Memory & Cognition
2006, 34 (7), 1431-1443

Effects of classification context on categorization in natural categories

JAMES A. HAMPTON
City University, London, England

DANIÈLE DUBOIS
CNRS, Paris, France

and

WENCHI YEH
Hewlett-Packard, Mountain View, California

The patterns of classification of borderline instances of eight common taxonomic categories were examined under three different instructional conditions to test two predictions: first, that lack of a specified context contributes to vagueness in categorization, and second, that altering the purpose of classification can lead to greater or lesser dependence on similarity in classification. The instructional conditions contrasted purely pragmatic with more technical/quasi-legal contexts as purposes for classification, and these were compared with a no-context control. The measures of category vagueness were between-subjects disagreement and within-subjects consistency, and the measures of similarity-based categorization were category breadth and the correlation of instance categorization probability with mean rated typicality, independently measured in a neutral context. Contrary to predictions, none of the measures of vagueness, reliability, category breadth, or correlation with typicality were generally affected by the instructional setting as a function of pragmatic versus technical purposes. Only one subcondition, in which a situational context was implied in addition to a purposive context, produced a significant change in categorization. Further experiments demonstrated that the effect of context was not increased when participants talked their way through the task, and that a technical context did not elicit more all-or-none categorization than did a pragmatic context. These findings place an important boundary condition on the effects of instructional context on conceptual categorization.

A phenomenon of major importance for psychological theories of concepts is the *vagueness* of many of our conceptual categories. Although every category can be said to have clear members (for example, a chair is clearly a type of furniture) and clear nonmembers (a cucumber is clearly not a type of furniture), there are also instances that are *borderline* to a category. For instance, when asked to decide whether rugs, paintings, or televisions are types of furniture, people are frequently uncertain about the answer. There is a vagueness in our use of common language terms that arguably makes such questions undecidable. The problem of vagueness poses serious threats to many accounts of the semantics of natural language (Keefe & Smith, 1996; Osherson & Smith, 1997), so the issue of what gives rise to the phenomenon is of central importance to theories of cognition.

There have been many demonstrations of vagueness. For example, McCloskey and Glucksberg (1978) presented two groups of students with lists of words, each list headed by a category name, such as "fruit" or "fish." One group was asked to give typicality ratings, identifying how typical or representative each word was of the category as a whole. The other group made a simple *yes–no* categorization decision about each word and returned 4 weeks later to make the same decision again. Many items in the lists showed high levels of disagreement between participants and poor test–retest reliability or consistency. These items also tended to be borderline in terms of their rated typicality in the category.

In a subsequent reanalysis of McCloskey and Glucksberg's (1978) data, Hampton (1998) showed that categorization probability for an item was closely related to rated typicality by a simple, monotonically increasing threshold function. List items that deviated from this standard function tended to be unfamiliar, or they might be parts or

This research was made possible by financial support to the first author from the French Ministry for Higher Education and Science, the Nuffield Foundation, the Fulbright Program, and the British Academy. Their assistance is gratefully acknowledged. The research was conducted while the first author was on sabbatical leave at the University of Chicago, whose hospitality is also acknowledged. The authors thank Lawrence Barsalou, Nick Braisby, Bradley Franks, Dedre Gentner, Douglas Medin, Gregory Murphy, Lance Rips, and Dan Sperber for helpful discussion, and our anonymous reviewers for helpful suggestions. Correspondence relating to this article may be sent to J. A. Hampton, Psychology Department, City University, Northampton Square, London EC1V 0HB, England (e-mail: hampton@city.ac.uk).

—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

properties of an instance rather than instances themselves. For biological categories, particular items also could deviate from the function if they had the appearance of a category member without technically belonging to the category or, conversely, if they technically belonged to a category but did not share its appearance features. From this analysis, Hampton (1998) argued that categorization decisions are to a large extent based on the same "family resemblance" semantic information that is used in judging typicality: An item is judged to be a category member if the similarity between the item and the prototype for that category passes some threshold value. Because both the concepts retrieved and the threshold criterion may vary across occasions, the probability of categorization rises as a monotonic threshold function of the semantic similarity of the instance with category concepts.

Why should categorization be so unstable at the category borderline? Barsalou (1987) argued that instability could reflect variation between individuals in their conceptual representations or in their recent experience with a category. However, it is also possible that instability in categorization results from the lack of a specific context with respect to which the categorization has to be made. In everyday language, words are used in specific contexts with specific communicative goals, and this contextual support is missing in standard categorization experiments. If individuals respond to the lack of context by arbitrarily constructing one of their own, differences in the resulting conceptual representations would create instability.

The purpose of the present study was to investigate variation in categorization resulting from a particular type of contextual source. It has been argued (Braisby, 1993; Braisby & Franks, 1997, 2000; Braisby, Franks, & Harris, 1997) that a major source of instability and vagueness in categorization judgments is the lack of any explicit context for the categorization. If asked whether a television is furniture, someone may give a different reply if the question is asked in the context of designing the look of a living room, as opposed to planning the need for electrical outlets in the home. The purpose for which a classification is made may be crucial to how it is performed. Braisby and Franks (1997) went so far as to argue that the lack of a clear context, or perspective, may be a major reason that categories appear to be so vague (see also Rey, 1983). According to their position, the observed vagueness of categories is in large part the result of categorizers' selecting at random different well-defined concepts relevant to different contexts or "perspectives." Since individuals recruit their own default context to the task, differences of opinion about categorization may be more apparent than real. This hypothesis resonates with Barsalou's (1987) proposal that people construct category representations "on the fly" as different tasks are presented to them, so that there is inherent instability in the information represented in working memory on any one occasion of categorization.

Several studies have deliberately manipulated context in categorization tasks (for a review, see Murphy, 2002, pp. 413–422). For example, Roth and Shoben (1983) varied sentential contexts, as in "The *bird* crossed the farmyard," and showed via measures of sentence processing that the context could lead to reversal of normal typicality effects (for example, in the sentence above, *chicken* was read faster than *robin*). Barsalou and Sewell (in a study described by Barsalou, 1987) showed that asking participants to take the point of view of (say) a suburban housewife rather than a "redneck" farmer produced marked changes in the typicality ranking of instances within categories such as vehicles or foods. In both of these studies, situational context was manipulated and marked effects were observed on the relative typicality and ease of processing of different category instances. Another important demonstration of shifts in categorization with context was provided by Medin, Lynch, Coley, and Atran (1997). In their study, different groups of tree experts sorted trees by similarity. Depending on whether the experts were taxonomists, landscape gardeners, or park maintenance staff, the structures observed were very different. Each group had important dimensions of similarity, not used by the others, that were relevant to their own profession.

Our experiments differed from these studies in several ways. First, we focused not on the situational context of a classification, but rather on its *purpose*. It is clear that when considering animals in the situational context of eating them, as opposed to in the context of inviting them into your home, you will adopt very different views of what instances make typical candidates. However, it is a largely unexplored question whether, differences in expertise aside, changing the purpose for which a classification is to be used will generate such shifts. Second, our experiments differed from many previous studies in that rather than measuring typicality structure, we measured changes in categorization itself. We wished to explore whether an item would be considered a member of a category in one purposive context but not in another. If category vagueness derives in part from contextual ambiguity, it is clearly important to show that context can affect categorization decisions as well as typicality structures. Finally, we wished to explore the possibility that disagreement and inconsistency would be reduced when a clear purposive context for classification was provided.

In an unpublished study, Braisby and Franks (2000) found evidence that categorization could be strongly influenced by shifting perspectives. They contrasted two types of borderline instances for natural kind classes: those that had the appearance but not the essence of a category (e.g., an Easter egg as an egg) and those that had the essence but not the appearance (e.g., a scrambled egg as an egg). People were asked to judge whether it was appropriate to use the word *egg* to talk about each object. In a series of experiments, the researchers showed that the relative frequency with which these two types of instance were categorized as eggs depended on a number of factors. Taking the perspective of a sculptor rather than a biologist naturally put more weight on appearance. Imagining speaking to an adult nonnative speaker led to more weight on essence, whereas imagining speaking to a 4-year-old child led to more weight on appearance. If the purpose of using the word was in conversation or in defining a mean-

ing, more weight was placed on essence, whereas using a word for picking an object from an array switched the weight to appearance. Focusing on "true classification" rather than appropriateness of word use shifted weight onto essence. It is clear, therefore, that communicative setting and purpose can be very influential in affecting categorization, at least as shown in people's judgments of appropriate words to use for objects. Our aim was to determine how general this effect may be by looking at a wider range of categories and borderline cases, as well as by measuring the consistency with which people make their judgments.

To test whether vagueness results from a lack of information about the purpose of classification, we aimed to provide participants with a clear perspective from which to make their categorizations. If the purpose of categorization is made clear, there should be less vagueness. The first experiment therefore employed three categorization conditions. One condition (the *no-context control*) was a simple, context-free categorization task, whereas the remaining two conditions offered different scenarios for the purpose and importance of the classifications. The first prediction was that providing a specific context of some kind would lead to less individual disagreement and inconsistency—that is, to less vagueness.

The second prediction concerned the kind of categorization context provided and the degree to which categorization would be dissociable from similarity to the category prototype. Hampton (1998) argued that the degree to which categorization probability is a simple monotonic function of typicality can be taken as a test of the degree to which participants are simply categorizing on the basis of similarity to prototype, as opposed to using a more complex, explanation-based or theory-driven decision process. For example, in biological categories (but not others) categorization probability was not well correlated with typicality for items that were technically in categories different from their appearance (for example, whales and bats).

The importance of causal-explanatory theories for categorization has been well established (Ahn, Kim, Lassaline, & Dennis, 2000; Murphy & Medin, 1985). Rips (1989) has argued that similarity is only a crude approximation of the basis on which people categorize the world. Under the right circumstances, it is possible to show that categorization and typicality judgments may be dissociated, as people turn for categorization to deeper core information about a concept and ignore superficial appearance (Ahn & Dennis, 2001). We hypothesized that depending on the classification context, similarity may turn out to be a more or less appropriate basis for categorization. For example, when setting up a news/interest group on the subject of fish, it would be appropriate to apply a "loose" interpretation of the category that could include shellfish or dolphins along with "true" fish, such as cod or trout. On the other hand, when preparing a scientific report on the ecological status of different species, it would be more appropriate to use a quasi-biological definition of fish, which would exclude shellfish or dolphins but might instead include seahorses. Our choice of contexts was designed to take advantage of this intuition that there may be more technical and more pragmatic forms of categorization. In the *technical condition*, the purpose of categorization had a technical foundation, based on scientific or legislative goals, but in the *pragmatic condition*, the purpose was more loosely practical, based on providing a classification that would be easy to use and would match people's general expectations. We predicted that a pragmatic context would be more likely to reveal similarity-based categorization, since it would place things into categories in which the mass of people would expect to find them. Categorization would therefore rely more on the "identification schema" or prototype of a concept, and less on its core definition. A technical context, however, should encourage participants to use deeper causal-explanatory schemas in which the relation of categorization to typicality is less direct. In order to test the degree to which categorization is based on similarity, we calculated correlations across items between the probability of a particular categorization and the mean rated typicality provided by an independent group of participants.

A further measure to be compared between these conditions was the overall threshold criterion used for categorization in each category (Hampton, 1995). We expected that in a pragmatic context, people would take a broad view of what may be included in a category, whereas in a technical context, the category boundary would be drawn more tightly. The no-context condition was predicted to be intermediate between these two.

A final source of interest in the task came from possible differences between categories in the degree to which they would be affected by contextual instructions. There has been considerable interest in domain differences in how concepts are represented (Barr & Caplan, 1987; Estes, 2003; Kalish, 1995). Although the design of the experiment was too small to permit adequate sampling from different semantic domains, we deliberately chose categories from four different ontological domains in order to provide a broad range of materials. Two biological kinds, fish and insects, were expected to show marked differences between technical and other contexts, because of the existence of biological definitions for these terms. Two artifact kinds, tools and furniture, were expected also to show changes across condition, because technical contexts would place greater weight on the utility or function of the objects, which may provide the central core of artifact concepts (Bloom, 1996). Finally, we also included two categories of edible plants, fruits and vegetables, and two social activities, sports and sciences, in which the technical contexts were expected to tap other possible forms of theoretical knowledge and beliefs. Although these pairs of semantic categories were clearly not unbiased samples of their respective domains, it was intended that analysis at the level of individual categories could provide indicative evidence of any strong and systematic domain differences that may exist.

To recap, we predicted first that adding any context at all would reduce vagueness, and second that technical

contexts would contrast with pragmatic contrasts by inducing tighter category boundaries and a reduction in the dependence of categorization on similarity, as measured by the correlation with typicality.

EXPERIMENT 1

Method

Participants. One hundred students at the University of Chicago volunteered for the experiment in return for a small payment. All were fluent speakers of English.

Materials. The categories and items used are shown in Appendix A. There were 24 items in each of 8 categories. Items were selected from category norms (Battig & Montague, 1969; Hampton & Gardiner, 1983; McCloskey & Glucksberg, 1978; Rosch & Mervis, 1975) and were designed to include clear members and nonmembers of the category, together with a substantial number of possible borderline cases to provide a measure of how vagueness and category membership change with context. Examples of scenarios for each condition are given in Appendix B. The aim for the technical condition was to provide a legalistic or scientific context stressing the important consequences of making a correct classification. For fish and insects, the classification was to be used by a government agency for monitoring the ecological performance of different nations. For fruits and vegetables, the classification was concerned with economics and trade. For furniture and tools, the context involved tax regulations, and for sports and science, the context concerned appropriate use of funds by government agencies. In the pragmatic condition, the stories were concerned with placing things in categories where people would expect to find them, so that they would be easily found. A variety of contexts were used, including an Internet news group (for fish and insects), a mail-order catalog (for fruits and vegetables), a department store database for monitoring stock (for furniture and tools), and a library index (for sports and science). Finally, the no-context condition had the same instructions for all categories: "Consider each of the following items and decide whether they belong in the category of _____."

Design. There were four groups of participants. One group of 40 students provided typicality ratings. Three other groups of 20 students each made *yes–no* categorization judgments, according to the three conditions. Categorization was retested 3–4 weeks later.

Procedure. The participants were given booklets to complete under supervision. The orders of categories within booklets and words within categories were balanced. The participants in the categorization conditions were asked to read the scenario and judge each item by circling one of three choices: *Y* (yes), *N* (no), or *Ø* (meaning of word unknown). After 3–4 weeks, they repeated the task with instructions to make fresh judgments without trying to recall their earlier decisions. A new random order of categories was used.

The typicality rating condition used a scale from 1 to 10 (10 = *most typical*) and standard typicality instructions similar to those of McCloskey and Glucksberg (1978): We told our participants to "rate each word according to how typical or atypical it is as a member of the category . . . decide how good or representative an example each word is of the category named." An example was then given for the category "flowers." Item and category orders were counterbalanced, as in the other conditions. The typicality ratings were not subjected to a retest.

Results

For each item in each of the three categorization conditions, the probability of a "yes" response was calculated on the basis of the two responses made by each participant in each condition. The reliability of the probabilities (mean .99) and the mean typicality ratings (.96) were uniformly high.

Intersubject agreement. The first measure of category vagueness was intersubject agreement. The addition of contextual instructions was predicted to reduce contextual ambiguity, so people would agree more on the categorization of items. Columns 2–5 in Table 1 show disagreement measured as the proportion of nonmodal responses (NMR), the proportion of participants giving a "no" response when the majority said "yes," or a "yes" response when the majority said "no." Overall, as predicted, there was slightly more disagreement in the no-context condition (18.6%) than in the other conditions (17.1% and 18.0% for the pragmatic and technical conditions, respectively, *SE*s = 1%). These means are for all items, including clear members and nonmembers. NMR rose to a maximum of around 35% for items at the center of the typicality scale (as in McCloskey & Glucksberg, 1978), leaving plenty of opportunity for a reduction in the item's value. We ran ANOVAs by participants and items, with category and condition as factors. There was no overall effect of condition ($F_1 < 1$), nor was the planned comparison between the no-context and the other two conditions significant [$t(57) = 0.9$, n.s.]. There was therefore no evidence that providing a context increased intersubject agreement. With $\alpha = .05$, the estimated power to detect a difference in condition means of as much as 5% between the no-context and the other conditions was greater than 97%. (In reporting power estimations, the lower of the two powers—for items or participants—is always quoted.)

Otherwise, there was a significant main effect of category [$F_1(7,399) = 9.78$, $p < .001$; $F_2(7,184) = 2.34$, $p = .026$] attributable to greater agreement about the four biological

Table 1
Percentages of Nonmodal Responses and of Categorization Responses That Remained the Same at Retest, As a Function of Condition and Category for Experiment 1

	Nonmodal Responses				Categorization Responses Remaining the Same			
Category	No Context	Pragmatic	Technical	Overall	No Context	Pragmatic	Technical	Overall
Vegetable	17	15	14	16	92	90	92	92
Fruit	15	17	14	15	94	92	88	91
Fish	9	11	11	10	97	94	95	95
Insect	18	17	18	18	89	94	89	91
Sport	23	19	24	22	86	88	89	87
Science	26	23	24	24	88	90	85	87
Tool	24	17	20	20	87	90	86	87
Furniture	17	18	19	18	89	93	85	88
Overall	18.6	17.1	18.0	17.9	90	91	89	90

categories than about the activity and artifact categories (mean NMRs = 14.6 and 21.2, respectively). There was no interaction between condition and category. In summary, in none of the categories was the no-context condition clearly subject to greater disagreement than the other conditions. Our expectation that a more specific categorization context would generally reduce category vagueness, as indexed by levels of disagreement, was not supported.

Within-subjects consistency. The second measure of vagueness was the test–retest consistency of categorization responses. We had predicted that the addition of a context would improve the stability of categorization decisions by reducing contextual ambiguity. If an item is categorized in relation to a contextualized concept, then recall of the same context should facilitate consistent responding. Note that if different individuals contextualized the concepts in idiosyncratic ways, NMR would remain high, but consistency should still improve relative to the no-context condition. Columns 6–9 in Table 1 show consistency as the percentage of responses that were the same on retest. On average, the same response was given 90% of the time. This level of consistency compares with a mean level of 88% in McCloskey and Glucksberg's (1978) data, and includes clear members and nonmembers. At the middle of the typicality scale, mean consistency fell to around 82%. There was no tendency for contexts to increase consistency. Mean values across conditions varied little, from 91% for the pragmatic condition to 89% for the technical condition (SE = 0.7%). In ANOVAs, only the main effect of category was significant in both analyses. Estimated power to detect the contrast between the no-context and the other conditions was 70% for a difference in means of 3%, and 97% for a difference of 5% (α = .05, two-tailed). The effect of category again showed up as a difference between the biological and food categories (92%) and the other categories (88%). In summary, the main prediction of lower consistency in the no-context control condition was also not supported.

Relation of categorization to typicality. The next analysis considered the correlation between categorization probability in each of the categorization contexts and the mean rated typicality of the items. To the extent that a high correlation is observed between categorization probability and typicality, it may be concluded that categorization is based on similarity, and for this reason we had argued that pragmatic contexts should show higher correlations with typicality than do technical contexts. To illustrate these correlations for "default" categorization, Figure 1 shows scatterplots for each of the categories between probability of categorization in the no-context condition (y-axis) and rated typicality (x-axis). It can be seen that for all categories except sport, there was a strong and systematic relation between the two measures. Columns 2–5 in Table 2 show the correlations between typicality and categorization probability. All correlations were high (the mean correlation of .95 was close to the theoretical maximum imposed by the reliabilities of the measures), and differences between the three overall correlations by condition were slight and not statistically significant

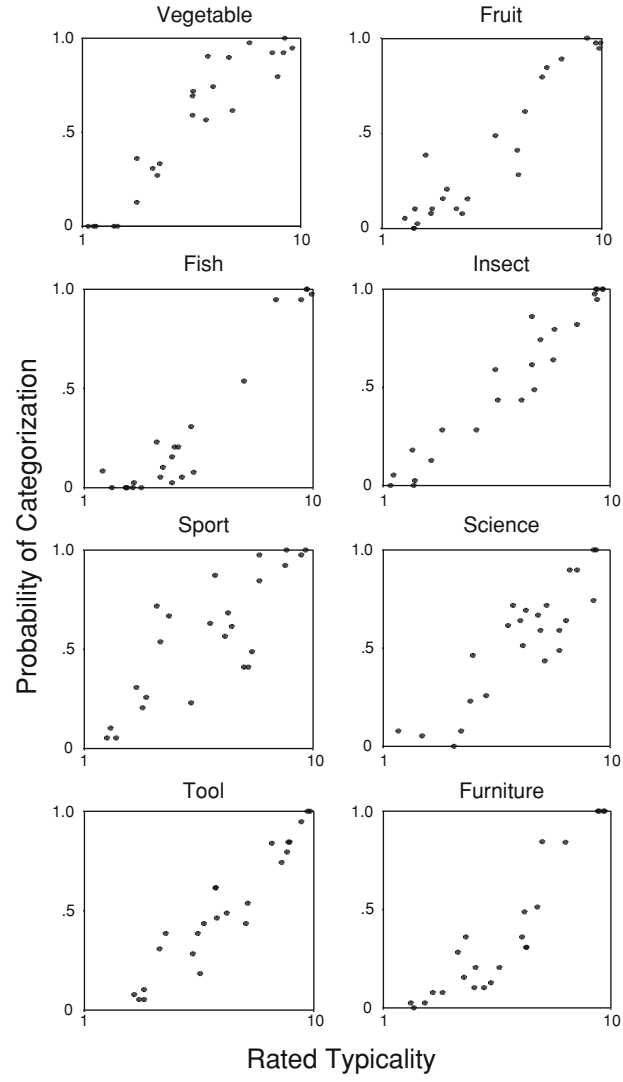

Figure 1. Scatterplots of categorization probability with typicality for each category in the no-context condition of Experiment 1.

[$\chi^2(2) = 2.01, p > .10$; Weatherburn, 1961, pp. 203–205]. Estimated power was over 90% for detecting a difference in mean correlations of as little as .95 versus .90 (significant at .05) between any two conditions. Taking the full set of 24 correlations as a whole, however, there *was* evidence for nonhomogeneity [$\chi^2(23) = 43.4, p < .01$], suggesting that some individual correlations were significantly lower than the rest.

In this instance, it was quite easy to find plausible accounts for the cells in the correlation matrix with lower coefficients. The biological and food categories had uniformly high correlations with typicality (.94 or greater). The categorization of sports correlated less well with typicality (except in the pragmatic condition), whereas for tools and furniture it was the pragmatic condition that showed a lower correlation. Borderline sports activities could be divided into two groups, those involving physi-

Table 2
Correlation of Categorization Probability With Typicality and Percentage of Positive Categorizations, As a Function of Condition and Category in Experiment 1

	Correlation of Categorization With Typicality				Percentage of Positive Categorizations			
Category	No Context	Pragmatic	Technical	Overall	No Context	Pragmatic	Technical	Overall
Vegetable	.96	.97	.94	.96	53	53	57	53
Fruit	.94	.96	.97	.96	40	45	42	42
Fish	.95	.96	.95	.95	29	33	33	33
Insect	.98	.97	.99	.98	55	59	52	54
Sport	.87	.96	.91	.91	56	57	56	57
Science	.92	.92	.96	.93	54	61	60	61
Tool	.97	.86	.94	.94	52	45	50	51
Furniture	.95	.90	.96	.94	39	39	39	39
Overall	.94	.94	.95	.95	47	49	49	49

cal exercise but little skill (e.g., aerobics and jogging) and those involving skill but no physical effort (croquet, billiards, and darts). In the pragmatic condition, rated typicality was a good predictor of categorization probability for both groups. For the other contexts, however, the categorization placed more emphasis on skills than on physical effort, so rated typicality overpredicted the categorization of aerobics and jogging as sports, and underpredicted that of croquet, billiards, and darts. The upper panel of Figure 2 shows the effect for the technical condition, in which a categorization was required for use in sports funding decisions. The similarity between the technical and the no-context conditions for sports suggests (interestingly) that the default categorization of sports also places greater weight on the skills aspect of sports than is seen in typicality judgments.

Tools and furniture showed reduced correlations with typicality in the pragmatic condition only. These scenarios involved devising a database for employees in a department store so they could check on the availability of different items. Participants appear to have adopted the layout of department stores as a guide to categorization. For example, electrical appliances of various kinds were less likely to be counted as furniture, since in many department stores they would be in the electrical goods department rather than the furniture department (see lower panel, Figure 2). This strategy was also accompanied by higher consistency for these two categories in the pragmatic condition (.90 and .93) than in the other conditions (values ranging from .85 to .89), and for tools it also was accompanied by a reduction in nonmodal responses. Note that the reduction in vagueness and the reduced correlation with typicality were produced only in those context–category combinations that allowed a *situational* context to be imagined. Providing a purpose for the classification per se had no discernible effect overall, yet the dependent measures *were* sensitive to situational context effects.

The high overall level of the correlations was obviously affected by the presence in the lists of clear members and nonmembers. Removing items with less than 10% NMR left between 9 and 19 borderline items in each list. With reduced range and reliability, the mean correlation between typicality and categorization fell to .85, and the four correlations identified before were the only ones to fall below .8.

Criterion. A further important way in which context could affect categorization was in terms of the breadth of the categories. We had predicted that technical contexts might lead to narrower category criteria, since the instructions stressed the importance of producing a fair categorization. Columns 6–9 in Table 2 show the mean percentages of positive categorizations. Contrary to expectation, the no-context condition had the *tightest* criterion overall, with 47% positive responses versus 49% in the other conditions

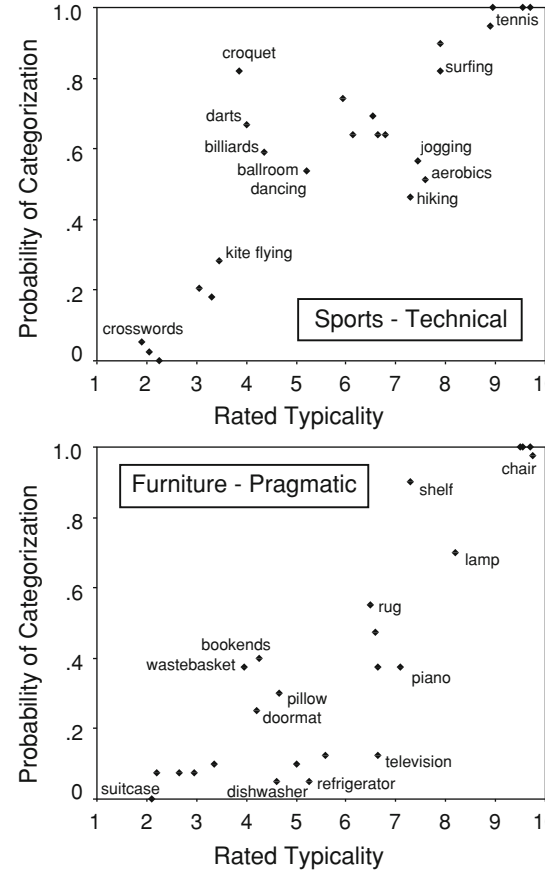

Figure 2. Scatterplots of categorization probability with typicality in Experiment 1, for sports in the technical context condition and furniture in the pragmatic context condition.

($SE = 2.6\%$). Only in the insect category did the technical condition, which had been predicted to be the tightest, have the lowest number of positive responses. The probability data (with an arcsine transformation) were subjected to ANOVAs by participants and items. None of the effects were consistently significant across both analyses. Our estimated power to detect a contrast between the no-context and the other conditions was 70% for a difference of as much as .05 in mean probability, and over 90% for a difference greater than .07 ($df = 57$, $\alpha = .05$).

Discussion

The experiment set out to examine a number of different measures to test the potential effect of a purposive scenario on categorization. First, we looked at levels of agreement and consistency in categorization, and no systematic effects were found of supplying a contextual scenario. Second, we examined correlations between typicality and category criterion as a test of the idea that "loose" pragmatic contexts would encourage use of default similarity to a common prototype and a broad criterion, whereas more technical contexts would encourage explanation-based categorization and a narrow criterion. The results also failed to support this notion. In fact, for five of the eight categories, technical contexts were more similarity based than were pragmatic contexts.

Only for one category and one context—tools in a department store scenario—did the provision of the context systematically affect all of our measures in the predicted way, by (1) reducing the vagueness in categorization in terms of less disagreement (17% vs. 20%) and greater consistency (90% vs. 87%), (2) reducing the correlation with typicality (.86 vs. .94), and (3) tightening up the category (45% vs. 51%). This consistent pattern is evidence that the manipulation of context can work in the predicted way. The interesting point to note is that the department store context was in many ways very similar to the types of context used by Roth and Shoben (1983), in which provision of a situational setting, such as a farmyard, affected the typicality of different birds. Thus, although we had stressed the *purpose* of the classification in our instructions, the existence of a familiar situational context was probably the key factor in changing categorization. It remains to be seen whether this one case can be generalized.

Our failure to obtain the predicted effects of context in general across the range of measures, however, suggests either that participants were ignoring the instructions to imagine themselves in the given scenarios or that their "default" way of thinking of each category was sufficiently powerful to be recruited into the different contexts relatively unchanged. Against the conclusion that the context had no effect at all, one can point to systematic effects observed for sports, furniture, and tools. These effects were associated with changes in the categorization of particular subcategories of items, such as exercise sports versus skilled sports, or electrical appliances versus more decorative furniture. The most remarkable finding was that the no-context condition showed no systematic differences overall from the other conditions in its correlation with typicality, its category breadth, its between-subjects disagreement, or its within-subjects consistency. The claim that vagueness in the standard categorization task might generally be the result of contextual ambiguity is hard to reconcile with this demonstration.

One might still argue that the context instruction was not taken sufficiently seriously by the participants. Perhaps they were uninvolved in the task, and so relied on default categorization. Accordingly, in Experiment 2 participants were instructed to speak aloud as they read the context stories, and then to spend a minute describing the basis on which they would categorize before starting to categorize each list. It would be hard under these circumstances for the participants to ignore the stories. Smith and Sloman (1994), for example, found that when they were asked to "think aloud" in one of Rips's (1989) categorization tasks, participants were more likely to show deeper rule-based rather than similarity-based reasoning.

EXPERIMENT 2

Experiment 2 was a partial replication of Experiment 1. We felt that if an effect of context were to be observed, the most powerful manipulation would be the contrast between the pragmatic context, in which items should be categorized where people would expect to find them, and the technical context, in which equitable rules and regulations relating to financial and professional interests were required. These were also the only two conditions in which there was any kind of story provided for participants to read out loud. Therefore, we now considered only these two conditions.

Method

Participants. Forty undergraduate students from the same population at the University of Chicago were paid to participate in the study. None had taken part in Experiment 1.

Design and Materials. The materials and participant population were identical to those of Experiment 1 so that comparisons could be made between the experiments. Twenty participants served in each of the two conditions, pragmatic and technical. Each task was only performed once.

Procedure. The participants were given the booklet and asked to explain out loud the situation described. They then spent a minute reflecting on how they were going to approach each task and what aspects of the category would be important. They were given the example of classifying weapons either for a museum display or for legislating about the legal age of possessing them. The participants then performed *yes–no* categorization judgments to the 24 words listed in one of two orders, as in Experiment 1. The sessions were tape-recorded.

Results

Intersubject agreement. Columns 2 and 3 in Table 3 show percentages of NMR. ANOVAs were run by participants and items with the factors condition and category. Only the effect of condition was significant across both analyses [$F_1(1,184) = 21.08$, $p < .001$; $F_2(1,38) = 8.65$, $p < .01$]. NMR was higher in the technical condition (21%) than in the pragmatic condition (17%, $SE = 0.3\%$). As in Experiment 1, the biological categories (17%) showed less

disagreement than did the others (22%), but in this case the category factor was not significant by items. Unlike the present experiment, Experiment 1 showed no statistically reliable difference in NMR between the pragmatic (17%) and technical (18%) conditions. An ANOVA with the factors experiment, condition, and category showed a two-way interaction between experiment and condition that was significant by items [$F_2(1,184) = 9.72, p < .01$] and marginally significant by participants [$F_1(1,114) = 2.9, p < .10$]. Comparing Experiments 1 and 2, the level of NMR for the pragmatic condition was the same (17%), whereas NMR for the technical condition increased from 18% to 21%. Thus, one effect of requiring participants to pay greater attention to the context was paradoxically *greater* disagreement among the participants, but only in the technical context condition.

Correlation with typicality. Columns 4 and 5 in Table 3 show that correlations of categorization probability with typicality were generally high (mean $r = .95$). If a technical context leads to more theoretically based categorization and less emphasis being placed on superficial similarity, correlations with typicality should be lower for the technical contexts. In the event, all but one of the categories showed *higher* correlations with typicality in the technical condition than the pragmatic condition [$t(7) = 4.29, p < .005$, across categories], a trend that was present but not significant in Experiment 1. Thus, the effect of paying more careful attention to the task and verbalizing the process of performing the categorization was that participants in the technical condition appeared to adhere even more closely to default similarity to prototype as the basis for their categorizations.

Examination of individual category data showed that tools and furniture were subject to the same "department store" effect as before, with reduced correlations of categorization with typicality in the pragmatic condition (.89 and .91, compared with .93 to .97 for the rest of the categories). For example, the items telephone, dishwasher, refrigerator, and piano were all less likely to be categorized as furniture than predicted by their typicality, and these were all items that would not be normally found in the furniture section of a store.

Criterion. The overall proportion of positive categorizations in each condition was identical (.49) and no different from those in Experiment 1. Columns 6–8 in Table 3 show the mean categorization probabilities. Categorization probabilities were also compared with (1) the no-context condition of Experiment 1 and (2) the conditions in Experiment 1 with the identical context stories. There were no systematic changes discernible, although some category criteria grew larger and some smaller when participants were required to verbalize the task.

Transcripts. Transcripts of the sessions indicated that participants were clearly aware of the requirements of the task and had fully understood the scenarios. When asked about the intended basis for categorization, participants in both conditions tended to say that it would be based on the characteristics of the item and their own "gut feeling" about the category membership. There were some attempts in the technical scenarios to find defining features, but they were not applied systematically. In sum, the transcripts revealed that participants were taking the task seriously and attempting to engage with the scenarios appropriately. However, the categorization data suggest that the basis on which they were categorizing was not markedly different from the one they would use in the absence of any particular context.

Discussion

Experiment 2 replicated two conditions of Experiment 1 under instructional conditions that would encourage greater attention to the contextual manipulation. If the failure of Experiment 1 to show effects of scenario contexts resulted from participants' not reading the stories and just proceeding to categorize by default, we expected that the effects originally predicted would emerge in this experiment. The results showed that if anything, participants were *more* likely to rely on their default prototype representation of the categories when the task was made more explicit and verbal. Especially in the technical condition, participants appeared to adhere even more closely to default similarity to prototype as the basis for their categorizations, and there was an increase in disagreement for this condition in comparison with Experiment 1. There was no evidence for contextual ambiguity as a source of vague

Table 3
Percentage of Nonmodal Responses, Correlation of Categorization Probability With Typicality, and Mean Categorization Probability As a Function of Condition and Category for Experiment 2, As Well As Mean Probabilities for the No-Context Condition in Experiment 1

	Percentage of NMR		Correlation With Typicality		Categorization Probability		
Category	Pragmatic	Technical	Pragmatic	Technical	Pragmatic	Technical	No Context (Exp. 1)
Vegetable	15	20	.95	.96	.51	.51	.53
Fruit	13	16	.95	.97	.38	.41	.40
Fish	16	19	.96	.97	.39	.41	.29
Insect	16	21	.97	.98	.46	.50	.55
Sport	20	26	.93	.96	.64	.51	.56
Science	19	26	.93	.93	.64	.60	.54
Tool	20	27	.89	.95	.47	.58	.52
Furniture	18	16	.91	.95	.42	.37	.39
Mean	17	21	.94	.96	.49	.49	.47

categories. Nor was there evidence for more explanation-based categorization in the technical condition, unless the increased level of disagreement in that condition reflected a greater diversity of individual rules being used by the participants. In that case, the categorization probabilities could have resembled the defaults through the process of averaging across individual differences.

EXPERIMENT 3

The third experiment used a different measure of the vagueness of categories. Kalish (1995) presented participants with a categorization scale in which they could either make a clear-cut yes or no decision or make a graded category membership judgment. This procedure was adapted in Experiment 3 to measure whether participants' view of the gradedness of categorization might be influenced by the different contexts. Specifically, we expected that in the pragmatic scenarios, where looser similarity-based classification might be considered appropriate, participants would tend to select graded category membership, whereas in the technical scenarios, which emphasized the importance of a correct and fair classification, responses would indicate clear-cut categorization.

Method

Participants. Forty undergraduate students at the University of Chicago were paid a small amount to participate in the study, with 20 in each of the conditions.

Materials. The same categories, items, and contexts were used as in Experiment 2.

Design and Procedure. The participants were given the same booklets used in Experiment 2, but with a different response scale. Specific instructions (adapted from Kalish's, 1995, Study 1) were given on how to use the response scale, which consisted of nine boxes. At one end was a box marked *not at all* and at the other a box marked *completely*. These response boxes were to be used if an item was clearly in or out of the category, and if category membership was felt to be an all-or-none affair just those options were to be used. Alternatively, if a categorization was felt to be a matter of degree, a graded response was to be chosen from the intermediate boxes, labeled 1–7 to indicate increasing degrees of membership from *barely*, through *sort of*, to *very much*.

Results

Use of extreme responses. The prediction of the experiment concerned the use of the two extreme response boxes (*not at all* and *completely*), as opposed to selection of a graded response. Figure 3 shows the proportional use of extreme responses as a function of the typicality of items (taken from Experiment 1) for each of the conditions. Judgments of graded membership were most common in the center of the typicality scale, where they occurred 25%–30% of the time. Table 4 shows the percentage of extreme responses given in each condition. The greater overall use of extreme responses for the pragmatic condition (85%, $SE = 2.9\%$) than for the technical condition (81%, $SE = 2.7\%$) was not significant in the ANOVA by participants ($F_1 < 1$), nor was the interaction with category. Estimated power for detecting a difference of 10% or greater between condition means was 80%. The only reliable effect was a main effect of category [$F_1(7,266) = 4.83$, $p < .001$; $F_2(7,184) = 5.18$, $p < .001$]. Extreme responses were more common for biological categories (87%) than for the others (79%)—a result consistent with earlier research (Estes 2003; Kalish, 1995). Table 4 also shows the number of participants who gave extreme responses to *all* the items in a category. Five of the 40 participants never used graded responses, 3 in the pragmatic and 2 in the technical condition. In summary, there was no evidence that a technical context led to the perception that the categorizations were less graded.

Discussion

Once again our predictions were unsupported. When participants adopted a technical context for classification, they were, if anything, *less* likely to treat the categorization of individual items as all-or-none than when giving a pragmatic classification. If technical contexts promote categorization based on deeper theoretical considerations, this result is very difficult to explain. Choice of a graded categorization response is generally interpreted as indicating similarity-based classification, in which items may fit into a class more or less well (Estes, 2003; Kalish, 1995).

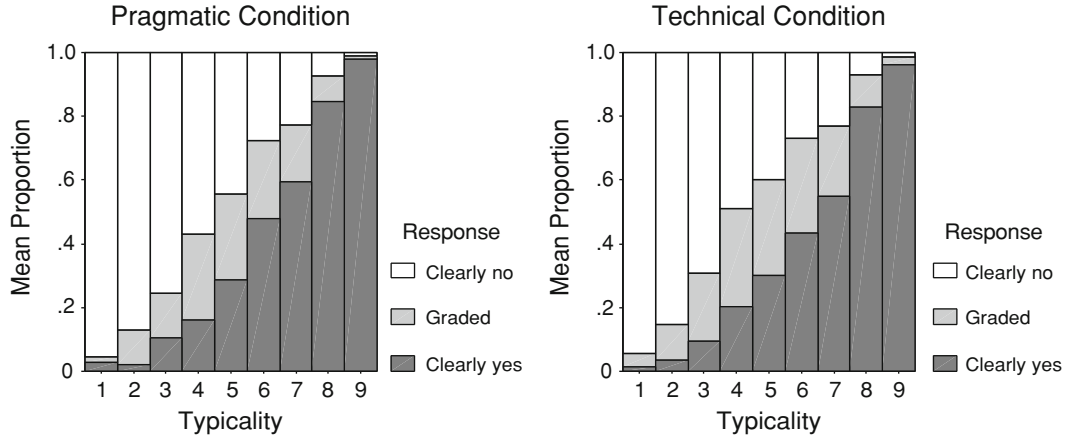

Figure 3. Stacked bars representing the distribution of responses for each condition in Experiment 3, between clearly "yes," clearly "no," and graded (1–7) categorization responses.

Table 4
Percentage of Extreme Responses and Number of Participants (Out of 20) Giving All Extreme Responses in Experiment 3, As a Function of Condition and Category

	Pragmatic		Technical	
Category	Percentage	Number	Percentage	Number
Vegetable	84	5	84	7
Fruit	89	7	90	10
Fish	89	13	79	6
Insect	93	10	87	10
Sport	79	6	70	5
Science	86	9	82	7
Tool	82	8	78	3
Furniture	76	7	80	7
Mean	85	8.1	81	6.9

It is striking that there was no significant difference in the use of graded responses between the pragmatic scenarios (where it would be reasonable to use graded categorization) and the technical scenarios (where the use of the categorization for trade or commercial regulations suggests an all-or-none categorization needs to be used).

GENERAL DISCUSSION

The results of the three experiments described here provide no support for either the hypothesis that the vagueness of everyday categorization reflects contextual ambiguity in the purpose for which the classification is being made or the contention that people switch to a deeper causal-explanatory basis for categorization when asked to consider categories from a more technical perspective. On the contrary, it appears that similarity, as indexed by context-free typicality judgments, provides a powerful predictor of categorization probability across the groups in all three of the categorization conditions used. Asking participants to pay more explicit attention to the scenarios (Experiment 2) did have an effect on the results, but one that showed no sign of bringing them into line with the expected effects. Indeed, in the technical condition of Experiment 2, the amount of disagreement between participants actually increased relative to Experiment 1, and the correlation with typicality was also stronger. Measuring the degree to which people made all-or-none as opposed to graded categorizations (Experiment 3) likewise showed no evidence that they were more likely to use all-or-none classification in technical than in pragmatic contexts. In fact, throughout the three experiments, it was the technical contexts that tended to show the closest relation between categorization probability and context-free typicality.

What conclusions may be drawn from these experiments? First, the notion of a robust "default" conceptual representation for the semantic categories used here seems to gain considerable support. The pattern of categorization probability changed relatively little as a function of different categorization scenarios, suggesting that people were using a similar concept representation in each case. There were also some notable effects of condition for particular categories. In Experiments 1 and 2, the pragmatic condition (which used a department store scenario for tools and furniture) generated a different pattern of data for those categories. Even though the classification task referred to creating a stock index rather than to the actual layout of items in the store, the existence of this prior system of classification proved too tempting to ignore, and participants tended to classify tools and furniture in terms of what would be found in the corresponding departments within the store. Roth and Shoben (1983) demonstrated a similar effect of situational context on category structure. However, whereas those researchers showed that *typicality* within a category shifted with context, our department store effect was reflected in *yes–no* categorization of borderline items.

Context also had a systematic effect on the classification of sports in which typicality was highly predictive of categorization when the categorizing was done for the purposes of a library index, but in other contextual conditions more weight was placed on skill and less on physical effort when classifying activities as sports. This result, however, did not apparently relate to ambiguity in the meaning of the concept "sport." If the context had provided disambiguation, there should have been less disagreement and greater consistency between the context and no-context conditions. But this was not the case. It could perhaps be argued that the no-context condition allowed participants to recruit the same default context, whereas the scenarios may have themselves been ambiguous, leading to greater disagreement and inconsistency in the context conditions. Although this is of course possible, it does not fit easily with explaining vagueness in the absence of context in terms of contextual ambiguity. If the account is unable to predict when contexts will increase vagueness and when they will decrease it, the explanation is empty.

The overall pattern of our results clearly indicates a failure to find evidence supporting our hypotheses. As such, the reader may feel that we have done little but fail to reject the null hypothesis. It is worth pointing out, therefore, the great contrast between the relatively small and subtle effects of the manipulations of purposive context attempted here and the large and robust effects of other manipulations of context. When concepts are placed in a situational context—birds on the seashore or vehicles seen from a farmer's point of view—a major restructuring of the typicality of category members takes place (Barsalou, 1987; Roth & Shoben, 1983). Barsalou found very low correlations between the typicality ordering of category items when very different points of view were adopted. Likewise, Medin et al.'s (1997) tree experts generated completely orthogonal structures for classifying trees, depending on the domain of their expertise. We believe that similarly large effects would be observed in our experiments if the scenarios had highlighted one particular subset of items over another. For example, if the task had been to consider what activities should count as sports, in the context of a foundation whose aim was to promote public fitness and health, it would be easy to imagine that activities that meet this need (e.g., disco dancing) would be more likely to be included, and those that do not (e.g., chess) would be more likely to be excluded. However, this was not our aim.

Our primary aim was not only to demonstrate context effects per se, but to test a hypothesis concerning category vagueness—namely, that at least a part of the disagreement and instability observed in categorization results from a failure to provide a context indicating the purpose or function of a classification. We approached this hypothesis by designing scenarios that, taken at face value, presented participants with very different ways of conceiving of the purpose of the categorization. Although we deliberately did not provide a strong bias toward any particular subset of the category or any particular feature of category members, we made the manipulation of the difference between our two types of scenario as strong as we could. For example, there is a very clear difference between deciding whether books on economics should be placed in a library index under the general category "science" and deciding whether a national funding body with responsibility for the support of science should be giving grants for research in economics. One purpose, prima facie, calls for a pragmatic, similarity-based approach, but the other calls for the construction of a classification rule that would need to be defended and justified.

In our second experiment, we took steps to ensure that participants took the scenarios seriously by requiring them to talk about the task for a minute before starting their classification. In the third experiment, we tried using a different dependent variable to measure the degree to which participants saw the classification as clear-cut or vague. In none of our experiments, however, did we find a lack of purposive context contributing to vagueness in categorization in the standard, context-free task, as had been observed in countless previous studies in this field. The scenarios were direct and easy to follow, but they had no effect on consistency. Nor were our experiments underpowered. The fact that subtle effects such as the pragmatic context effect on tools and furniture were identified indicates the power of these experiments to detect effects. Indeed, there were significant effects of context (Experiment 2), but they did not occur in the predicted direction. The results were consistent across categories and experiments, and the high correlations observed between typicality and categorization probability are another indicator of the low error variance in these data. Our measures were accurate, and they revealed that instructional context has little detectable effect on categorization probability.

Taken as a whole, our results strongly suggest that there is a common default way of representing conceptual categories and of making category decisions. No matter whether the classification was being created for a tax regulation or for a newsgroup search index, the same underlying pattern of categorization probabilities emerged, and the same degrees of vagueness and instability were observed at the category boundary. Context effects in categorization can be readily demonstrated in paradigms in which the context invokes a situation that particular subsets of the category are strongly associated with (Barsalou, 1987; Roth & Shoben, 1983). Birds in a farmyard have a different graded typicality structure from birds on the seashore, but they are all still birds. Effects on *yes–no* categorization are harder to demonstrate.

One of the only successful demonstrations of such effects is the research by Braisby and Franks (2000) described in the introduction. They deliberately created borderline cases that either shared only appearance, with a category (e.g., a plastic flower or an Easter egg) or shared only essence, (e.g., a dried flower or a scrambled egg). In their study, the weight given to surface appearance, as opposed to underlying essence, in categorizing instances was shown to change as a function of a number of factors affecting the perspective of word use. Our results suggest that there are limits to the generality of these effects. First, vagueness in categorization was not reduced by providing purposive contexts. Second, there was no shift away from similarity-based categorization when more serious consequences could arise from the classification. It is possible, therefore, that the effects described by Braisby and Franks (2000) are primarily related to word use—when should you use the word *egg* to refer to a chocolate egg or the word *chicken* to refer to a rubber chicken?

We see our results as contributing to the general debate about the stability versus the context dependence of conceptual representations (Barsalou, 1987). At least with regard to changing purposive contexts, concepts appear to be remarkably stable.

REFERENCES

Ahn, W.-K., & Dennis, M. J. (2001). Dissociation between categorization and similarity judgement: Differential effect of causal status on feature weights. In U. Hahn & M. Ramscar (Eds.), *Similarity and categorization* (pp. 87-107). Oxford: Oxford University Press.

Ahn, W.-K., Kim, N. S., Lassaline, M. E., & Dennis, M. J. (2000). Causal status as a determinant of feature centrality. *Cognitive Psychology*, **41**, 361-416.

Barr, R. A., & Caplan, L. J. (1987). Category representations and their implications for category structure. *Memory & Cognition*, **15**, 397-418.

Barsalou, L. W. (1987). The instability of graded structure: Implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*. Cambridge: Cambridge University Press.

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).

Bloom, P. (1996). Intention, history, and artifact concepts. *Cognition*, **60**, 1-29.

Braisby, N. (1993). Stable concepts and context-sensitive classification. *Irish Journal of Psychology*, **14**, 426-441.

Braisby, N., & Franks, B. (1997). What does word use tell us about conceptual content? *Psychology of Language & Communication*, **1**(2), 5-16.

Braisby, N., & Franks, B. (2000). *Perspectives and categorization*. Unpublished manuscript.

Braisby, N., Franks, B., & Harris, J. (1997). Classification and concepts: Fuzzy or perspectival? In D. Dubois (Ed.), *Catégorisation et cognition: De la perception au discours* (pp. 163-189). Paris: Kimé.

Estes, Z. (2003). Domain differences in the structure of artifactual and natural categories. *Memory & Cognition*, **31**, 199-214.

Hampton, J. A. (1995). Testing the prototype theory of concepts. *Journal of Memory & Language*, **34**, 686-708.

Hampton, J. A. (1998). Similarity-based categorization and fuzziness of natural categories. *Cognition*, **65**, 137-165.
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.
Kalish, C. W. (1995). Essentialism and graded membership in animal and artifact categories. *Memory & Cognition*, **23**, 335-353.
Keefe, R., & Smith, P. (1996). Introduction: Theories of vagueness. In R. Keefe & P. Smith (Eds.), *Vagueness: A reader* (pp. 1-57). Cambridge, MA: MIT Press.
McCloskey, M. E., & Glucksberg, S. (1978). Natural categories: Well defined or fuzzy sets? *Memory & Cognition*, **6**, 462-472.
Medin, D. L., Lynch, E. B., Coley, J. D., & Atran, S. (1997). Categorization and reasoning among tree experts: Do all roads lead to Rome? *Cognitive Psychology*, **32**, 49-96.
Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.
Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289-316.
Osherson, D., & Smith, E. E. (1997). On typicality and vagueness. *Cognition*, **64**, 189-206.
Rey, G. (1983). Concepts and stereotypes. *Cognition*, **15**, 237-262.
Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). Cambridge: Cambridge University Press.
Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, **7**, 573-605.
Roth, E. M., & Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology*, **15**, 346-378.
Smith, E. E., & Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition*, **22**, 377-386.
Weatherburn, C. E. (1961). *A first course in mathematical statistics*. Cambridge: Cambridge University Press.

APPENDIX A
Items for Categorization Used in Experiments 1–3

Vegetable	Fruit	Fish	Insect	Sport	Science	Tool	Furniture
spinach	strawberry	trout	ant	swimming	chemistry	axe	chair
celery	orange	salmon	wasp	tennis	mathematics	screwdriver	table
lettuce	banana	sardine	mosquito	skiing	astronomy	rake	bed
asparagus	watermelon	catfish	grasshopper	surfing	medicine	hammer	lamp
potato	pomegranate	goldfish	termite	jogging	meteorology	shovel	desk
artichoke	tomato	shark	silkworm	croquet	psychology	sewing needle	rug
soybean	avocado	eel	caterpillar	billiards	nutrition	funnel	television
watercress	pumpkin	sea horse	moth	ballroom dancing	geometry	scalpel	shelf
parsley	olive	squid	dust mite	frisbee	sociology	pitchfork	bookends
dandelion	coconut	shrimp	head lice	wrestling	mineralogy	calculator	curtains
seaweed	cucumber	jellyfish	maggot	darts	economics	dictionary	waste basket
bamboo shoot	eggplant	lobster	scorpion	hunting	geography	tractor	dishwasher
chili pepper	acorn	starfish	centipede	bullfighting	dentistry	toothbrush	cushion
cloves	almond	clam	spider	weightlifting	pharmacy	broom	door mat
garlic	walnut	crab	tarantula	aerobics	architecture	scissors	painting
sage	date	tadpole	snail	fishing	archaeology	key	ashtray
apple	pine cone	whale	earthworm	mountaineering	agriculture	varnish	telephone
turnip	rhubarb	seal	leech	hiking	criminology	screw	refrigerator
peanut	sugar beet	plankton	tapeworm	bridge	astrology	string	piano
bread	carrot	alligator	lizard	kite flying	literature	umbrella	suitcase
pineapple	mushroom	oyster	bat	conversation	advertising	photograph	plate
milk	ginger	sponge	hamster	chess	palm reading	trunk	bucket
rice	onion	gull	amoeba	crosswords	religious studies	pen	pillow
cereal	mint	frog	bacterium	picnicking	philosophy	stone	book

APPENDIX B
Examples of Pragmatic and Technical Context Scenarios Used in All Experiments

Technical Scenario for Vegetables

The National Administration of Food and Agriculture is planning to regulate the growth of various kinds of agricultural produce, so that the quality of the produce in the market can be monitored. Imagine that you belong to a panel of advisors for the Administration of Food and Agriculture to provide help in drafting the regulation. In the chapter for **vegetables**, you want to include all produce that should be considered as vegetables, excluding other kinds of agricultural produce, which would be covered under other chapters. Because the regulation affects vegetable farmers nationwide, a clear categorization of vegetables will thus ensure a fair and reasonable regulation.

Consider each of the following items, and decide whether acting in the panel of advisors, you would classify the item in the category of **vegetables**.

Pragmatic Scenario for Furniture

Klein, which is a department store, is designing a sorting system to list the items in stock, as well as their prices and quantities. Marketing persons in the store can quickly obtain information about these items by using such a system. Imagine that you work for the department store to develop the sorting system. You have to categorize selling items under different headings, so that marketing persons can search for information easily and quickly. The following is a list of items from the department store. You have to decide whether or not they should be included in the category of **furniture**, so that most marketing persons would be able to find things under the category heading where they expected to find them.

(Manuscript received July 5, 2004;
revision accepted for publication August 8, 2005.)