6	2.971	5.0	8.1	3.8
Track	3	6.0	3.784	5.0	4.2	4.8
Gymnastics	2	5.4	3.598	3.7	3.9	3.4
Rugby	2	5.4	4.441	2.7	4.2	2.2
Swimming	18	5.3	3.441	7.0	5.4	5.9
Softball	4	5.2	4.324	6.3	6.5	4.3
Wrestling	2	4.9	3.990	2.6	2.8	2.1
Jogging	4	4.6	3.814	5.2	4.9	5.0
Golf	2	4.1	3.167	4.8	5.1	2.5
Martial arts	2	4.1	3.020	2.8	4.1	2.6
Horseback riding	2	3.4	2.490	3.3	6.5	5.2
Badminton	3	3.3	3.647	2.9	4.0	3.4

Fruit

Exemplar	OD	EG	CT	FOI	I	FAM
Apple	32	8.7	4.491	7.0	7.8	6.5
Orange	32	8.4	4.565	7.6	8.0	6.8
Strawberries	12	8.1	4.426	7.3	7.8	5.6
Banana	19	7.1	3.435	7.3	6.9	5.3
Pear	14	6.8	4.611	4.0	5.9	4.5
Peach	10	6.2	5.121	5.7	5.8	5.3
Pineapple	3	6.2	3.639	5.2	6.2	5.1
Plum	2	6.2	5.009	4.6	4.7	4.4
Apricot	2	6.1	4.898	4.9	4.9	4.0
Nectarine	2	5.9	4.843	4.4	5.5	3.6
Tangerine	4	5.9	4.556	4.3	4.7	4.0
Grapes	10	5.7	4.370	6.1	7.3	6.0
Cherry	11	5.1	4.824	5.9	6.9	4.7
Watermelon	5	5.1	2.908	5.8	6.7	5.4
Berries	2	4.7	4.389	4.2	5.6	3.9
Lemon	2	4.1	3.611	5.4	3.3	6.1
Blueberries	3	4.0	4.361	3.2	4.7	4.1
Raisins	2	3.3	3.722	5.3	5.0	4.5
Tomato	4	2.0	2.111	4.3	6.2	6.2

Tools

Exemplar	OD	EG	CT	FOI	I	FAM
Hammer	35	8.9	4.042	8.1	7.8	4.9
Screwdriver	25	8.2	4.570	8.5	8.6	4.6
Pliers	5	7.6	3.681	7.1	7.7	4.4
Wrench	22	7.3	3.847	8.0	6.9	4.0
Saw	14	6.6	3.403	5.7	7.0	3.3
Drill	7	5.4	4.083	5.1	5.0	3.0
Philips screwdriver	2	5.4	3.958	6.6	7.0	4.3
Shovel	2	5.2	3.250	5.4	6.9	3.5
Chisel	2	4.5	4.722	2.7	3.3	1.6
Socket wrenches	2	4.5	3.653	6.3	5.4	1.7
Nails	5	4.4	2.917	7.4	6.5	5.5
Crowbar	2	4.1	4.153	4.3	4.3	1.6
Knife	3	3.1	4.000	5.3	6.3	7.0

Furniture

Exemplar	OD	EG	CT	FOI	I	FAM
Couch	22	7.8	4.454	6.9	5.5	6.8
Chair	30	7.7	4.815	7.4	6.0	6.9
Sofa	10	7.6	4.491	5.3	5.0	5.9
Dresser	6	7.4	4.352	6.7	5.9	6.2
Easy chair	2	7.1	4.334	5.0	3.5	5.9
Table	27	7.1	5.130	8.4	5.8	7.1
Desk	13	6.8	4.768	7.9	6.0	7.6
Coffee table	5	6.4	4.676	6.6	3.9	5.1
Dining table	2	6.3	4.639	6.0	5.7	6.3
Bed	18	6.1	4.222	7.7	8.6	8.9

Furniture (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Rocking chair	2	5.7	4.083	4.4	2.4	4.6
Stool	5	4.7	4.213	3.9	3.7	4.2
Cabinet	3	4.2	4.352	5.4	5.4	5.8
Bed stand	2	4.1	4.722	5.7	3.4	4.6
Lamp	13	3.5	2.871	7.6	7.5	7.2
TV	4	3.2	2.796	6.5	4.2	8.2
Rug	2	2.5	2.102	6.3	3.8	5.7
Refrigerator	2	2.4	2.509	6.7	8.0	7.9
Stereo equipment	2	2.1	2.806	6.3	5.2	7.4

Clothing

Exemplar	OD	EG	CT	FOI	I	FAM
Pants	28	8.4	4.746	8.4	8.0	7.4
Shirt	28	8.2	5.000	7.5	7.4	7.9
Blouse	6	7.8	4.746	3.7	7.1	7.5
Jeans	3	7.7	4.404	8.0	6.3	7.9
Dress	13	7.5	4.307	3.6	6.3	7.2
Underwear	14	6.7	4.149	7.7	7.0	7.7
Sweater	4	6.6	4.886	5.2	5.2	7.2
T-shirt	2	6.5	4.526	6.4	5.2	6.5
Trousers	2	6.3	4.675	6.5	6.7	6.5
Skirt	0	5.9	4.237	3.6	5.9	6.2
Shorts	5	5.3	4.176	7.1	4.2	6.4
Bra	2	4.9	3.956	3.3	4.8	8.1
Jacket	6	4.8	4.456	5.7	4.4	7.6
Shoes	21	4.8	2.921	7.7	6.4	8.6
Coat	3	4.7	4.298	5.8	5.2	7.1
Socks	22	4.0	3.518	7.4	4.4	6.6
Tie	7	3.1	3.211	3.8	2.3	4.0
Belt	2	2.6	3.377	6.0	3.4	6.5
Gloves	2	2.1	3.088	1.8	1.7	5.7
Hat	9	2.1	3.079	3.5	2.5	4.6

Weapons

Exemplar	OD	EG	CT	FOI	I	FAM
Gun	33	8.6	4.192	6.4	7.3	3.2
Pistol	2	8.5	3.725	6.4	6.8	3.4
Bomb	6	7.9	4.000	4.4	7.7	3.1
Rifle	3	7.8	3.808	5.8	6.9	2.3
Nuclear bomb	4	7.4	3.608	3.6	9.0	3.6
Nuclear missle	2	7.3	3.633	4.2	8.7	3.0
Missle	4	7.0	4.017	3.5	8.0	2.7
Hand grenade	3	6.4	3.708	2.7	7.0	1.6
Bazooka	2	5.9	3.775	3.9	7.5	1.5
Switchblade	3	5.6	3.933	5.8	4.6	1.4
Sword	4	4.9	3.992	5.0	4.0	2.0
Knife	29	4.8	3.908	7.4	4.9	7.0
Bow and arrow	3	3.9	3.692	3.7	3.7	2.9
Cannon	2	3.5	3.642	3.3	6.6	2.0

(*Appendix Continues*)

Weapons (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
Spear	2	3.2	4.217	2.8	3.1	2.2
Stick	10	2.9	3.858	5.8	2.1	3.9
Ax	2	2.8	3.458	3.2	3.7	2.3
Hands	2	2.8	2.958	5.5	3.2	8.6
Chain	4	2.7	2.975	5.1	3.3	5.1
Jet	3	2.3	3.150	2.8	5.1	5.8
Rocks	2	2.1	3.317	5.3	1.8	3.9

Vehicles

Exemplar	OD	EG	CT	FOI	I	FAM
Car	31	9.0	5.563	8.4	6.1	8.6
Truck	14	6.9	4.646	6.4	5.8	3.9
Bus	11	6.6	4.094	5.1	6.5	5.4
Sports car	3	6.4	4.885	6.0	4.8	6.1
Porsche	2	6.3	4.854	5.9	5.7	5.7

Vehicles (*continued*)

Exemplar	OD	EG	CT	FOI	I	FAM
VW	2	6.2	4.927	6.8	5.2	4.2
Camaro	3	6.1	5.042	6.5	5.2	4.2
Motorcycle	15	5.7	4.052	6.2	6.3	3.5
Mustang	2	5.6	5.177	6.8	5.5	5.3
Plane	11	5.5	2.135	6.6	7.8	7.4
Jeep	3	5.4	4.709	4.1	5.4	4.7
Bike	24	4.3	3.084	6.7	5.4	5.1
Train	8	4.3	3.271	3.5	7.0	4.2
Moped	6	3.7	3.521	3.7	5.9	4.4
Boat	7	3.3	2.209	6.3	4.1	5.1
Tractor	2	3.2	3.458	1.7	2.5	2.0
Skateboard	3	1.4	1.875	4.9	3.9	3.3

Received February 14, 1984
Revision received March 27, 1985 ■