COGNITIVE PSYCHOLOGY 27, 41–70 (1994)

Water Is Not H_2O

BARBARA C. MALT

Lehigh University

What makes a liquid water? A strong version of "psychological essentialism" predicts that people use the presence or absence of H_2O as the primary determinant of what liquids they call "water." To test this prediction, subjects were asked to judge the amount of H_2O in liquids called "water" and liquids not called "water." Neither their beliefs about the simple presence/absence of H_2O nor about the proportion of H_2O in the liquids accounted well for which ones are normally called "water." Typicality ratings and an extended tree solution on similarity ratings suggested that use, location, and source of a liquid may also influence whether it is considered to be water. Sentence acceptability judgments further suggested that there may be a sense of "water" that corresponds to the strong essentialist view, but that there is also a more general sense in common use encompassing mixtures with varying amounts of H_2O. These findings indicate that essentialist beliefs alone may not fully explain category membership judgments and word use, and they suggest a modified version of psychological essentialism.

What is water? Most people, if asked this question, would probably respond something like "It's a clear, odorless liquid found in rivers and lakes, which humans and other animals drink." But the philosopher Hilary Putnam argued in an influential paper (Putnam, 1975) that such responses only describe what water is typically like; they do not constitute the meaning of the word "water." According to Putnam, "water" and other natural kind terms (e.g., "tiger," "gold," "lemon") refer to sets of entities that share an "essence" or hidden underlying trait such as a particular atomic weight, genetic structure, or chemical composition. In the case of water, its essence (to the best of current scientific knowledge) is H_2O.

This research was supported by NSF Grant BNS-8909360. I thank Yongming Gao for extensive assistance with programming and data collection in Experiments 1–4, Bethann Schultz and Christopher Santi for assistance with data collection for Experiment 3, Donna Martin for assistance with data collection for Experiment 4, and James Corter for carrying out the scaling solution in Experiment 3. I also thank the Beckman Institute and the Department of Psychology, University of Illinois, for their hospitality during the completion of this paper. Earlier drafts of this paper benefitted from comments by Susan Barrett, Raymond Gibbs, Gregory Murphy, and Padraig O'Seaghdha. Reprint requests should be addressed to B. C. Malt, Department of Psychology, 17 Memorial Drive East, Lehigh University, Bethlehem, PA 18015.

0010-0285/94 $9.00

Putnam's approach takes as fundamental the idea that word meanings are not a matter of mental representation; they are a matter of truths about the world. However, Putnam also assumes that people use words in a way consistent with this claim: According to Putnam, people would only consider a clear, odorless liquid found in a lake to be water if it were composed of H_2O. If discovered to actually be some other compound, they would not want to call it water. Although psychologists have avoided commitment to Putnam's general approach to word meaning, many have recently adopted a position derived from the more psychological parts of his claim. They argue that people *believe* there are essences shared by sets of things, and this belief plays a role in determining what things will be accepted as members of a category and called by the category name (e.g., Carey, 1985; Keil, 1989; Malt, 1990; Medin & Ortony, 1989; Neisser, 1987). Thus, whether or not there really is an essence shared by all the things normally called "water" or by any other natural kind term, people may believe there is, and they may take into account whether they believe something has the appropriate essence when they decide whether the thing is a member of a particular category.

This approach helps provide a resolution to the dilemma faced by traditional approaches to word meaning and conceptual structure: It has been impossible to find explicit necessary and sufficient conditions for membership in most common categories (see, e.g., Smith & Medin, 1981), but it has also become clear that pure family resemblance accounts are too unconstrained to account for category membership judgments (e.g., Keil, 1989; Murphy & Medin, 1985; Rips, 1989). The essence possibility provides an additional constraint on category membership and word use without requiring that people know sets of necessary and sufficient features.

Support for the existence of a belief in essences comes from several sources. Keil (1989) found that when subjects were given a description of a raccoon that had had surgery to make it look and act just like a skunk, they still considered the animal to be a raccoon. The subjects appealed to a raccoon essence: They believed there was something hidden and internal that made the raccoon stay a raccoon despite superficial changes. Malt (1990) found that subjects felt an animal or a plant with an ambiguous appearance would nevertheless have only one correct name. For instance, when told of a plant that is halfway between a marigold and a dandelion in appearance, they judged that it made more sense to say "We'd have to ask an expert to tell us which it is" than to say "I guess you can call it whichever you want." Again, people seemed to believe there is some hidden property that makes these entities belong to only one category, despite superficial ambiguity. As for Putnam's intuition that the presence of H_2O, not more readily observed properties like being clear and odorless, is a critical factor for willingness to call a liquid "water," this intuition is so widely shared that it is often assumed without further comment (e.g., Wellman & Gelman, 1988, p. 113; Carey, 1985, p. 174).

Neither the laboratory studies, though, nor the intuition that H_2O is a critical factor for determining what liquids are water necessarily provides a complete picture of the role of essentialist beliefs in category membership judgments and word use. People's intuitions about the way their language works are generally poor; for instance, speakers of a language do not have insight into either the syntactic structure or the phonological system of their language. And people's explanations for their own judgments and behaviors are not always complete; factors may influence their judgments that they do not have introspective access to or cannot verbalize (Ericsson & Simon, 1980; Nisbett & Wilson, 1977). Essentialist beliefs elicited in a laboratory task thus may not be the only factor at work in determining classification. Given the available evidence, the role that beliefs in an essence play in category membership judgments and word use could be either large or small.

The case of water provides an ideal testing ground for investigating the role of essentialist beliefs, since there is a general consensus both that belief in an essence should be relevant for water, and that the particular essence involved is the chemical composition, H_2O. The studies presented below test a strong version of psychological essentialism that takes belief in an essence to be the primary determinant of category membership and word use. Under this version, if a person believes a liquid has the appropriate essence, H_2O, he or she will consider it to be water; if the person believes it doesn't, he or she won't consider it to be water. This position is most true to the spirit of Putnam's original analysis, and it is at least implicit in many discussions of psychological essentialism that draw heavily on Putnam's analysis (e.g., Carey, 1985; Keil, 1989). The studies presented below indicate that the liquids people actually consider to be water are only partly determined by a belief about the presence of H_2O. They also provide evidence for the importance of several dimensions other than H_2O in determining what is water. These data argue against the strong version of psychological essentialism, and they suggest a modified view of the role of essentialist beliefs in category membership and word use.

EXPERIMENT 1: BELIEFS ABOUT H_2O

Given the strong version of psychological essentialism, there are three possible interpretations of how the liquids considered to be water could depend on a belief in the presence of H_2O. The most restrictive possibility is that people believe all water is pure H_2O. Under this interpretation, any liquid that people believe to contain ingredients other than H_2O will not be considered to be water and will be called by another name. This possibility does not, of course, entail that all the liquids commonly considered to be water actually *are* pure H_2O, only that people think of them as such. The second interpretation is that people realize many liquids containing H_2O are not pure, but they consider a liquid to be water whenever they believe that H_2O is the dominant ingredient of the liquid. Any liquid they believe to be less than 50% H_2O will therefore be called by another name. Putnam (1988), in fact, makes essentially this suggestion about the use of the word "milk"; he suggests that people recognize that milk may vary in composition somewhat, but if a liquid is not composed of at least 50% of the normal milk ingredients, they will not want to call it "milk." The third, and loosest, interpretation is that people do not necessarily restrict "water" to liquids that they believe have H_2O as the majority ingredient. They do, however, have some threshold for what is acceptable as water, so that the liquids they call "water" are ones they believe have a larger amount of H_2O than liquids they do not call "water." In other words, people may simply distinguish between liquids they believe have "a lot" of H_2O and those they believe have "a little." Under this interpretation, the cutoff for calling a liquid "water" could be anywhere, but the amount of H_2O believed to be present for waters and non-waters should not overlap.

To test the three possibilities just outlined, a large set of examples of liquids that are normally called "water" was compiled. A large set of examples of liquids that are similar to water in one or more ways (e.g., in appearance, odor, potability, etc.) but that are not normally called "water" was also compiled. The sets of examples themselves provide certain clues about the role of H_2O in determining what is considered to be water. To provide a more definitive picture of beliefs about the presence of H_2O and how the beliefs relate to what liquids people call "water," college students were also asked for explicit judgments of the amount of H_2O in each liquid on the lists.

Part 1: Examples of Water and Non-water

Method

Examples of water were collected from three sources. The first was a computer search of the Brown text corpus, source for the Kucera and Francis (1967) word frequency counts. The corpus contains 500 two-thousand-word samples of American English prose from novels, newspaper articles, textbooks, etc. The texts were searched for each occurrence of the word "water." The kinds of water referred to were recorded, as determined either by presence of a modifier-plus-noun construction (e.g., "lake water," "rain water") or by the context available in the 3 lines above and below each occurrence (e.g., "water" in a report on the quality of water in a lake would be recorded as "lake water").

The second source of examples was a laboratory task designed to encourage subjects to think of examples of water other than the most obvious, easily retrieved ones. Fourteen

Lehigh undergraduates were asked to imagine that they were playing a word game in which the goal was to guess the name of the type of water that their challenger had in mind. They were told that the trick was that their challenger was not thinking of an easy, obvious example; therefore, they would have to try to think of less obvious examples. As a hint, they were told that their challenger was NOT thinking of tap water or ocean water. Subjects generated the examples for water as part of a larger task in which they also generated examples for five other common English nouns. Since subjects sometimes generated responses that did not seem to be sincere examples of water, another group of 15 subjects was asked to read each response and evaluate whether it was truly an example of water. Selection was restricted to those that received at least 10 "yes" responses.

The third source of examples was observation by the author and a graduate assistant of uses of the word "water" in everyday conversations, newspaper, television, etc. Again, if the instance did not occur with a modifier attached, the type of water was determined from context.

Examples of liquids similar to water but not called "water" (henceforth referred to as examples of "non-water") came from two sources. The first was a laboratory task similar to that for generating examples of water. Fourteen Lehigh undergraduates were asked to imagine that they were playing a word game, and that the goal was to guess the name of the thing their challenger had in mind. In this version, they were told that the trick was that the thing was similar to water, but was not actually called "water." As a hint, they were told that their challenger was NOT thinking of vodka or lemonade. Subjects generated the examples as part of a larger task in which they generated examples for five other common English nouns. Since subjects sometimes generated responses that did not seem to be sincere examples of non-water, another group of 15 subjects was asked to judge the similarity of each example to water. Selection was restricted to those that received mean ratings of at least 3.00 on a 7-point scale and that did not name a type of water.

The second source of non-examples was informal observation. Since there was no prespecified target word, observation consisted of looking for use of names of liquids similar to water in conversations, etc., and also searching drug, grocery, and other store shelves for examples of such liquids.

Results and Discussion

The final list of waters consisted of 43 liquids judged by the author to have "water" as their most common, everyday label. Examples having other dominant labels (e.g., "body water," more commonly known as "sweat" or "perspiration") were eliminated. About half the examples came from the computer search and half from the laboratory task, with one additional example from observation. The final list of 55 non-waters consisted of seven examples collected in the laboratory task and 48 collected by observation.

The examples of waters and non-waters are given in Table 1. These examples strongly suggest that "water" does not refer only to liquids that people believe to be entirely H_2O. The fact that the set of common water types obtained includes liquids referred to as "pure water," "purified water," and "distilled water" suggests in itself that there are waters believed to contrast in composition with these pure forms. The names for the other waters further suggest that they do not all refer to liquids that people see as exactly the same substance. Although some of the names

TABLE 1
Water and Non-water Examples and Mean Judged Percentage of H_2O

Water example	Mean judged percentage of H_2O
Pure water	98.1
Purified water	94.8
Natural spring water	92.6
Bottled water	92.3
Rain water	90.9
Ice water	90.4
Soft tap water	89.9
Drinking water	89.4
Fresh water	89.1
Water fountain water	88.8
Reservoir water	88.0
Distilled water	87.9
Tap water	87.7
River water	87.7
Well water	87.7
Waterfall water	86.9
Stream water	86.4
Water from melted snow	86.3
Babbling brook water	85.6
Mountain water	85.5
Hard water	85.0
Mineral water	85.0
Bath water	83.2
Lake water	83.0
Salt water	82.7
Ground water	82.3
Flood water	82.3
Pool water	81.6
Soapy water	81.4
Swimming pool water	81.0
Stagnant water	80.4
Lagoon water	80.0
Puddle water	79.6
Pond water	78.8
Ocean water	78.7
Chlorinated water	78.1
Dish water	77.1
Polluted water	70.6
Muddy water	70.3
Unpurified water	69.2
Swamp water	68.8
Radiator water	67.3
Sewer water	67.0

TABLE 1—*Continued*

Non-water example	Mean judged percentage of H_2O
Tea (cup of)	91.0
Saliva	89.3
Coffee (cup of)	89.1
Tears	88.6
Sweat	87.3
Lemonade	86.9
Chicken broth	81.3
Saline solution for contact lens	80.1
Urine	79.1
Cranberry juice	76.9
Pickle juice	76.9
Apple juice	76.7
Clam juice	76.5
Eye drops	76.3
Mouthwash	75.8
Grapefruit juice	75.8
Gingerale	73.9
Windshield wiper fluid	73.9
Juice from canned pineapple	72.3
Sprite	72.0
Onion juice	71.5
Windex	70.1
Contact lens cleaner	69.1
Blood	68.7
Radiator coolant	67.8
Ice milk	66.8
Cologne	66.4
Cider	65.1
Soy sauce	64.0
Garlic juice	64.0
Skim milk	63.9
Steak juice	63.9
Disinfectant	63.2
Fantastik spray cleaner	61.8
Hydrogen peroxide	61.7
Syrup from canned fruit cocktail	61.3
Cough syrup	60.6
Antiseptic face lotion	59.5
Brake fluid	58.5
Plasma	58.4
White vinegar	58.0
Bug repellant	55.0
Transmission fluid	54.2
Corn syrup	52.8
Aftershave lotion	52.6
Bleach	51.3

TABLE 1—*Continued*

Non-water example	Mean judged percentage of H_2O
Creme rinse (hair conditioner)	50.7
Vodka	48.5
Tree sap	48.2
Nail polish remover	46.8
Lighter fluid	42.3
Cool Whip	41.8

could be interpreted as describing location or use (e.g., spring water, mountain water, dish water), others can only be interpreted as referring to ingredients in a mixture (e.g., salt water, soapy water, mineral water, chlorinated water, polluted water, muddy water). Even if the average speaker of English uses some of the names for water by rote, without ever reflecting on their ingredients, it is unlikely that they do so for examples that are familiar household mixtures. The contents of chlorinated water are known to everyone who owns a swimming pool, and the difference between mineral water and tap water is known to everyone who drinks Perrier. Thus the examples indicate that people do not hold a belief, at least for the more familiar members of the lists, that water consists only of H_2O.

Part 2: Judgments of H_2O

The observations above address the most restrictive interpretation of how a belief in an essence could constrain what is considered to be water. They do not address either the possibility that people call "water" only those liquids that they believe have H_2O as the dominant substance, or the possibility that they call "water" only those liquids that they believe have some threshold amount of H_2O, even if not above 50%. To obtain additional evidence about people's beliefs about the presence of H_2O in the liquids on the lists, and the relationship of those beliefs to use of the word "water," subjects were next explicitly asked to judge the H_2O content of the liquids.

Method

Subjects. Forty-six Lehigh undergraduates participated for course credit.

Stimuli. Stimuli were the 43 examples of water and 55 examples of non-water collected in Part 1. Two random orders of each set were constructed.

Procedure. Half of the subjects received the water examples to judge, and half received the non-water examples. This division was to prevent subjects from using the labels as the basis for their responses. If subjects received both, they might simply rate all the liquids with "water" in their name as high in H_2O and all those without "water" as low.

Subjects who received the water examples were given instructions drawing their attention to the fact that liquids may be composed of more than one substance. They were reminded that, for instance, rubbing alcohol actually contains both pure alcohol and water. They were then asked to carefully consider each example of water listed and judge what percent of it is actually H_2O.

Subjects who received the non-water examples were given parallel instructions pointing out that liquids may contain more than one substance, but the instructions were modified to draw their attention to the fact that liquids may often contain a substance even though their name does not directly reflect the presence of that substance. Subjects were asked to carefully consider each liquid listed and judge what percent of it is actually H_2O.

Subjects were randomly assigned to the water or non-water group. They responded by writing a percentage from 0 to 100 in the space next to each example. They were run singly or in small groups and completed their judgments at their own pace.

Results and Discussion

The mean judged percentages of H_2O for the waters and non-waters are given in Table 1. Three non-waters received responses from less than 15 subjects and were eliminated from the data.

Consistent with evidence from Part 1, it is clear that people believe liquids called "water" can have ingredients other than H_2O. Even "pure water" was judged to be only 98% H_2O, and "purified water" and "distilled water" were judged to be 95 and 88% H_2O, respectively. Only five other waters were judged to be 90% or more H_2O; the remaining 35 examples were judged 89% or less H_2O.

What about the possibilities either that H_2O must be the dominant ingredient, or that it must be at least always higher in proportion than in liquids not called "water?" The overall judged percentage of H_2O was somewhat lower for non-waters than for waters, with non-waters judged to have an average of 67% H_2O and waters an average of 83%. This observation might initially suggest that belief in something about the percentage of H_2O does underlie the distinction. However, closer examination shows that the judgments of H_2O are not consistent with either interpretation of the essence view.

With respect to the first possibility, judgments do show that liquids called water are uniformly believed to have H_2O as their dominant ingredient; the range is from 67 to 98%. At the same time, though, the judgments reveal that this fact cannot account for what liquids are considered water and what ones are not. Although the average judged percentage of H_2O is lower for non-waters, it is nevertheless over 50%, and the great majority of individual non-waters (47 out of 52) were judged to have H_2O as their dominant ingredient. If liquids were considered water whenever their dominant substance is H_2O, all but five of the non-waters would be considered water instead.

With respect to the second possibility, the judgments do show a higher average percentage for waters than for non-waters, as already noted. Again, though, it is clear that this fact is not the whole story. Despite the higher average for waters, there is substantial overlap for individual waters and non-waters. Some waters are judged lower in H_2O than many non-waters, and some non-waters are judged higher in H_2O than many waters. The ranges illustrate this overlap: The range for non-waters is from 42 to 91%, and for waters it is from 67 to 98%. Among the most striking individual low judgments for "water" are those for sewer water, radiator water, swamp water, unpurified water, muddy water, polluted water, and dish water, which were judged to have only 67, 67, 69, 69, 70, 71, and 77% H_2O, respectively. For non-waters, the most striking high judgments include tea, saliva, coffee, tears, sweat, and lemonade, which were judged to have 91, 89, 89, 89, 87, and 87% H_2O, respectively. This overlap indicates that waters are not simply liquids that people believe have some sufficiently high proportion of H_2O while non-waters are liquids that people believe have somewhat lower proportions. If this were so, a number of liquids called "water" should be called by another name, and a number of non-waters should be called "water."

These results clearly indicate that beliefs about the presence of H_2O do not by themselves fully account for what liquids people consider to be water.[1] If not, though, what other factors may be contributing to whether a liquid is considered to be water? Experiments 2 and 3 were aimed at investigating the concept of water in order to help answer this question.

EXPERIMENT 2: TYPICALITY

As one source of information about the concept of water, typicality ratings were collected for the 43 examples of water used in the previous two experiments. Typicality rating tasks ask subjects to judge what examples are the best examples of a concept; the best examples are generally considered to be those most central to the concept and to reflect the concept prototype (e.g., Rosch & Mervis, 1975; Smith & Medin, 1981). Thus, if people's concept of water centers on the idea of H_2O, those

[1] As noted under Methods, it was necessary to have separate groups of subjects rate the waters and non-waters so that they could not use presence or absence of the word "water" as the basis for a judgment about the amount of H_2O. It is possible as a result that there was some slight adjustment of the range of judgments given by each group. However, subjects clearly did not feel compelled to use the entire scale in their judgments: The bottommost water examples were rated substantially higher than the bottommost non-waters, and the topmost non-water examples were rated somewhat lower than the topmost water examples. Small adjustments in range would not have been sufficient to change the average of substantially over 50% for non-waters to under 50%. Most importantly, small adjustments could not have created the extensive degree of overlap of judged percentages for waters and non-waters. It is highly unlikely that subjects would have preferred to rate the topmost non-water examples (e.g., tea, rated as 91% H_2O) as lower than the bottommost water examples (e.g., sewer water, rated as 67% H_2O). Only extreme adjustments of this sort would falsify the conclusion of substantial overlap between water and non-water judgments.

examples that are believed to contain the most H_2O should be judged most typical.[2] If the concept of water involves other important factors in addition to H_2O, then these factors should also be reflected in the examples that are considered to be most typical.

Method

Subjects. Twenty-six Lehigh undergraduates participated for course credit.

Stimuli. Stimuli were the 43 examples of water from the previous experiment.

Procedure. Subjects received the 43 exemplars in one of two random orders, along with similar lists for five other common English nouns as part of a larger study. They were given standard typicality rating instructions (e.g., Rosch & Mervis, 1975; Malt & Smith, 1984) asking them to rate each example on a scale of 1 (low) to 7 (high) according to how good or typical an example of the specified concept each was. They were told not to rate any example that was unfamiliar to them. Subjects responded by placing a rating in the blank next to each item.

Results and Discussion

Three subjects did not rate "babbling brook water" and one each did not rate "ground water," "hard water," and "lagoon water." All other examples were rated by all subjects. The average rating for each example is given in Table 2.

There is an overall positive correlation between judged H_2O from Experiment 1 and typicality rating, $r = .69, p < .01$, indicating that typicality is influenced by purity of the liquid. However, the proportion of variance accounted for by this correlation is only .48, indicating that other factors are also contributing to typicality. Inspection of Table 2 clearly indicates that the waters people believe are closest to pure H_2O are not necessarily those they will consider to be the best examples of the concept of water. The three waters with the highest judged percentage of H_2O in Experiment 1 (pure, purified, and natural spring water) are ranked 8th, 5th, and 15th in typicality. Conversely, the three with the highest typicality ratings (drinking, tap, and rain water) were ranked 8th, 5th, and 13th in the judgments of percentage H_2O in Experiment 1.

If not only a high H_2O content, then what else may contribute to making the highest rated waters so central to the concept of water? It is striking that the most typical one, drinking water, is the one most important for human survival and is also the most familiar and frequently encountered in urban American life. The other highly rated waters share these two characteristics to a large extent. Tap water, bottled water, and ice water

[2] If people believed simply that all water is pure H_2O, there would be no reason for a range of typicality to exist; any example of water would be as good as any other. However, the results of Experiment 1 show that people do not have this belief. Given that they believe that different examples have differing amounts of H_2O present, a range of typicality can be expected.

TABLE 2
Typicality Ratings for Examples of Water

Example	Rating
Drinking water	6.5[a]
Tap water	6.5
Rain water	6.1
Water fountain water	6.1
Purified water	6.0
Bottled water	6.0
Ice water	6.0
Pure water	5.8
Ocean water	5.8
Soft tap water	5.8
Mineral water	5.7
Distilled water	5.7
Salt water	5.7
Fresh water	5.7
Natural spring water	5.6
Pool water	5.6
River water	5.6
Swimming pool water	5.4
Stream water	5.3
Lake water	5.2
Reservoir water	5.2
Pond water	4.9
Bath water	4.9
Puddle water	4.7
Waterfall water	4.7
Well water	4.6
Dish water	4.5
Soapy water	4.4
Chlorinated water	4.2
Mountain water	4.2
Water from melted snow	4.0
Polluted water	3.8
Sewer water	3.8
Unpurified water	3.8
Flood water	3.7
Muddy water	3.7
Ground water	3.7
Hard water	3.6
Swamp water	3.5
Stagnant water	3.4
Babbling brook water	3.4
Lagoon water	3.4
Radiator water	2.7

[a] 1 = low, 7 = high.

are all typically found in the home; the latter two are used primarily for drinking, and tap water is used both for drinking and for a variety of other household functions. Water fountain water is not usually found in homes, but it is found in other familiar indoor locations such as schools and office buildings and is used primarily for drinking. Rain water is not (usually) found inside buildings, but it is a form of water that people have close and frequent contact with, and that serves valuable functions for them such as watering their lawns and gardens. Finally, subjects may think of purified water as a product of home tap water purification systems and so consider it similar to ordinary tap water. In sum, it seems that the most typical waters are those that occur in the vicinity of human dwellings and play the most central role in human lives, even if they are not the purest examples of H_2O.

Many of the other typicality ratings appear to reflect the same influences. For instance, pool water and swimming pool water received relatively high typicality ratings (5.6 and 5.4, respectively) although they were judged to be relatively low in H_2O content (82 and 81% H_2O, respectively). In contrast, babbling brook water, lagoon water, and mountain water received very low typicality ratings (3.4, 3.4, and 4.2, respectively), despite having been judged as high in H_2O as the pool waters (86, 80, and 86%, respectively). Here again, the two with high ratings relative to their judged percentage of H_2O are waters that occur in human-devised locations, and the three with low ratings relative to their judged percentage of H_2O are waters that are much less frequently encountered by urban Americans and do not serve a direct purpose to them.

Finally, it is noteworthy that ocean water and pond water were among the lowest in judged H_2O (35th and 34th, respectively), but they were substantially higher in rated typicality (9th and 22nd, respectively). They, along with stream, lake, river, and several other waters that flow or occur in bodies, make up the top half of the typicality distribution other than the waters found in and around human dwellings. These waters are probably the ones that were most familiar to people and essential to their survival up until the last 50 to 100 years. Thus these naturally occurring forms of water may be the ones that in the recent past held the same relation to people that tap water, bottled water, etc., do today, and they may be relatively central to the concept for that reason.

The pattern of typicality thus suggests that the waters most central to the water concept are those that have not only a moderate to high amount of H_2O, but are also found in or near the home and that serve important human functions (and, secondarily, are found in nature and in the past served the same functions). To further evaluate whether the importance of a liquid in human lives influences its typicality as water, 20 University of Illinois undergraduates were asked to rate each water example for its centrality in human lives, and another 20 rated the frequency with which humans encounter each water example. Centrality ratings correlated more strongly with typicality than the judged percentage of H_2O did, $r = .72, p < .01$, and they correlated significantly with typicality when judged percentage of H_2O was held constant, partial $r = .56, p < .01$. Centrality ratings substantially improved the ability to predict typicality, with centrality and judged percentage of H_2O together yielding a multiple R of .78, accounting for 62% of the variance. Frequency ratings were less helpful; they were not as strongly related to typicality ($r = .52, p < .01$) and added little predictive power beyond centrality (multiple $R = .79$, $R^2 = .62$ for all three predictors together). The predictive value of the centrality ratings strongly supports the proposal that the role a liquid plays in human lives, in addition to its composition, influences how well it represents the concept of water.

The typicality results thus indicate one general dimension, importance in human lives, that is relevant to the water concept in addition to H_2O content. To obtain more detailed information about the concept of water, a scaling solution for a subset of the water examples was obtained in Experiment 3.

EXPERIMENT 3: THE DIMENSIONS OF WATER

Scaling solutions provide a means of identifying the dimensions that underlie a semantic domain. A scaling solution for water was therefore obtained as a means of confirming or disconfirming the two dimensions suggested by the typicality data, and as a means of seeking more detailed information about the contents of the concept. An extended tree (Corter & Tversky, 1986) was used to represent the structure in the data since this scaling method reveals both nested and non-nested relationships among stimuli.

Method

Subjects. Twenty Lehigh undergraduates participated for course credit.

Stimuli. Since it is not feasible to collect similarity ratings on more than 20 stimuli per subject, 20 examples of water were selected from the 44 used in the preceding experiments. The list was narrowed from 44 to 20 by including only one item from any pair that seemed to be near synonyms. For instance, of "babbling brook water" and "stream water," only "stream water" was included, and of "salt water" and "ocean water," only "salt water" was included. The choice of which member of a pair to exclude was arbitrary. The final list of 20 examples of water was: auto radiator, bath, chlorinated, dish, distilled, flood, lake, mineral, pond, puddle, rain, river, salt, sewer, stagnant, stream, swamp, swimming pool, tap, and well. All 190 possible pairs of the 20 stimuli were formed for presentation.

Procedure. Subjects were instructed that they would be reading pairs of phrases referring to types of water, and that their task was simply to judge how similar the two types of water in each pair seemed. Ratings were made on a scale of 1 (low similarity) to 9 (high similarity). Subjects received three practice trials using examples from the category "animal."

The pairs of phrases were presented on an IBM-compatible microcomputer using the Micro Experimental Laboratory system. Each pair remained on the screen until the subject pressed a key to register a rating number. Subjects were given the option of taking a break halfway through data collection.

Results and Discussion

The extended tree structure obtained is given in Fig. 1. The stress value for the solution is .056 (formula 1) or .297 (formula 2), indicating a good fit of the data to the solution. The proportion of variance accounted for is .8733.

Three major clusters are clearly present in the tree structure. The bottommost contains eight examples of water that can be thought of as rel-

```
                  NNNNNNNNNNNN----------      swamp
              CCC|
        EEEEUU|   ZZZZ-------------------      sewer
IIII--|       |
|     |       DDDDDD-------------     stagnant
|     |
|     -------------------------------     auto radiator
|
|           XXXXXXX---------     stream
|             |
|         ---|   --------------     river
|         |  CCC|
|         |     ZZZZ-------------     flood
|         |
|UU-----|         HHHHHHHHH-     rain
|         |CCC---|
|         |         DDDDDDEEEEIIII-----     puddle
|         |
|         |  NNNNNNNNNNNN-------     pond
|         ---|
|             ---------------     lake
|
|-------------------------------     salt
|
|         HHHHHHHHH------     well
|             |
| OOOOOOO|   XXXXXXX--------     mineral
| |          ----|
| |              ----------------     distilled
| |
| |                  ---------     swimming pool
--|  -------------|
  |  |               -------------     chlorinated
  |  |
  ---|      OOOOOOO---------     tap
     |          |
     -----|----------------     bath
     |
     -----------------     dish
```

FIG. 1. Extended tree solution for similarity ratings of 20 examples of water. Segments marked by letters indicate features shared by nonadjacent nodes.

atively "tame" or "domestic" waters. These types of water share the properties of occurring in and around the house, being brought into their locations by humans, and being used for drinking and other domestic functions such as washing. Tap water, which is the most typical water from Experiment 3 that was included in the stimulus set, is within this cluster. Its occurrence here is consistent with the fact that the most typical waters in Experiment 2 were ones that can also be thought of as "domestic." The middle cluster consists of types of water that can be thought of as more "wild," in that they share the properties of occurring outdoors, in natural settings, and being deposited in those settings by natural forces. Finally, the third cluster seems to consist of waters that are neither exclusively "tame" nor exclusively "wild," but that share the property of an especially large proportion of ingredients other than H_2O that make them "dirty" or unhealthy to humans. These major clusters thus suggest four dimensions central to the water concept. The first is location (domestic or in the wild); the second, immediate source (deposited in that location by nature or by human intervention); the third, function (used domestically or not); and the fourth, relative purity or composition of the water.

Smaller clusters within the structure suggest the same dimensions. For instance, within the "domestic" waters, there are three sub-clusters. The members of the bottommost have in common their interior location, faucet source, and domestic uses. The members of the middle have in common their composition and location. The members of the topmost seem to be linked by their use (for drinking) and their source (wells/springs in the ground).

Clusters indicated by the marked feature segments also seem to pick out similar features. For instance, tap water is linked to well, mineral, and distilled water by being used for drinking. Swamp, sewer and stagnant water are linked to the "wild" water cluster by their outdoor location. Pond water is linked to swamp water by similarity of location, source, composition, and function in nature, and puddle water is linked to the "dirty" waters by composition (that is, muddiness).

The scaling solution thus is consistent with the suggestions from the typicality ratings, and points to four basic dimensions involved in the concept of water: composition, current location, source, and function for humans. Only the first of these dimensions corresponds to the idea of an underlying essence as described by Putnam (1975) and as carried into many discussions of psychological essentialism. The other three dimensions are more superficial in the sense of being readily observable. Furthermore, function is a dimension that depends on human interaction with the liquid and so is not an inherent quality at all, but rather is endowed upon the liquid. Both location and function, in fact, are fairly readily changeable, so that a liquid that occurs in one location and is used for one purpose might at some point be transported to a different location and/or be used for a different purpose. Such changeable, context-dependent properties are not candidates for the essence of a category in Putnam's analysis and are not typically considered as such in discussions of psychological essentialism.

To the extent that such properties influence whether people consider a liquid to be water or not, this outcome would be inconsistent with the strong version of psychological essentialism. The scaling solution, of course, does not in itself demonstrate that these properties are used in determining what liquids people consider to be water. In fact, one might be tempted to suggest that these sorts of features are merely "identification features" (Smith & Medin, 1981; Miller & Johnson-Laird, 1976) that help pick out examples of water in a quick and dirty fashion; they are not the essential features that really make a liquid water. However, Experiment 1 showed that the proposed essential feature for water (a belief about H_2O) does not alone distinguish between water and non-water. Thus, there must be other features that not only help identify examples of water but are integral to making them be water. The properties identified here are candidates to be those other important features.

One piece of anecdotal evidence supports the involvement of the particular properties identified here in determining whether a liquid is considered to be water: A news item from the April 1991 issue of *Greenpeace Magazine*[3] reproduces a photograph of the Toronto skyline and states that *water* from polluted Lake Ontario served as the photographic developing agent. As long as the liquid involved comes directly from Lake Ontario, thereby emphasizing its lake source and its usual location in which it retains at least some of its original lake functions, people apparently continue to consider the liquid to be water. If the same impure liquid were bottled and sold in a photography shop for use in developing pictures, though, it would more likely be labeled and thought of as a weak developing solution. Calling it "water" in its new location and with its new intended function would seem both inappropriate and misleading. The General Discussion gives a more detailed illustration of the ability of these dimensions to explain what liquids are considered to be water versus non-water.

EXPERIMENT 4

The data to this point indicate that there is a general sense of "water" that refers to liquids containing mixtures of H_2O and other ingredients. However, the data do not rule out the possibility that there may also be

[3] Brought to my attention by Padraig O'Seaghdha.

a more restricted sense that conforms more closely to the strong essentialist view of the meaning of "water." It is likely that such a sense exists in scientific contexts. Chemists talking about water, for instance, may mean only pure H_2O, since other ingredients could affect the chemical processes under consideration. It is not clear, though, whether this more restricted sense exists as part of everyday English vocabulary. If both senses exist, then the strong version of the essentialist view should not be completely rejected. Instead, the conclusion would be that it is an appropriate description of one sense of the word but not another. The final study was designed to explore whether two senses of "water" exist in everyday English.

Part 1: "X Is Only Partly Water"

A sentence acceptability task was used to evaluate whether people have two senses of "water." Subjects judged the acceptability of statements of two forms, "X is a type of water" and "X is only partly water," where "X" represents water examples from the previous experiments. If people have only one sense of the word "water" in which it refers to a mixture of H_2O plus other ingredients, then they should find sentences such as "Lake water is a type of water" acceptable, since lake water is one possible mixture that would be an example of water. At the same time, they should find sentences such as "Lake water is only partly water" unacceptable, since a substance cannot both be an example of a mixture and at the same time have that mixture as one of its ingredients. In contrast, if people have only one sense of the word "water" in which it refers to H_2O alone, then they should find sentences such as "Lake water is only partly water" acceptable, since H_2O can be an ingredient in the mixture called "lake water." At the same time, they should find "Lake water is a type of water" unacceptable, since something that is a mixture cannot both have H_2O as one of its ingredients and be that ingredient. Finally, if people have two senses of "water," one in which it refers to a mixture and one in which it refers to pure H_2O, then they should find both types of sentences acceptable: "Lake water is a type of water" is acceptable on the first sense of the word, and "Lake water is only partly water" is acceptable on the second sense.

Method

Subjects. Forty-eight Lehigh undergraduates participated for course credit.

Stimuli. Since there is a limit on how many acceptability judgments subjects can meaningfully make, and since it was necessary to include a large number of filler sentences to disguise the comparison of interest, a subset of the 43 water examples from Experiment 1 was used. Eighteen examples were randomly chosen, with the constraint that 6 come from each third of the distribution of judged percentage of H_2O in that experiment. The 18 waters chosen were (in order of judged percentage of H_2O): pure, natural spring, ice, drinking, distilled, tap, stream, babbling brook, mineral, lake, salt, ground, swimming pool, stagnant, puddle, ocean, polluted, and swamp.

For comparison, 18 non-water examples were also used. These were chosen from among those that had judged percentages of H_2O within the range of the water examples. The non-water examples were randomly selected from this set with the constraint that 6 come from each third of the judged percentage of H_2O. The 18 non-waters were (in order of judged percentage of H_2O): tea, saliva, tears, sweat, lemonade, saline solution, cranberry juice, pickle juice, apple juice, clam juice, mouthwash, grapefruit juice, windshield wiper fluid, pineapple juice, Sprite, onion juice, contact lens cleaner, and radiator coolant.

Three stimulus sets were formed, each containing six water and six non-water examples, chosen so that two of each came from each portion of the judged H_2O distribution. Each example appeared twice in each stimulus set, once in the "X is a type of water" form and once in the "X is only partly water" form, yielding 24 target stimuli per set.

Mixed with the 24 target stimuli in each set were 96 filler sentences asking about substances other than water (e.g., milk, alcohol, cheese, paper). Half the fillers were of each sentence form. Within each form, half were designed to be answered "yes" (e.g., "Chocolate milk is a type of milk"; "Air is only partly oxygen") and half were designed to be answered "no" (e.g., "A hotdog is a type of dog"; "A towel is only partly cloth"). Within those divisions, half the sentences involved items that contained the substance being asked about in their name (as in the chocolate milk and hotdog examples), and half involved items that did not contain the substance being asked about in their name (as in the air and towel examples). Two random orders of the targets plus filler sentences were created for each stimulus set. The 120 sentences were presented in a five page booklet with a response blank next to each sentence.

Procedure. Subjects were told that the experiment concerned how names for things relate to what those things are. They were given the example that "tuna sandwich" clearly reveals what type of thing is being named (a type of sandwich), whereas "hoagie" does not (since a hoagie is also a type of sandwich). Subjects were also given the example that "helium" fully reflects the substance that is being named (since helium is an element), whereas "root beer float" does not (since there is more in a root beer float than root beer). They were told that they would be reading sentences stating something about a named object, such as "A hoagie is a type of sandwich" or "A root beer float is only partly root beer," and their task was to decide if each statement made sense. They were also instructed not to be concerned if some of their judgments seemed to contradict each other. They were told that some words have more than one meaning, and two superficially contradictory sentences might both make sense as a result.

Subjects responded by placing a check next to each sentence that they felt made sense and an "x" next to each one that they felt did not. They were run singly or in small groups and completed the task at their own pace. Sixteen subjects received each stimulus set.

Results and Discussion

Data were compiled as the percentage of target stimuli for the two sentence forms that each subject judged to be sensible. One water example was inadvertently omitted in the "type of water" form from one random order of stimuli, and a second was omitted in the "type of water" form from another random order. Percentages for the water examples in "type of water" judgments for subjects who received these two random orders were therefore based on five rather than six stimuli.

Non-water examples. Judgments for the non-water examples were clear-cut. Subjects consistently denied that it made sense to say that the non-water examples were types of water: They gave a mean of only 17% positive responses to sentences of the "type of water" form. At the same time, subjects consistently felt it made sense to say that the non-water examples were only partly water: They gave a mean of 84.3% positive responses to sentences of the "only partly water" form. The responses to the two forms of sentence differed significantly by a sign test, $z = 5.54$, $p < .01$. The positive responses to the "only partly water" sentences indicate that subjects do recognize that these liquids contain water (in some sense). The negative responses to the "type of water" sentences indicate that, consistent with the results of Experiment 1, recognition of some amount of water in a mixture is not enough for it to be considered a type of water.

Water examples. Judgments for the water examples contrasted sharply with the non-water judgments. Subjects consistently indicated that it made sense to say that the water examples were types of water, with a mean of 86.2% positive responses to sentences with the "type of water" form. However, they generally felt that it did not make sense to say that the water examples were only partly water: They gave a mean of only 35.2% positive responses to the "only partly water" forms. The responses to the two forms of sentence differed significantly by a sign test, $z = 6.43$, $p < .01$.

The acceptability of the "type of water" sentences is consistent with the outcome of the previous experiments and supports the idea that there is a sense of water corresponding to a mixture. Table 3 shows that the acceptability was consistent across examples, with all eighteen receiving positive responses from a majority of subjects. Furthermore, the degree of agreement (as reflected in the percentage of positive responses) was not related to either typicality of the example ($r = .23$, n.s.), or judged percentage of H_2O ($r = .26$, n.s.) from Experiments 1 and 2. The lack of correlation indicates that even mixtures recognized as having a relatively large amount of other ingredients and/or considered to be somewhat atypical are still sensible examples of water.

The low acceptability of sentences describing the water examples as only partly water, on the other hand, indicates that a sense of "water" corresponding to pure H_2O is not highly available in this context. Since subjects consistently judged the non-water examples to be acceptable in the same sentence context, they clearly do not view the sentence form itself as ungrammatical or otherwise anomalous. Instead, it seems that subjects simply did not think of the water examples as having water as one ingredient among others. In fact, one subject corrected two of the sentences on the page: (S)he circled the word "is" and drew an arrow from "is" to "water" so that the sentences would read "Puddle water is

TABLE 3
Mean Percentage of Positive Judgments for Water Examples in "Type of Water" and "Only Partly Water" Sentences (Part 1) and "Mostly But Not Entirely Water" (Part 2) Sentences

Water example	"Type of water"	"Only partly water"	"Mostly but not entirely water"
Pure	75	0	7
Natural spring	100	6	27
Ice	81	6	27
Drinking	69	25	47
Distilled	100	0	13
Tap	100	19	40
Stream	88	31	53
Babbling brook	88	25	47
Mineral	100	13	73
Lake	81	44	60
Salt	100	81	87
Ground	94	50	40
Swimming pool	63	50	67
Stagnant	75	25	40
Puddle	81	56	67
Ocean	88	75	67
Polluted	81	63	80
Swamp	69	63	80

Note. Water examples are in order of mean judged percentage of H_2O from Experiment 1.

water" and "Tap water is water." This subject's corrections, and the generally negative judgments, suggest that the liquids were thought of *as* water in this context, rather than as only partly water. In other words, it seems that the primary interpretation of "water" in these sentences was as a mixture that can include ingredients other than H_2O.

The pattern of responses across the eighteen individual water examples supports this conclusion. In contrast to the outcome for "type of water," the degree of acceptability for sentences describing water examples as only partly water was negatively correlated with the judged percentage of H_2O from Experiment 1 ($r = -.77$, $p < .01$). (There was also a non-significant relationship to typicality in the same direction, $r = -.32$, n.s.). Table 3 shows that the majority of the water examples were found not sensible (by the majority of subjects) in these sentences, but some of the examples with the lowest judged amounts of H_2O were relatively acceptable. This pattern indicates that there can be a substantial divergence between the use of the term "water" and what people believe about the H_2O content of the liquids. People recognize that many liquids such as drinking water and spring water are not pure H_2O, as shown by H_2O judgments in Experiment 1, but they do not necessarily think of these as being "only partly" water. In these sentences, only when the H_2O component is especially low, and the non-H_2O component is large, do they think of the liquid as having a water part and a non-water part. In fact, the five waters accepted by more than half the subjects are ones in which the non-H_2O ingredients seem to be especially salient and identifiable: Ocean and salt water are the most accepted, with polluted, puddle, and swamp water following. Ground and stagnant water, which may have less familiar ingredients, fall within the same range of judged percentage of H_2O, but they are accepted by less than half the subjects. This distribution suggests that people may think of a liquid as having a part that is not water only if the non-H_2O component is perceptually obvious.

The fact that subjects do accept some "only partly water" statements for water examples suggests that people may at times retrieve a sense of "water" as H_2O alone in interpreting these sentences. However, this sense appears to be retrieved only under limited circumstances. The more usual interpretation that subjects give to "water" in these sentences seems to be as a mixture that can include ingredients other than H_2O.

Part 2: "X Is Mostly But Not Entirely Water"

The results of Part 1 were interpreted as meaning that people often take "water" to refer to a mixture when they read sentences such as "X is only partly water." An alternative interpretation of the results might be that the subjects read "water" as referring to pure H_2O, but they rejected sentences of this form on the basis of pragmatics: The phrase "only partly" implies that the liquid in question has a relatively small amount of H_2O, which is incompatible with their belief. Two points argue against this interpretation. First, subjects in Part 1 did find the "X is only partly water" sentences acceptable for some water examples that had low judged amounts of H_2O, but at the same time, they rejected others that had equally low judged amounts. As discussed above, the acceptable ones seemed to be limited to those that had particularly conspicuous non-H_2O ingredients. Second, intuition suggests that comparable sentences that make the pure H_2O reading explicit (e.g., "Natural spring water is only partly H_2O"; "Lake water is only partly H_2O") are slightly odd but not so anomalous that they would be rejected as unacceptable sentences. To test the alternative interpretation more fully, however, subjects in Part 2 judged the acceptability of sentences of the form, "X is mostly but not entirely water." If subjects in Part 1 rejected sentences of the form "X is only partly water" only because the sentences implied a low amount of H_2O, they should find sentences of the new form acceptable.

Method

Subjects. Forty-five University of Illinois undergraduates participated for course credit.

Stimuli. All sentences of the form "X is only partly water" from Part 1 (including water, non-water, and filler sentences) were changed to "X is mostly but not entirely water."

Because some filler sentences previously accepted would have been rejected with this change, fillers were modified where necessary to maintain the balance of positive and negative responses. Stimulus booklets were otherwise identical to Part 1 booklets.

Procedure. The instructions were modified to accommodate the new wording of sentences. The procedure was otherwise identical to Part 1 procedure. Fifteen subjects received each stimulus set.

Results and Discussion

Non-water examples. Judgments for the non-water examples closely paralleled those from Part 1. Subjects gave a mean of only 16% positive responses to sentences of the "type of water" form, but they gave a mean of 80.7% positive responses to sentences of the "mostly but not entirely water" form. The responses to the two forms of sentence differed significantly by a sign test, $z = 6.10$, $p < .01$.

Water examples. As in Part 1, judgments for the water examples reversed the pattern for the non-water judgments. Subjects indicated that it made sense to say that the water examples were types of water, with a mean of 83% positive responses to sentences with the "type of water" form. However, they felt that it made less sense to say that the water examples were mostly but not entirely water: They gave a mean of only 51% positive responses to the "mostly but not entirely water" forms. The responses to the two forms of sentence differed significantly by a sign test, $z = 4.62$, $p < .01$.

The 51% overall level of acceptance of "mostly but not entirely water" for waters is somewhat higher than the 35.2% level for "only partly water" in Part 1. Although this comparison is across experiments, the fact that acceptance levels for non-waters for both sentence forms and for the waters for the "type of water" sentence form were virtually identical in the two experiments suggests that the difference is meaningful. By itself, this observation might be taken as support for the idea that results for the "only partly water" sentences of Part 1 were due to a mismatch of the sentence pragmatics with subjects' beliefs about H_2O content.

However, two other observations are inconsistent with this conclusion. First, despite the somewhat higher level of acceptance, the absolute level is still substantially (and significantly) lower than "type of water" judgments, with only about half of all judgments being positive. Second, the data for individual waters, given in Table 3, show that judgments for nine of the eighteen water examples were predominantly negative, and judgments for a tenth were only slightly over 50% positive. These relatively low levels of acceptance suggest that although the pragmatics factor may have had some influence in Part 1, it was not pragmatic fit alone that kept "only partly water" sentences in Part 1 from a high level of acceptability.

Notably, as in Part 1, the sentences that are most accepted tend to be those at the low end of the range of judged H_2O. This fact is additional evidence against the idea that the results of Part 1 were due to pragmatic incompatibility: The Part 2 sentence form does not imply a relatively low amount of water (in either sense), yet subjects still find most acceptable the sentences with waters with low judged H_2O. Given that there is no pragmatic reason to favor those sentences, this pattern also supports the idea that "water" is often given a mixture interpretation in these sentences, with an interpretation separating H_2O from other ingredients elicited mainly when the non-H_2O component is especially salient. Consistent with this idea, and as in Part 1, sentences referring to ground and stagnant water were relatively less accepted than sentences with other water examples in the same range of judged H_2O. For these two water examples, the particular non-H_2O ingredients may be less familiar and therefore less salient.

In sum, the results of Part 2 support the interpretation originally given for Part 1: That people often interpret "water" as referring to a mixture in these sentences, although they also at times retrieve a sense of "water" corresponding to pure H_2O. This fact is especially noteworthy given that the "X is mostly but not entirely water" sentence form should in itself tend to encourage retrieval of the more restricted sense of water. The results of this experiment therefore suggest that there may be a sense of "water" in ordinary English that corresponds to the strong essentialist view, but it is not the only, and may be the less frequent, sense of the word. Theories of concept structure and of word meaning clearly need to account for all standard senses of the word, not just one.

GENERAL DISCUSSION

The four experiments just presented suggest several conclusions about the meaning of the word "water" and about the nature of concepts and word meanings in general.

The Limited Role of Belief in An Essence

The first conclusion from this study is that belief in a particular essence, H_2O, does not by itself account for what liquids people consider to be water (in one common sense) and what liquids they do not; judgments of the amount of H_2O in a liquid do not predict well whether that liquid is considered to be water. To the extent that Putnam's (1975) account of word meaning assumes that people's use of words is based on belief in an essence, his account is inconsistent with the data. Of greater direct significance to psychological theories, this conclusion argues against a strong version of psychological essentialism that takes essentialist beliefs as the only factor constraining what things will be granted category membership. Results from the study of one concept do not, of course, show that essentialist beliefs will never be the sole critical factor for any concept. However, they do place limits on the generality of the approach, and they indicate that a complete theory of concepts or of word meanings will need to look beyond beliefs in an essence alone.

A possible response to this conclusion is that people's concept of water is, in fact, simply H_2O, and the apparent counterexamples presented here are uses produced only due to communication needs. That is, perhaps lack of a better word in English vocabulary leads to referring to a variety of liquids as "water" even though they do not fit people's concept of "real" water. However, several observations argue against this interpretation. First, if this interpretation were correct, subjects in Experiment 4 should have been much less willing to accept sentences such as "Lake water is a type of water" and much more willing to accept ones such as "Lake water is only partly water," since the latter would reflect their "real" concept better than the former. Second, as noted earlier, a number of waters judged to have a relatively low percentage of H_2O are nevertheless considered to be relatively good examples of water as reflected in typicality judgments. If ocean water, for example, were called "water" only due to lack of a better name for the liquid, it should be considered an atypical rather than a fairly typical example of water. Third, this alternative interpretation cannot account for why some liquids judged high in H_2O are *not* called "water." Liquids such as tea and coffee that are judged as high in H_2O as typical waters such as tap and rain water should, under this account, be called "water." Finally, the examples of tea, coffee, and other non-waters with high H_2O content illustrate that English has, in fact, developed contrastive names for many liquids that share their major ingredient. Ocean water, river water, and other naturally occurring bodies of water have been central to English-speaking people's lives for centuries, and it is implausible that no better vocabulary for them would have developed if the label "water" were only an ill-fitting extension of a more conceptually appropriate use. Together, these observations suggest that there must be some principle(s) that determine when distinctive labels will exist (as for "tea" and "coffee") and when they will not (as for "ocean water" and "river water").

Four Dimensions for Classification

A second conclusion from this study is that at least four dimensions appear to be involved in the concept of water. These dimensions include source, current location, and function, in addition to the composition of the liquid. As noted earlier, the failure of H_2O to fully account for what liquids are considered to be water implies that these dimensions are not just part of a quick and dirty identification procedure for recognizing examples of water, but rather have a role in actually determining category membership.

If this proposal is correct, it should be possible to explain why the non-waters are not considered to be water in terms of the dimensions just described. In Experiment 1, 25 of the non-water examples were judged to have percentages of H_2O within the range of the waters (67% or greater); these are therefore the best test cases for the explanatory adequacy of the other dimensions. Strikingly, these 25 fall into several distinct groups that differ prominently from the examples of water on one or more of the dimensions identified. The juices (apple, cranberry, grapefruit, pineapple, claim, onion, pickle) are squeezed from plant or animal tissues (a contrasting source) and used only for consumption (a more restricted function); the body fluids (blood, sweat, urine, tears, saliva) are produced within the body (a contrasting source) and serve special cell and organ functions (a more restricted function). The general beverages (Sprite, gingerale, lemonade, coffee, tea, chicken broth), cleaners (Windex, contact lens cleaner), personal care products (mouthwash, eyedrops, saline solution for contact lenses), and car products (radiator coolant, windshield wiper fluid) are created and used exclusively as beverages or as cleaners, for personal hygiene, or for auto care, respectively (more restricted functions), and the non-H_2O ingredients are specifically added to obtain those functions (contrasting sources). The contrasts on these dimensions appear to serve as the basis for concepts that are separate from the concept of water. The concept *juice*, for instance, centers on the idea of drinkable liquid extracted from tissue, and the concept *cleaner* centers on the idea of a substance made for a specific cleaning purpose. The name a liquid is given, then, may be driven primarily by the concept that the liquid fits into. A liquid that is used as a beverage will be named to contrast with other beverages, whereas one that is created as a cleaner will be named to contrast with other cleaners, and so on. This account explains why many liquids judged high in H_2O are not called "water": Even though people may be aware that the liquids are largely H_2O, as long as they fall into conceptual domains that are separate from *water*, they will not be named as types of water.

The dimensions of source, location, and function, then, help provide a more complete explanation of what liquids are considered to be water than composition of the liquids alone can. The apparent importance of these multiple dimensions suggests that it may be necessary to revive the possibility of a family resemblance structure in describing what liquids are considered to be water (see also Lakoff, 1987). It should be noted, though, that H_2O was judged to be present in all the water examples collected. It may be that one property, the presence of H_2O, is necessary, but not sufficient, for a liquid to be considered water. If H_2O is believed to be present, the remaining factors may then determine whether the liquid will be considered water or a member of some other category.

These observations suggest a version of psychological essentialism that is somewhat weaker than that implied in many existing discussions. The modified version includes essentialist beliefs as a factor that influences category membership judgments and word use and may provide an important constraint on them. However, it also includes other properties of objects, including ones that are perceptually obvious, context-dependent, and changeable, as potentially important in helping to determine membership and word use. This modified version incorporates an important insight from Putnam's original analysis, but is much less true to the spirit of his approach. In fact, to the extent that it takes dimensions such as use by humans as influencing category membership (even for natural kinds), this version has some resemblance to Lakoff's (1987) diametrically opposed approach emphasizing the role of human experience in classification.

Natural Kinds and Artifacts

A third suggestion from the experiments presented here is that natural kind concepts may be more similar in nature to artifact concepts than has often been assumed. Within philosophy, the essence possibility has been taken most seriously for natural kind terms. Likewise, within psychology, the existing empirical evidence about essences has focussed on natural kinds, and discussion of essences has also tended to emphasize natural kinds (e.g., Carey, 1985; Markman, 1989; Medin & Ortony, 1989). Artifact terms, and their associated concepts, have more often been taken to be somewhat more loosely defined (e.g., Gelman, 1988; Keil, 1989; Malt & Johnson, 1992). However, the data just presented suggest that, like artifacts, natural kind terms such as "water" may not be easily characterized in terms of a single factor that by itself provides clear category boundaries.

It might be argued that belief in an essence is more central to living kinds such as plants and animals than it is to elements such as gold or compounds such as water (Atran, 1990). Under this argument, the strong version of psychological essentialism may be wrong for water, but it nevertheless may be relevant for many other natural kinds. However, several examples suggest that there will be limits on the role of essentialist beliefs even for living kinds. On the one hand, there are cases such as of caterpillars versus butterflies and tadpoles versus frogs, where people recognize shared hidden properties such as genetic endowment but maintain separate categories for the superficially different forms. On the other hand, there are cases such as of trees, fish, and vegetables (see Keil, 1989; Malt, 1991), where people may come to recognize that their beliefs about shared hidden properties among the category members are wrong, but they do not abandon the categories. For instance, realization that there is no biological basis for the grouping of things we call "trees" does not typically cause people to declare the category nonexistent. Instead, they adjust their beliefs about the nature of the category. These examples suggest that factors including morphological differences (as for caterpillars versus butterflies), morphological similarities (as for trees), habitat (as for fish), and use by humans (as for vegetables) all have the potential to influence classification of living things in addition to beliefs about essences. The relative importance of these factors versus essentialist beliefs remains to be explored. However, it is likely that some degree of continuity, rather than a strict dichotomy, between living and non-living kinds will need to be recognized.

Beliefs vs Word Use and Categorization

Finally, the notion of essence-based definitions obviously has great intuitive appeal both to theoreticians and in folk psychology; versions of essence theories have existed for many years (see, e.g., Hull, 1965; Keil, 1989; Lakoff, 1987 for reviews). However, the data suggest that there is an important gap between conscious beliefs about the meaning of words such as "water," which may be based entirely on the idea of an essence, and the knowledge on which categorization and word use is actually based. If people do in fact hold a strong belief that natural kind terms are fully defined by essences, then understanding the reason for this belief is itself of interest. However, to explain actual category membership judgments and word use, it may be necessary to separate the question of conscious belief from the question of the knowledge involved in the categorization and word use processes.

APPENDIX

Mean Pairwise Similarity Ratings for 20 Examples of Water (1 = very dissimilar; 9 = very similar)

1 = bath water	11 = rain water
2 = chlorinated water	12 = river water
3 = dish water	13 = salt water
4 = distilled water	14 = sewer water
5 = flood water	15 = stagnant water
6 = lake water	16 = stream water
7 = mineral water	17 = swamp water
8 = pond water	18 = swimming pool water
9 = puddle water	19 = tap water
10 = auto radiator water	20 = well water

1																				
2	4.20																			
3	6.05	4.50																		
4	4.55	4.75	4.70																	
5	3.25	2.95	3.60	3.15																
6	3.95	2.95	3.45	3.30	5.20															
7	4.45	3.65	3.40	6.90	3.10	4.45														
8	2.75	2.60	3.05	2.75	5.00	6.75	3.40													
9	3.25	2.50	3.35	2.65	5.50	5.15	3.00	5.60												
10	3.15	2.90	3.95	3.65	2.95	2.60	3.20	3.15	3.05											
11	4.05	3.35	3.50	3.90	7.10	6.80	4.15	6.25	7.20	3.85										
12	3.35	2.50	3.70	3.55	6.70	6.25	4.50	5.65	5.05	3.05	7.15									
13	2.90	3.00	2.40	2.80	2.45	3.85	3.65	3.05	2.70	2.55	3.25	3.20								
14	2.60	2.55	3.05	1.85	4.70	2.65	1.85	3.35	4.65	4.30	4.40	3.75	2.45							
15	4.05	3.05	3.85	3.10	3.35	4.70	2.85	5.20	6.20	4.55	2.75	3.40	3.10	4.65						
16	4.30	2.95	3.25	4.20	6.15	6.35	5.55	5.45	4.75	2.45	6.45	6.95	3.05	2.65	3.15					
17	2.15	2.30	3.05	1.80	4.35	4.40	2.55	5.70	5.35	3.40	4.60	4.40	2.90	5.40	5.70	3.70				
18	5.50	8.15	4.00	4.80	3.30	4.35	4.40	3.10	2.95	3.25	4.45	3.65	3.35	2.00	3.80	4.15	2.20			
19	6.90	4.65	6.75	6.55	3.35	3.60	5.80	3.55	2.60	3.85	4.30	4.10	2.90	2.80	2.60	4.70	1.95	5.60		
20	5.35	4.25	4.50	5.55	4.25	4.75	6.95	4.80	4.20	2.95	6.75	4.95	2.95	2.15	4.10	6.00	2.65	4.00	6.75	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

REFERENCES

Atran, S. (1990). *Cognitive foundations of natural history*. Cambridge, England: Cambridge University Press.

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Corter, J. E., & Tversky, A. (1986). Extended similarity trees. *Psychometrika*, 51, 429–451.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, **87**, 215–251.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, **20**, 65–95.

Hull, D. L. (1965). The effect of essentialism on taxonomy—Two thousand years of stasis. *British Journal for the Philosophy of Science*, **15**, 314–326.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Kucera, N., & Francis, W. N. (1967). *A computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Malt, B. C. (1990). Features and beliefs in the mental representation of categories. *Journal of Memory and Language*, **29**, 289–315.

Malt, B. C. (1991). Word meaning and word use. In P. Schwanenflugel (Ed.), *The Psychology of word meanings* (pp. 37–70). Hillsdale, NJ: Erlbaum.

Malt, B. C., & Johnson, E. C. (1992). Do artifact concepts have cores? *Journal of Memory and Language*, **31**, 195–217.

Malt, B. C., & Smith, E. E. (1984). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior*, **23**, 250–269.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A.

Ortony (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Miller, G. A., & Johnson-Laird, P. (1976). *Language and perception*. Cambridge, MA: Cambridge University Press.

Neisser, U. (1987). From direct perception to conceptual structure. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization*. New York: Cambridge University Press.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, **84**, 231–259.

Putnam, H. (1975). The meaning of 'meaning.' In H. Putnam (Ed.), *Mind, language, and reality: Philosophical papers, vol. 2*. Cambridge, England: Cambridge University Press.

Putnam, H. (1988). *Representation and reality*. Cambridge, MA: MIT Press.

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

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

Wellman, H. M., & Gelman, S. A. (1988). Children's understanding of the nonobvious. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 99–135). Hillsdale, NJ: Erlbaum.

(Accepted June 30, 1993)