2008

The real deal : is psychological essentialism for artifacts 'really' real?

Micheal R. Paquet
Lehigh University

Paquet, Michael R.

The Real Deal: Is Psychological Essentialism for Artifacts "Really" Real?

May 2008

The Real Deal:

Is Psychological Essentialism for Artifacts "Really" Real?

by

Michael R. Paquet

A Thesis

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Psychology

Lehigh University

May 2008

Michael R. Paquet

This thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science Degree.

4/15/2008
Date

Barbara Malt, Ph.D.
Professor of Psychology
Adviser

Catherine Arrington, Ph.D.
Assistant Professor of Psychology
Co-Adviser

Heidi Grant-Halvorson, Ph.D.
Assistant Professor of Psychology
Co-Adviser

Diane Hyland, Ph.D.
Professor of Psychology
Department Chair

Table of Contents

Abstract

Although people seem to possess essentialist beliefs about many natural kinds, it is not as clear whether they do so for artifacts. The experiments presented here tested hypotheses about artifact cognition generated by three competing theories: strong essentialism, weak essentialism, and radial theory. In Experiment 1, participants rated the acceptability of various ways of talking and thinking about photographed artifacts that had been created for one purpose but were being used for a different purpose. When thinking about how to talk about the artifacts, people preferred the new name. This trend was reversed when asked what the artifacts *really* were, consistent with strong essentialism. However, participants rated the new category higher as what the object *really* is when the artifacts were typical with respect to that category than when they were atypical, a finding not consistent with either essentialist perspective. Participants consistently rated both categories as generally acceptable, also not consistent with essentialism. In Experiment 2, participants rated ways of talking or thinking about artifacts created for one purpose but used for another purpose by one or many people of varying entitlement; these artifacts were highly typical of both categories. Participants tended to rate the objects as *really* both categories, as well as liking both names for talking about them. Consistent with radial theory's predictions, people were willing to allow artifacts to belong *really* to multiple categories for non-essentialist reasons. The results support radial theory; there is no universal psychological essentialism for artifacts.

The Real Deal: Is Psychological Essentialism for Artifacts "Really" Real?

Imagine a table. You might be thinking of a solid object with four legs and a flat surface, perhaps made of wood, roughly three feet in height. Now imagine that one leg has been sawed off. Is it still a table? Imagine that it is made of Jell-o, not wood. Still a table? Imagine that all the legs have been removed, and it has been shrunk to the size of a nickel. Still a table? If it is not really a table, then what is it?

Many philosophers and psychologists have argued that people believe that categories in the world have an essence – some property or set of properties without which an object in the category would cease to be a member of the category. This psychological essentialism is proposed to exist regardless of whether such essences actually exist (metaphysical essentialism) (Medin & Ortony, 1989). Is psychological essentialism really real? In this paper, I argue that psychological essentialism is not real for artifacts. I review evidence for and against three competing theories: strong essentialism, weak essentialism, and radial theory. In two experiments below, I present circumstances that will tease apart these theories and their predictions. I conclude by proposing that radial theory can best account for the trends in these experiments.

Essentialist and Non-Essentialist Evidence for Natural Kinds

There is evidence from the psychological literature suggesting that people believe natural kinds (i.e., naturally occurring things, such as rabbits and rocks) have an essence (Keil, 1989) and that they may look to experts for information

about what that essence is (Putnam, 1977; Gelman & Coley, 1991). Until recently, astronomers officially considered Pluto a planet, so lay people should have considered Pluto a planet, too. Now that Pluto has been officially relegated to "minor planet status" by the International Astronomical Union (Shiga, 2006), non-experts should change their beliefs about Pluto, too, if Putnam's argument about deference is correct. However, to the extent that people do have essentialist beliefs about natural kinds, they do not apply them consistently or to all natural kind categories (Malt, 1994; Kalish, 1995; Kalish, 2002). For instance, Malt (1994) had participants judge the amount of H_2O in liquids typically called "water" (e.g. rain water) and liquids typically not called "water" (e.g. Sprite). Although an essentialist model would predict that people would use the presence or absence of H_2O to determine what is or is not water, Malt's participants did not consistently judge as "water" those liquids that they also judged to contain the most H_2O (e.g. tea). Likewise, they did not consistently reject as "water" those liquids they judged to contain relatively less H_2O (e.g. sewer water).

In addition to making category judgments inconsistent with essentialist predictions, people also do not simply defer to expert biological opinion about natural kind categories when that opinion is at odds with common sense. Braisby (2001) asked participants whether a natural kind that had been altered genetically with or without physical changes was still a member of its original category. He then gave putative "expert opinion" from a biologist about the item's categorical status and asked the participants the same question again. Some participants'

subsequent responses were affected by expert opinion, but not when experts suggested that a genetically altered natural kind that no longer looked, tasted, or smelled like its original category was still a member of its original category. Thus, people seem to care about expert opinion to some extent but are capable of thinking the expert opinion is wrong.

In sum, people may hold some essentialist beliefs about natural kinds, but these beliefs are not generic to all cognition about natural kinds. Even when a natural kind such as water appears to have a scientifically well-defined essence, people do not stringently apply the definition when deciding what is and what is not water (Malt, 1994). Furthermore, people may not even accept the opinion of an expert such as a scientist with respect to a natural kind's essence if a more plausible opinion is accessible to them (Braisby, 2001).

Essentialist and Non-Essentialist Evidence for Artifacts

There has also been a debate in the literature as to whether artifacts (i.e., human-made objects) have a psychological essence. Kripke (1980) proposed that an artifact's essence is contained in the names we learn to give artifacts when we are young; those names could theoretically be traced back to some primitive societal decision to call particular objects by particular names. Some psychologists (e.g. Keleman, 1999; Matan & Carey, 2001) have suggested that the psychological essence for artifacts is the original intended function. If someone created an object with the intention for it to be sat upon, then its essence is the function "to be sat upon," which may coincide with a typical concept of *chair*. Bloom (1996) revised

Dennett's (1987) "design stance" to suggest that the original intended *category*, not the function, is in fact the essence. Intended function has relevance, but intended category encompasses other factors, such as physical properties and a name associated with a particular category. In his intentional-historical hypothesis, Bloom (1996) posited that a person assumes that objects belonging to a given category are "those entities created with the intention to belong to the same [category] as the original object or objects she was exposed to" (p. 11). Supporting this hypothesis, children consider the category membership intended by the artist when choosing names for drawings (Bloom & Markson, 1998) and by the creator when choosing names for novel artifacts (Diesendruck, Markson, & Bloom, 2003).

However, there is some evidence that linguistic categorization can be affected by factors other than original intended category, even when that intention has been made salient. Kemler Nelson, Herron, and Morris (2002) found that children and adults were affected by both original intended category and the nature of object damage in name-extension tasks. Participants were more willing to include broken objects in the objects' original category than objects that appeared intentionally dysfunctional. The authors concluded that intended function – though not necessarily the original one – provided the essence for categorization.

Dennett (1990) suggested that when an artifact takes on a new function, its original intended function may merely serve as historical interest; the users, not the inventor, decide what the object is. This implies that the "essence" might be

taken over by the new function. Although it does not seem central to his original intentional-historical hypothesis, Bloom (1996) does allow for artifact essence to change in a situation in which an agent (either the original creator or someone else) decides that the artifact belongs to a new category (e.g. a penny being used as a pawn in a chess set). It follows intuitively that an act that consciously alters the purpose of an artifact substantially might justify membership in an alternate category. For example, consider an object created to be a stool with a round seat. If someone were to begin using the object as an end table, it is conceivable that people might consider the object to be an end table and not a stool, particularly as the duration of the new use increased. In this weaker form of essentialism, artifacts have essences that are psychologically real, but subject to change. In order for this still to be considered essentialism, however, it must be the case that an object has one and only one essence at a given time (i.e., following a categorical decision by the creator). If psychological essentialism is real, and it permits essence to change, it remains to be seen how many changes are permissible and who is eligible to make such a change. But once an essence has changed or become ambiguous enough as to lead different people to believe that it has different essences, one must ask whether there is really any essence at all. The idea of an "essence" spanning multiple categories and being potentially changeable by multiple people contradicts the notion of an essence as the *one* true nature of an object.

Recent evidence, however, suggests that neither intended category nor intended function consistently behave as essences in judgments of category membership for artifacts. On the contrary, people are affected by information about how many people are using an object in a particular way when deciding what the object is "really for." Siegel and Callanan (2007) presented adults and children with pictures of novel artifacts. Participants learned that each object was created for a particular function but was being used for a different function by either one person or many people. Participants were asked to decide what each object was "really for." Adults and children were significantly more likely to endorse the original function when just one person had altered the function; when many people had altered the function, participants preferred the new function as the "real" one. Although this study did not record judgments of category membership, the evidence for functional flexibility suggests that neither original intended function nor new intended function can be considered the essence of an artifact. If there is any psychological essence demonstrated in this study, it is flexible and not bound by theories of intentional origin.

Even if an essence is subject to change, a weak essentialist theory should posit that an object can only have one essence at a time – otherwise, there would be no such thing as an essence (Markman, 1989; Rothbart & Taylor, 1992). If people believe in multiple "essences" for the same artifact, then whatever these things are, they are not essences. Determining whether people believe that an artifact has one or more essences can be difficult. Malt and Sloman (2007) had participants

read descriptions of artifacts that had been accidentally or intentionally changed so as to reduce their association with one category and increase their association with another category. For example, one description depicted a piece of plastic being shaped into a knife, either accidentally or intentionally. Participants rated the appropriateness of statements about the artifact – whether it was really a piece of plastic, really a knife, really both, sort of a piece of plastic, sort of a knife, or sort of both. Participants were, in some circumstances, willing to allow artifacts to belong "really" to both categories.

In a different experiment from the same series, participants read about an artifact that was created to belong to one category (e.g. teapot) but was being used to fulfill a function associated with another category (e.g. watering can). When asked to choose ways of referring to the object, participants were more likely to prefer the new name to the name associated with the original category as the new use approached permanence. If, for instance, the fictional character did not just use the teapot as a watering can one time only, but continued to use it in the new use, preference for the new name increased. These results suggest that neither a traditional essentialist theory nor a weak essentialist theory can adequately account for the linguistic description of artifacts in discourse contexts.

Possible Implications of Previous Research

Option A: Essentialism for artifacts is real!

It is possible that psychological essentialism for artifacts is a real phenomenon, but the aforementioned studies that suggest otherwise have been asking the wrong

questions. For instance, Siegel and Callanan's (2007) participants might have been thinking too much about function and appropriate reference, and not enough about what the experimental items *really* were, to reveal their essentialist beliefs. That is, artifact judgments in the literature that have been used to argue against essentialism (Malt & Sloman, 2007) may have been tapping preferences for *names* people will call artifacts, or functions that *could be fulfilled* by artifacts, but perhaps there is an underlying psychological essence about what artifacts *really* are that cannot be swayed by other, pragmatic factors (Bloom, 2007; see also Malt & Sloman, 2007). There might be a difference between what people might call something and what they might judge that it *really* is.

Language and discourse constraints allow the objects referred to by a particular word to be somewhat flexible (Malt, 1989). When asked to imagine a table, most people would probably picture a table with four legs. Yet we would still consider a three-legged table part of the larger category referred to by the word "table." Are three-legged tables *really* tables, though? We can imagine cases in which we might refer to an object with a particular name, but when thinking about it more deeply, we might suppose that the object is not *really* a member of the given category.

For example, imagine a large disk suspended from the ceiling by cables, with a fold-down seat attached, and manila folders tacked around the perimeter. Although this object varies considerably from a typical desk, might people be willing to call it a desk? Malt and Johnson (1992) found that some participants

were willing to call the described object a desk, as measured by their response to the question "Is this thing a desk?". The authors reasoned that such participants might have viewed the object as a futuristic desk. However, asking participants whether the object was "really" a desk might have prompted different responses. They might have been inclined to say it was not "really" a desk if they had envisioned some other use for the object when reading the description, even though this scenario involved no consideration of intentional origin.

Option B: Essentialism for artifacts is not real!

It is also possible that the studies that have argued against essentialism for artifacts have tapped into a psychologically real phenomenon: People do not consistently behave in an essentialist way when thinking about artifacts. If that is the case, there must be alternative explanations for what governs people's understanding of artifacts. One plausible alternative is radial theory, which is an offshoot of prototype theories that posit that how much an object is considered a member of a category depends on how much its features match those of a typical member of that category. In Rosch's (1975) description of prototypes, artifacts have graded category membership; category members are considered better exemplars of the category as their similarity to the category prototype increases. For instance, Rosch's participants rated *chair* as a better example of the category *furniture* than *telephone* or *refrigerator*, ostensibly because what comes to mind when people think of chairs is closer to their prototype for the broader category *furniture*.

In contrast to essentialist theories, prototype theories of naming posit that people have some central ideas about what makes an artifact belong to a particular linguistic category, but these ideas are flexible and tolerant of instances in which an artifact might belong to more than one category simultaneously (Lakoff, 1987; Rosch, 1975). In Lakoff's radial theory, there is no single factor that can predict how a name is extended. On the contrary, names can be extended for a host of reasons – contextual, historical, cultural, pragmatic, and so on. Through a combination of centrality and chaining, names are extended in time and space; the larger culture may or may not incorporate a particular extension into its collective semantic bank. An object is identified as a chair because it is featurally close to the central features of typical chairs. A new object may be identified as a *chair* because of its "chained" similarity to the first chair, even though it lacks some of central features of chairs that the first chair possessed. In this way, objects called *chairs* may, over time, look less and less like typical chairs. They are called "chairs" because of their chained similarity to other chairs that had been closer featurally to typical chairs. Extending a name to an object can also fail. Because name extensions are flexible and potentially overlapping, it is possible that (1) two different category names could be considered acceptable for the same object, or (2) a once-acceptable name use could become less acceptable in light of contextual changes. Therefore, there can be no a priori reason to subscribe to an essentialism of any sort. In this paper, I propose that radial theory is most suited to an accurate explanation of naming phenomena. To understand what is really

going on with name use, we must look beyond one-word categorizations and describe other features of the object that would explain its position in categorical space.

An essentialist theory would predict that asking people to think about how they would refer to an artifact could elicit responses different from what people *really* think about the same artifact (Bloom, 2007; cf. Malt & Sloman, 2007). In other words, what something is *called* is not necessarily the same as what it *really* is. Radial theory would predict no such distinction between what something is called and what it *really* is. On the contrary, this view suggests that judgments of what something *really* is could not reliably be ascribed a stability as a function of any single feature of the artifact. That is not to say that there is no stability for an artifact, in terms of what we are willing either to call it or say it *really* is. However, because many features of an artifact interact to determine what we are willing to call it and what we think it *really* is, changes to the features, including simply how the artifact is being used at a given moment, or how typical the artifact is with respect to some category, can also interact to effect a change in what we are willing to call it and what we think it *really* is. No matter whether people are *talking* about an artifact or thinking about what it *really* is, they should be trying to capture the most accurate description of the artifact based on what they currently know about it. According to radial theory, both of these activities would be subject to the same sorts of influences. It follows, then, that accurate linguistic descriptions, including categorization, are necessarily dynamic.

Judgments of what an artifact *really* is can change, and an object can belong *really* to more than one category simultaneously. It is still possible, and likely, for people to judge that something was *created to be* an X, but is now *really* a Y. Such knowledge of intentional history, in contrast to an essentialist account, would not necessarily be the overarching factor in determining what something *really* is. Intentional history, in the radial theory, is one of many features that can affect *really* judgments. If the artifact had been intended to be an X, but is highly atypical of that category, then people should be less likely to say that it is *really* an X.

Bloom (1998) admitted that judgments of initial intended category are often assumptions and therefore fallible. This begs the question of why an assumption about a creator would even be necessary in situations in which the act of creation has not been made salient. Why, for instance, couldn't people just assume that an artifact is a chair because it looks so much like other chairs, or has the capacity to function like other chairs, or was just called a chair by someone else (Malt & Johnson, 1998)? If people truly give precedence to a psychological essence when thinking about artifacts, their judgments of what an artifact *really* is should bear this out.

Another prediction generated by radial theory is that two members of the same category could conceivably be called by the same name for different reasons. Consider, for example, a miniature model boat encased in a glass bottle. Its shape allows it to be called a boat, but its size makes the typical function of a boat (to

carry one or more people over a body of water) impossible. On the other hand, we can imagine an object that functions as a boat despite not having the typical physical characteristics of a boat (e.g. a hexagonal-shaped raft with a propeller, six-foot-high sides, and a sail, that can safely accommodate two people). The features of the object and what is known – and assumed – by the person who needs to refer to it will interact to affect to what extent a particular name is deemed appropriate for reference to an artifact.

The radial view's predictions have been supported in several experiments. Siegel and Callanan's (2007) results implicate the number of users of an artifact as an important factor in people's judgments of the artifact's function. In the absence of other information (e.g. typicality, familiarity, pragmatics), if people know what the artifact was created for and how many people are using it in some new capacity, they are more likely to say that the object is *really* for the new use if many people, not just one person, are using it in that capacity.

In addition to the numerosity factor, people's judgments about category membership are also affected by the expertise of an opinion, although not consistently. As noted earlier, Braisby (2001) presented participants with information about a natural kind (e.g. a potato) that had had 50 percent of its genetic material replaced with material from another natural kind that was either a member of the same superordinate category (e.g. another vegetable) or was not (e.g. a fruit). The item was said to continue to look, taste, and smell like its original form. Participants then judged the item's membership in the original

category. Later, when presented with membership judgments by "most biologists," two-thirds of participants switched their categorizations to agree with the experts – but only in the conditions in which the genetic modification came from the same superordinate category as the original item. When the expert opinion suggested that the item was still a member of the original category even though it now contained the genetic material of another superordinate category, many participants did not switch their category judgments. This obduracy was even more pronounced in another experiment in the same series in which the appearance of the original item was said to have changed along with the genetic material. Therefore, expert opinion seems to matter to a certain extent, but when expert opinion flies in the face of common sense, people will reject the supposed expertise. Furthermore, Rothbart and Taylor (1992) suggested that if people look to experts to determine category membership in the first place, then adjustment of categorization on the advice of an expert would also be possible.

A third factor that has been shown to be relevant to category membership judgments is communicative goals. As discussed earlier, Malt and Sloman (2007) presented participants with stories about artifacts that were created for one purpose (tea kettle) but were being used for a different purpose (watering can). Participants rated possible ways for referring to the objects and also rated what the objects "really" were. Participants consistently preferred the name associated with the new use when thinking about how to talk about the object. Furthermore, ratings that the object was *really* both the new and old categories were higher in

scenarios in which it would have made sense to refer to the object with either the old or new name. For instance, when rating ways of talking about the watering can/teapot, participants thought both names would be acceptable when addressing someone unfamiliar with the story. When asked to think about what this object *really* was, participants were sometimes willing to say it was *really* both a teapot and a watering can. Participants still tended to prefer the original category in their *really* judgments, but this preference was mediated by referential pragmatics. As the referential utility of the new category name increased, ratings that the object was "really" the new category also increased, as did ratings that the object was "really" both categories. Ratings that the object was "really" the old category decreased, however. These findings contradict essentialist predictions.

People are also affected by information regarding whether one member of a category or a whole category is at odds with factors important to membership in that category. Braisby, Franks, and Hampton (1996) presented participants with counterfactual scenarios involving natural kinds in which either an individual category member (one cat) or the entire category (all cats) was found to lack either an essential property (the quality of being a mammal) or nonessential property (ability to meow) of the category. Participants then responded to true/false statements about the individual and category. Contrary to essentialist predictions, participants did not consistently endorse statements suggesting that an individual found to lack an essential category property was no longer a member of the category. For essential changes to an entire category, participants tended to

disagree with essentialist claims more often when the statements contained information beyond mere existential statements. For example, when given the information that all cats are not mammals but instead robots controlled from Mars, participants responded to the simple existential statement "Cats do exist" with high "true" ratings. But in the same scenario, participants ascribed some truth value to the statement "There are no such things as cats, only robots controlled from Mars." Thus, participants were willing to endorse seemingly contradictory statements in order to express their opinions as accurately as possible. The authors concluded that there are different senses to natural kind terms, and different experimental designs can elicit different types of responses about category membership.

In summary, there are potentially many factors that influence how people think about and choose names for artifacts. Although essentialist theories heavily weight intentional history to the exclusion of other factors, other research has demonstrated that people are also affected by the number of people using an artifact in a certain way (Siegel & Callanan, 2007) and pragmatic goals (Malt & Sloman, 2007) when thinking about artifacts and how to talk about them. In addition, there is evidence that people consider opinion expertise (Braisby, 2001) and the extent to which one member of a category or the whole category is at odds with factors important to membership in that category (Braisby, Franks, & Hampton, 1996) when categorizing natural kinds. In the experiments discussed

below, I explore the effects of these factors as well as intentional history on people's name judgments for artifacts.

Overview of Current Research

Really Judgments as Linguistic Description

Picture a flat, rectangular object made of foldable cloth, with a thin plastic veneer on one side. It is 7 feet long and 5 feet wide, with a red-and-white checkered pattern on one side. It was created to be a tablecloth but is being used as a picnic blanket from now on. People might be willing to call it a tablecloth and a picnic blanket. But do people think it is *really* a tablecloth and a picnic blanket? Or really just one or the other? Or really neither?

Judgments of what an artifact *really* is are potentially interesting because of what they may reveal about how people think about artifacts. One possibility is that by moving beyond what will make for the most successful reference in a conversation, and getting people to think about what something *really* is, a dissociation between discourse-based reference and linguistic description may be revealed. By "linguistic description," I mean whatever sort of verbalization deemed most accurate and succinct by a speaker for describing a particular artifact. This differs from more traditional linguistic categorization in that there is no reason to assume categorization must take place in order for an artifact to be described in a way that effectively conveys its properties. For instance, the strange desk described in Malt and Johnson (1992) might be called *desk* by some people, but its atypicality cannot be appreciated with such a parsimonious name. A

narrative elaboration, as given in the artifact's description, would be a more complete way of describing the artifact. In this way, linguistic description does not preclude categorization with a single noun; but it allows for more accurate description if needed. Linguistic categorization, then, is a form of linguistic description.

Because much of the foregoing research on the subject has dealt with categorization per se, I will address linguistic categorization, with the understanding that it need not be the only, or most accurate, form of linguistic description for an object. Although a *really* judgment could theoretically reflect non-linguistic categorization, I am using *really* judgments to examine whether there is any systematic difference between what people are willing to call an artifact and what they are willing to say it (really) is. An ideal non-linguistic categorization task would not involve language responses, but even then there would be no way to guarantee that participants are not using language, either explicitly or implicitly, to guide their decisions. In a traditional naming task, on the contrary, participants are often forced to choose one name for an artifact, which denies artifacts the opportunity to have more than one acceptable name. In construing *really* judgments as a measure of linguistic description, then, I am allowing linguistic responses to be an acceptable, as opposed to confounding, data modality. Simply asking participants to name an object, or asking them whether a particular name fits for an object, is not enough. The *really* prompt should activate deeper processing of the concept, going beyond the level of successful

discourse reference. If there is a reliable difference between what something is believed to be *really* and what it can be referred to in discourse, the *really* prompt should reveal it.

In addition to using *really* judgments, the work presented here treats judgments of artifacts' category membership as potentially graded. Much of the work from the essentialist domain forced participants to choose one category name only for experimental items (Keil, 1989; Kemler Nelson, Herron, & Morris, 2002; Diesendruck, Markson, & Bloom, 2003). But this type of design does not allow participants to express the non-essentialist opinion that an artifact can belong to more than one category simultaneously, or even the opinion that it would be acceptable to refer to an artifact using different names. In line with Braisby, Franks, and Hampton's (1996) suggestions, the experiments presented below were designed to allow for the expression of a fuller range of possible beliefs.

For instance, the stool/end table discussed at the beginning of this thesis might well be considered both a stool and an end table, although it may be considered more of a stool than an end table, or vice versa. This possibility of multiple category membership and gradation of category membership has often been ignored by essentialists, even though there is evidence suggesting that people judge artifacts in a graded fashion (Estes, 2004). Other researchers (Kalish, 1995; Hampton, Dubois, & Yeh, 2006; Malt & Sloman, 2007) have used rating scales in their designs to allow participants to express category membership gradation as well as membership in more than one category. When these designs have been

used, most participants tend to express some belief in the possibility of multiple category membership and category membership gradation for artifacts, although some participants seem to respond in an essentialist way. However, these trends have been largely ignored by essentialist researchers, who have continued to evaluate their hypotheses with forced-choice designs and demand characteristics that suggest or require that only one category can be endorsed for a particular item.

The Present Experiments

In the experiments presented below, I manipulated four factors that radial theory predicts are relevant to people's judgments of what an artifact is, but which essentialist theories consider less important. Both experiments involved artifacts that were created for one purpose but were being perceived or used as being for another purpose. In Experiment 1, I investigated whether people would be affected by the typicality of the artifact with respect to the new category. I was also interested in seeing whether people's responses about artifacts would differ as a factor of whether they were thinking about how to talk about the artifact or about what they or the person using the artifact would consider it to *really* be. In Experiment 2, I investigated whether people would be affected by the number of people using or perceiving an object in a new way and by the entitlement of those people to make such a decision. The person who created an object might be perceived as being more entitled to change its category than someone with less entitlement – for instance, a shopper.

Experiment 1

In Experiment 1, I tested an artifact-based version of Lakoff's (1987) radial theory of name extension, in which many factors should influence how people talk about artifacts and what they judge the artifacts to be *really*, including appearance, typicality, function, and what the object has been considered to be. In contrast to the radial view, strong essentialist theory would maintain that naming in discourse contexts and judgments of what an artifact *really* is are not necessarily related; successful reference need not take essence into account (Bloom, 2007). However, objects should be considered *really* members of their initial intended category and no other category. A weak essentialist theory would only diverge from its strong counterpart on one point: The weak version would hold that an object's essence, as reflected by *really* judgments, can change. People should be willing to accept a name change as "real" when someone (either the original creator or a secondary agent) recreates the artifact with the intention that it belong to a new category.

To discriminate among these three competing hypotheses (strong essentialism, weak essentialism, and radial theory), I presented college students with pictures of artifacts in a current use associated with one category along with a story indicating that the artifact was originally intended to belong to another category. For example, one story involved an artifact originally intended to be used as a candy dish being used as an ashtray from now on. Half of the test items were typical with respect to their new category (as judged with typicality ratings); the

other half were atypical. Participants rated the suitability of names for each artifact in three different measures: what someone should call it in order to talk to someone else about it, what the participants thought it *really* was, and what they felt the artifact's owner would think it *really* was. I refer to these three measures as Discourse Naming Method, Really Naming Method, and Character Really Naming Method. I included the Character Really Naming Method to assess any potential psychological differentiation of what the participant thought the object *really* was from what the participant felt the object's owner might think it *really* was. If people indicate a discrepancy between these two types of *really* judgments, it would provide additional evidence that any putative essentialism is psychologically relative and context-dependent, and therefore not essentialism at all. A complete list of stimuli can be found in Table 1, an example of the stimuli viewed by participants can be found in Figure 1, and a sample story with questions and rating scales can be found in Appendix A.

Predictions

Because none of the theories discussed above would predict any main effects of Category, Typicality, or Naming Method across all levels of the other variables, I present predictions below for each level of the Naming Method variable. This allowed predictions of main effects of Category and Typicality within each Naming Method. The interaction predicted for the entire data set (all three variables) will be discussed last.

Predictions for strong essentialism. In the Discourse Naming Method, the strong essentialist theory would predict a main effect of Category; participants should prefer the new category name because that is the name that allows them to communicate the current use. There should also be an interaction of Category and Typicality. Participants should prefer the new category name more when the object is typical with respect to the new category than when it is atypical because it is best suited for communication when the object is more typical of that name. Because of this interaction, there should be no main effect of Typicality.

Strong essentialism predicts no differences in results between the Really and Character Really naming methods. For both, there should be a main effect of Category, with the old category being preferred and the new category rejected, because participants should judge each artifact to be *really* whatever it was originally intended to be. There should be no main effect of Typicality or interaction with Category because judgments should be fully governed by beliefs about what the original intended category was.

When considering all three variables together, there should be an interaction of Category with Naming Method: Participants should prefer the old category in the *really* judgments, but the new category in the discourse. There should also be an interaction of Category, Naming Method, and Typicality. In the *really* ratings, participants should consistently prefer the old category names and reject the new category names with no effect of Typicality, but in the Discourse Naming

Method, typicality should affect name preference. The predicted trends for strong essentialist theory are presented in Figure 2.

Predictions for weak essentialism. In the Discourse Naming Method, weak essentialism would predict a main effect of Category: Participants should prefer the new category name because that is the name that allows them to communicate the current use. There should be an interaction of Typicality with Category: When the object is typical with respect to the new category, participants should endorse the new category more than when the object is atypical because it is best suited for communication when the object is more typical of that name. There should be no main effect of Typicality.

The predictions for the Really Naming Method and Character Really Naming Method are the same. The weak essentialist view would allow for essence to change; so when an object's owner has made a salient intentional category change, participants should endorse the new category in their judgments of *really*. There should be a main effect of Category: Participants should prefer the new category name. Specifically, there should be no main effect of Typicality on ratings of name acceptability in the two *really* naming methods. The predicted trends for weak essentialism are presented in Figure 3. Weak essentialism differs from strong essentialism here only in which category name is preferred: In strong essentialism, the old category name should be preferred; in weak essentialism, the new category name should be preferred.

Predictions for radial theory. The predictions of radial theory for the Discourse Naming Method are the same as those of strong and weak essentialism. There should be a main effect of Category: Participants should prefer the new category name. There should be an interaction of Typicality and Category: Participants should prefer the new category name more when the object is typical with respect to the new category than when it is atypical. There should be no main effect of Typicality.

The key predictions for radial theory lie in the predictions for the Really and Character Really naming methods. According to radial theory, the typicality of the test item relative to the new category will mediate the acceptability of either name in the Really Naming Method. Participants may select either or both names. There should be no main effect of Category or Typicality. There should be an interaction, however, with the effect of Category mediated by Typicality. Participants should be more willing to rate the object as *really* the new category when it is typical of that category than when it is atypical.

Radial theory would allow for a differentiation between what the participants think the object really is, and what they think its owner thinks it really is, in the Character Really Naming Method because different people may have different perspectives on the object. Because the character in the story has expressed a bias toward the object – he or she intends to use the object in the new category "from now on" – participants might feel that, from their perspective, the object is "really an X," while also believing that, from the character's perspective, it's "really a

Y." There should be a main effect of Category: Participants should rate the new category name higher than the old category name. There should also be an interaction: When the object is atypical with respect to the new category, participants should still rate the character's belief that the object is really the new category as acceptable, but less so than when the object is typical. There should be no main effect of Typicality. The predicted trends for radial theory are presented in Figure 4.

Method

Participants.

Participants were 24 Lehigh University undergraduates (13 men and 11 women) who participated for course credit.

Design. The first independent variable was Typicality with respect to the new category, with two levels: Typical and Atypical. The second independent variable was Category, with two levels: Old and New. The third independent variable was Naming Method, with three levels: Discourse Judgments, Really Judgments, and Character Really Judgments. The dependent variable was the rating of name acceptability for each of the two potential categories in each of the three response measures.

Category and Naming Method were completely within-subjects variables. That is, all participants evaluated the artifacts with respect to both the old and new categories in each of the three naming methods. Typicality was also a within-subjects variable. Each participant viewed half of the items in the Typical

condition and half of the items in the Atypical condition. However, assignment of pictures to stories was counterbalanced so that half of the participants viewed two typical and two atypical test item tokens, and the other half viewed the corresponding atypical and typical tokens of the respective types. For example, Participant 1 viewed the typical picnic blanket and atypical planter, but Participant 2 viewed the atypical picnic blanket and typical planter.

Materials. The list of stimuli is presented in Table 1. Stimuli were presented in color photographs pasted onto white paper. Each test item type had a typical token and an atypical token. For example, for the picnic blanket, there was one blanket that looked like a typical picnic blanket (see Figure 1a) and one that did not look like a typical picnic blanket (see Figure 1b). Ratings of typicality were obtained with 14 Lehigh University undergraduates who were not involved with the main study. Six pairs of stimuli were rated for typicality with respect to the new category. Of the six, four pairs demonstrated suitable differentiation along a dimension of typicality to the new category: *tablecloth*, *cookie jar*, *stool*, and *candy dish*. The other two stimuli (*desk* and *beer stein*) were included in the data collection, but their data are not presented here.

Test items included pictures of a *tablecloth* (being used as a *picnic blanket*), *cookie jar* (being used as a *planter*), *stool* (being used as an *end table*), and *candy dish* (being used as an *ashtray*). Participants read descriptions of the artifacts as being created for one purpose but used for another function. For example, the *tablecloth/picnic blanket* story featured a main character, Marla, who had decided

to use an old tablecloth as a picnic blanket from now on. The story ends with Marla about to refer to the object (while using it as a picnic blanket) to a friend who knows its intentional history. See Table 2 for sample stories and response options for each Naming Method.

After reading each story, participants then answered six questions about each artifact. All responses were made on a seven-point scale of acceptability of potential names for the artifact. The rating options ranged from "Not acceptable at all" to "Highly acceptable." For the Discourse ratings, participants rated the acceptability of the main character in the story referring to the artifact by using the original category name, and by using the new category name. The main character was addressing another person who knew the intentional history of the object. For the Really judgments, participants rated the acceptability of both saying the object is *really* the new category, and of saying it is *really* the old category. For the Character Really judgment, participants rated the acceptability of saying that the main character thinks that the object is really the new category, and really the old category. Text for all the experimental stories can be found in Appendix A.

Order of questions and the order in which the categories were mentioned in the response questions were counterbalanced across subjects. All participants made the discourse Name Judgments first because each story ended with an unfinished sentence, and it would have disrupted the syntax to have participants answer the discourse question in any position other than first. Half the participants

made the Really judgments second, and half made the Character Really judgments second. Half the participants rated the acceptability of the old category name in all three questions first, and half rated the acceptability of the new category name first.

Procedure. Participants completed the experiment in pairs. The experimenter told them that they were participating in an experiment on naming, and they would be looking at pictures of artifacts, reading descriptions of the objects, and then thinking about what the objects are. Participants then sat down and completed the paper-and-pencil task. When they had completed the ratings, participants were thanked and debriefed.

Results

Creation of Super-Subjects

Because each participant saw only half the items as typical and half as atypical, 12 "super-subjects" were created out of every two participants. Combining each odd-numbered participant (who viewed Typical objects 1, 3, and 5, and Atypical objects 2, 4, and 6) with a respective even-numbered participant (who viewed Typical objects 2, 4, and 6, and Atypical objects 1, 3, and 5) produced a "super-subject" who rated each category for all of the items with both typical and atypical objects.

To ensure that there were no systematic differences between the odd- and even-numbered participants that would preclude their combination, a 2 (odd participant vs. even participant) x 2 (old category vs. new category) x 2 (typical token vs.

atypical token) ANOVA was conducted at each level of the Naming Method variable. At the level of Discourse, there was no significant main effect of participant group (odd- vs. even-numbered participants), $F(1, 22) = 1.33$, $p = .26$. At the level of Really Judgments, there also was no significant effect of participant group, $F(1, 22) = .18$, $p = .67$. Finally, at the level of Character Really Judgments, there was no significant effect of participant group, $F(1, 22) = .02$, $p = .89$. I concluded, therefore, that participants were not behaving systematically differently depending on the actual stimuli they viewed, and elected to combine the participants into super-subjects as planned. This means that although Typicality was not a within-subjects variable, it was treated as if it were a within-subjects variable in the analyses below.

Computation of Means

Mean ratings across stimuli were computed for each super-subject for each Typicality and Category combination, at each level of the Naming Method variable. Thus, for each naming method, there were four cells: Old Category-Typical, Old Category-Atypical, New Category-Typical, and New Category-Atypical. The super-subject means were then averaged at each of the aforementioned cells. These means are presented in Tables 3, 4, and 5, and Figures 5, 6, and 7. I will discuss each level of Naming Method separately.

Discourse Naming Method

Strong essentialism, weak essentialism, and radial theory would all predict a main effect of Category in the Discourse Naming Method: Participants should

rate the new category higher than the old category because their concern is with the best way of achieving reference. The pattern of means was consistent with this prediction, as shown in Table 3. The difference between these means is significant, $F(1, 11) = 19.12$, $p < .01$, $\eta^2 = 0.64$. Participants consistently rated the name associated with the new category as more acceptable for use in discourse than the name associated with the old category.

There was no main effect of Typicality, $F(1, 11) = 0.91$, $p = .36$, $\eta^2 = .08$. Ratings for atypical objects were as high as for typical ones overall. However, there was a significant interaction between Category and Typicality, $F(1, 11) = 8.45$, $p < .05$, $\eta^2 = .43$. When rating objects that were typical with respect to the new category, participants rated the name associated with the new category higher than the name associated with the old category. When rating objects atypical with respect to the new category, the difference was reduced: Participants' preference for the new category name was reduced for objects atypical of that category. Paired-samples t-tests were conducted to examine whether the effect of Category was significant at either level of Typicality. For Typical objects, ratings for the new category were significantly higher than ratings for the old category, $t(23) = 3.20$, $p = .004$. For Atypical objects, the difference between ratings for the new category and ratings for the old category did not reach significance, $t(23) = 1.26$, $p = .22$. This outcome was consistent with the predictions of all three views.

Really Naming Method

If strong essentialism is correct, participants should like the old category name and reject the new category name across both levels of Typicality when considering what the object *really* is. If weak essentialism is correct, participants should endorse the new category and reject the old category across both levels of Typicality. If radial theory is correct, participants should not consistently reject either category, but rather should be affected by the extent to which the object is typical with respect to the new category when making their judgments. When an object is typical, participants should rate it more acceptable of belonging *really* to the new category than when it is atypical. Judgments of acceptability of the old category name should not be affected by typicality with respect to the new category.

The pattern of means is most consistent with the predictions of radial theory. As Table 4 shows, participants rated both the old and new categories as acceptable, but they rated the old category higher than the new category. There was a main effect of Category, $F(1, 11) = 28.03$, $p < .001$, $\eta^2 = .72$. Participants rated the old category consistently higher than the new category. They still rated the new category somewhat acceptable, however, with mean ratings around 4 for both levels of Typicality.

Furthermore, the degree of preference was affected by Typicality: There was a marginally significant interaction of Category and Typicality, $F(1, 11) = 4.17$, $p = .066$, $\eta^2 = .28$. Participants preferred the old category over the new category, but this effect was more pronounced for objects atypical of the new category.

Participants endorsed the old category more, and the new category less, for atypical than for typical objects. Thus when overall preference was for the *new* category (as seen in the Discourse Naming Method), the preference was reduced by low typicality with respect to the new category, and in parallel, here, when overall preference was for the *old* category, the preference was heightened by low typicality with respect to the new category.

There was also a marginal main effect of Typicality, $F(1, 11) = 4.62$, $p = .055$, $\eta^2 = .30$. When averaged across categories, the Atypical objects were rated higher than the Typical objects. This main effect reflects the fact that the Atypical-Old Category cell received the highest mean rating. Because I only manipulated Typicality with respect to the *new* category, these objects may have been considered very good examples of the old category.

Character Really Naming Method

If strong or weak essentialism is correct, then the same predictions that hold for the Really condition should hold for the Character Really condition. For strong essentialism, participants should prefer the old category regardless of typicality. For weak essentialism, participants should prefer the new category regardless of typicality. If radial theory is correct, people might deviate from a one-to-one correspondence between their Really ratings and their Character Really ratings: They might consider that the two perspectives can differ.

The pattern of means is similar to the means in the Really condition, with one notable exception. As Table 5 shows, participants endorsed the new categories to

the same extent in the Character Really condition that they did in the Really condition. But participants endorsed the old categories less in the Character Really condition (at both levels of Typicality) than they did in the Really condition. When considering what the character thought the object *really* was, participants ascribed lower ratings to the old category than when considering what they personally thought the object *really* was. This trend is predicted by radial theory, but not by either essentialist framework.

There was a main effect of Category, $F(1, 11) = 14.58$, $p < .01$, $\eta^2 = .57$. As in the Really condition, participants rated the old category higher than the new category, but they did not reject the new category. New categories received a mean rating of 3.89, reflecting a trend of rating the possibility of thinking of the object as *really* the new category as somewhat acceptable.

The interaction of Category and Typicality was in the same direction as the interaction in the Really condition, and reached marginal significance, $F(1, 11) = 3.18$, $p = .10$. Participants endorsed the old category, on average, one point more strongly when the object was Atypical than when it was Typical. Endorsement of the new category declined when the object was Atypical, as expected.

Once again, there was a main effect of Typicality, $F(1, 11) = 5.70$, $p < .05$, $\eta^2 = .34$. The mean rating for Typical objects was higher than the mean rating for Atypical objects, as in the Really Naming Method.

Differences Across Naming Methods

To investigate the differences across the three naming methods, a 2 (old vs. new) x 2 (typical vs. atypical) x 3 (Discourse vs. Really vs. Character Really) ANOVA was conducted. Naming Methods was a within-subjects variable. There was no main effect of Naming Method, $F(2, 33) = 1.11$, $p = .34$. There was also no significant three-way interaction of Category, Typicality, and Naming Method, $F(2, 33) = .06$, $p = .95$. However, these results would be expected given that there were no apparent systematic relationships among the Naming Methods in which participants made consistently higher or lower ratings across all levels of the within-subjects variables as a function of Naming Method. Rather, any effect of Naming Method would be manifested in an interaction. The interaction of Category with Naming Method was highly significant, $F(2, 33) = 25.82$, $p < .001$, $\eta^2 = .61$. This interaction was not significant, however, when the Discourse Naming Method was excluded from the analysis, $F(1, 33) = .77$, $p = .40$. Thus, the significant interaction of Category with Naming Method is a by-product of the differences between ratings in the Discourse Naming Method and ratings in the Really and Character Really naming methods. Collapsing across conditions of Typicality, the old category names were rated higher than the new names in the Really and Character Really naming methods, but the new category names were rated higher than the old names in the Discourse Naming Method. These means are presented in Table 6 and Figure 5.

Discussion

Strong essentialism predicts that people should endorse only the originally intended category when deciding what an artifact *really* is. Weak essentialism predicts that people should endorse only the new intended category. Radial theory predicts that people may endorse either, neither, or both categories, and should be affected by other factors such as typicality. In Experiment 1, I found that people think it is acceptable to use either of the names associated with the old or new category when thinking about what an object whose use has been permanently changed *really* is. Participants rated the old category names higher than the new category names, but this trend was diminished when the objects were typical with respect to the new category. Moreover, participants consistently preferred the new category names when thinking about making successful reference to the objects (Discourse Naming Method), reflecting consideration of pragmatic goals. These results, particularly the interaction of Category with Typicality in all three conditions, suggest that cognition is not governed by either a weak or strong brand of essentialism. Contrary to essentialist predictions, people's judgments of what an artifact *really* is are affected by factors such as the naming context and the object's degree of typicality to the new category.

By giving participants the option of endorsing both or neither categories rather than one or the other, this study overcame a methodological error of earlier work, in which participants are forced to endorse one and only one category for an artifact. The methodology in my study allows for a more ecologically valid

linguistic description of the artifacts, in which a gradation of multi-category preference is encouraged rather than precluded. Although participants preferred the old category names in the *really* conditions, they did not consistently reject the new category names. Means around 4 for the new category names reflect approval of the new category names; they were deemed acceptable although not preferable. Participants may have been indicating, "It would be OK to think that the object is really the new category, but it might make more sense to think of it as the old category." Evidence that participants favor the old category does not support strong essentialism; if strong essentialism were correct, participants would not only endorse the old category – they would reject the new category, too. Participants did not reject the new category here.

Two variables were manipulated in this experiment: the naming context and the degree of typicality to the new category. In most real-world situations, there would be many more factors that would influence cognitive decision-making about category membership (e.g. how many people think that an object belongs to a particular category, entitlement of a person to change an artifact's function, functionality, salience of physical features associated with either the old or new category, availability of language to categorize the object with respect to both categories simultaneously, and so on). In Experiment 2, I examined the effects of some of these other factors in radial theory.

Despite the presence of just two of these factors, Experiment 1 demonstrated that people are sensitive to both Naming Method and Typicality when considering

how to talk about artifacts and how to categorize them when prompted to think about what they *really* are. There is no simple, one-to-one correspondence between a creator's categorical intention and the cognition of another person, even when that person is aware of the object's intentional history. People clearly take intention into consideration. But this experiment demonstrates that to say creator's intention is the only factor people consider in naming or cognition of artifacts would be inaccurate.

Radial theory might suggest that people would endorse the category associated with the object's current use, but they did not always do this. There are plausible explanations for why participants did not endorse the new category names more readily. First, the typicality of the objects with respect to the old category most likely varied unintentionally with the typicality with respect to the new category. For instance, the typical picnic blanket started out as a tablecloth (see Figure 1a), but it might have looked less typical as a tablecloth than the atypical picnic blanket (see Figure 1b). Thus, participants may have been influenced by typicality with respect to the old category as well as typicality with respect to the new category, but I failed to take this into consideration when choosing the stimuli. When viewing objects that were atypical with respect to the new category, participants may have considered the object quite typical with respect to the old category, which would have resulted in their endorsing the old category more, and the new category less. Future research that seeks to expand on these findings should attempt to keep typicality with respect to the old category constant while

varying typicality with respect to the new category. This could present practical difficulties, as it seems intuitively unlikely to find many stimuli that vary along a featural dimension with respect to one category without varying along a featural dimension of a contrasting category.

The above explanation does not account for preference for the old category name in the Really Naming Method when the objects were sufficiently typical with respect to the new category. A possible explanation for the tendency to endorse the old category more than new category in the Really judgments is the dichotomy of categories in the response ratings. Although it was a step in the right direction to allow for gradation of response along two categories instead of forcing a unilateral categorical choice on the participants, the response forms used here may not have elicited truly valid measures of artifact cognition. It may be the case that participants did not think the object was *really* a picnic blanket or a tablecloth, but something else entirely, such as a "sheet" or a "cover for eating." While the rating forms gave participants the ability to reject both possible categories, no participants actually did so. In other words, participants most likely reacted to a perceived experimental demand characteristic that they should endorse at least one of the categories, even though the experimental instructions informed them that it would be acceptable to endorse both, either, or neither of the categories. Instead of using rating forms for two possible answers, it might be beneficial to present participants with more possible answers which they can simply endorse (with an "X"), given that they are permitted to endorse more than

one response. For example, participants could choose answers such as "It's really both," "It's more like a picnic blanket than a tablecloth," and "It's really a picnic blanket," which would present a more well-rounded, yet discrete, linguistic description of the objects. Malt and Sloman (2007) employed a similar design and obtained results somewhat less suggestive of original-category preferences than those reported here.

In a follow-up study not analyzed here, I presented participants with the same pictures and stories, but had them answer the questions orally instead of on rating forms. For instance, at the end of each story, participants actually finished the character's sentence in order to refer to the object. They then answered the questions "What is it really?" and "What does [the character] think it really is?". Although participants tended to choose one-word categorical answers to these questions, many participants responded with elaborations ("cover for eating"), superordinate categories ("cloth"), or expressions indicating both categories ("it's really both"). Thus, people may think about objects in ways that are not succinctly expressed by one category, or by ratings of two potential categories. By getting participants to verbalize, they should feel free to express what they are thinking as accurately as possible, without relying on a single word to sum up their thoughts.

In addition to having all participants rate both the old and new categories, Experiment 1 also required all participants to do so in all three Naming Methods. This may have contributed to the relatively high ratings for the old categories in the Really and Character Really naming methods, and the relatively high ratings

for the new categories in the Discourse Naming Method. Because Discourse Naming Method was always presented first, participants may have perceived an experimental demand to polarize their category preferences across Naming Method. For instance, a participant who rated the new category higher than the old category in the Discourse Naming Method may have perceived pressure to indicate that the object was really the old category in the Really and Character Really naming methods. In Experiment 2, I attempted to reduce this experimental demand by making Naming Method a completely between-subjects variable.

Another implication of the findings reported here is that there may be individual differences in degree of psychological essentialism. A small subset of the participants behaved more like strong essentialists than the majority of participants, consistently rejecting the new category (as indicated by a rating of 1 or 2) and strongly endorsing the old category (as indicated by a rating of 6 or 7). These participants may have been exhibiting a need for cognitive order (Webster & Stewart, 1973). An interesting idea for future research would be to separate strong essentialists from non-essentialists in a pre-test and then investigate potential differences in naming and artifact cognition between the two groups. Just as there may be no single predictor of how an artifact will be categorized in any situation, there may be no single framework to which all people conform when engaging in those activities.

In conclusion, then, the findings of Experiment 1 indicate that people, on average, do not behave like essentialists when naming artifacts that have changed

categories, or when thinking about what these artifacts *really* are. People are sensitive to other cues in the context, such as the pragmatic goal of referring to the object in discourse, or the extent to which the object is typical with respect to the new category. These results support the radial view of artifact cognition and naming, but not an essentialist view. Typicality to the new category and pragmatics are just two of many features that may play a role in artifact cognition, according to the radial view.

In Experiment 2, I sought to control the effects of typicality by using descriptions of artifacts instead of photographs. By controlling these descriptions for typicality with respect to both the old and new categories, I attempted to reduce the likelihood that either name could be considered more physically plausible to refer to the object based on the object description. In doing so, I investigated the effects of several variables that were not manipulated in Experiment 1.

Experiment 2

According to radial theory, typicality is neither the only nor the most important factor that can affect what an artifact is thought to be. Experiment 1 examined the effect of typicality on name choice in Discourse and Really scenarios when the artifact's owner had decided to change its use permanently. Experiment 2 explored the effects of Entitlement of a person to change an artifact's category and Numerosity of opinion on name choices for artifacts typical with respect to two categories (e.g. an object that could plausibly be both a stool and an end

table). By Entitlement, I mean how much authority a person has to decide that an artifact now belongs to a new category. By Numerosity, I mean how many people think that an artifact belongs to a particular category. In order to test essentialist predictions, information about original intended category was always included.

People are affected by expert opinion when thinking about natural kind categories (Braisby, 2001). In this experiment, I explored the possibility that people are affected by expertise with respect to artifacts – this expertise amounts to entitlement to change the artifact's category. I conceived three levels of Entitlement, listed with decreasing authority: creator, owner, and shopper. If Braisby's results can be replicated in the domain of artifacts, radial theory suggests that people should be influenced by a creator most of all, by an owner less, and by a shopper least of all, because of their decreasing levels of entitlement. Strong essentialist theory would predict that people side with the creator's original intent when deciding what an object really is (even if the creator has decided that the object is now something else). Weak essentialist theory would predict that a change in essence would be allowable, presumably by the creator or owner.

People's judgments of what an artifact is "really *for*" are affected by whether one or many people are using it in a given function (Siegel & Callanan, 2007). Those results pertained only to function; this experiment investigated whether the results could be replicated when asking participants to consider what the artifact really *is*. Thus, in the stories used in Experiment 2, some number of people

disagreed with the entitled person's opinion (the Numerosity variable). Radial theory suggests that numerosity should have an effect on such judgments: If many people think an object is a stool, then the participant should be more likely to agree than if only one person had thought so. Essentialism predicts that the number of people that share an opinion should not matter.

Experiment 2 tested the effects of Entitlement, Numerosity, and Naming Method on ratings of the suitability of potential names for artifacts. I presented college students with stories about artifacts created for one purpose but perceived as or used for another purpose. All participants read one story each. Half the participants read about an artifact that one person wanted to use for the new purpose; the other half read about an artifact that many people wanted to use for the new purpose. Within each half, a third of the participants read about the object's creator(s) making the change; a third read about the owner(s) making the change; and a third read about one or more shoppers desiring to make the change. After reading each story, all participants rated possible answers to questions about the artifact. A third of the participants rated the acceptability of talking about the artifact using the names associated with both the old and new categories. A third of the participants rated the acceptability of both thinking that the object is really the old category, and thinking that it is really the new category. The final third of the participants answered the same question without the *really* prompt. This served as a control measure to determine whether people were responding differently based solely on the presence of the word "really." The control scripts

were identical to the *really* scripts, but in the control scripts, the word "really" was deleted from the instructions and rating scales. A sample story with response options for the Discourse and Really naming methods can be found in Table 7.

Predictions

Predictions are presented below for each level of Naming Method because there are no predicted main effects of Naming Method across all levels of the other variables in any of the three theories. The notable difference between the predictions of the essentialist theories and radial theory is that radial theory is the only one that predicts judgments of both categories being mediated by Entitlement and Numerosity at all levels of Naming Method. Neither essentialist theory would allow for effects of both Entitlement and Numerosity on judgments of what the artifact (really) is. The predictions for Discourse Naming Methods are the same in all theories: People should respond to the pragmatic value of using a particular name for communicative purposes.

Predictions for Discourse Naming Method (all theories). All theories predict effects of pragmatic variables on what someone might choose to call an artifact in discourse. There should be an interaction of Category with both Numerosity and Entitlement. Liking for the new Category should increase with greater Numerosity and Entitlement, and liking for the old Category should decrease with greater Numerosity and Entitlement. The predicted trends for Discourse Naming Method are presented in Figure 6.

Predictions for strong essentialism, Really Naming Method. If strong essentialism is correct, people should respect the original intentions of the creator in the Really Naming Method. There should be a main effect of Category; the old name should be preferred over the new name. There should be no other main effects or interactions. Participants should never indicate that an artifact belongs *really* to the new category or to both categories. The predicted trends of strong essentialism for the Really Naming Method are presented in Figure 7.

Predictions for weak essentialism, Really Naming Method. If weak essentialism is correct, people should respect the decision of a creator or owner to change the categorical status of an artifact. The object is only really the new category when the creator or owner has made the decision; anyone else cannot change what the object *really* is. In the Really Naming Method, there should be an interaction of Entitlement and Category. Ratings should be high for the new category name at the creator and owner levels of Entitlement, and low for the new category name in the shopper level, regardless of Numerosity. Ratings for the old category name should be low in the creator and owner levels, and high in the shopper level, regardless of Numerosity. As with strong essentialism, weak essentialism predicts that no participants should ever endorse both categories in the Really Naming Method. The predicted trends of weak essentialism for the Really Naming Method are presented in Figure 8.

Predictions for radial theory, Really Naming Method. Radial theory does not recognize an essence; an object may shift category membership or belong *really*

to more than one category, and judgments of this should depend on the strength of the evidence favoring one category or another. Thus, Numerosity and Entitlement should both have an effect on preferences for the old versus the new category name, but there can be no stable prediction of consistent rejection or endorsement across all levels of any variable simply as a function of some other variable, as essentialism would predict. If radial theory is correct, ratings should be affected by all of the information made available to the participants. In effect, the predictions for radial theory for the Really Naming Method are the same as the predictions for the Discourse Naming Method. The same factors that affect people's judgments of what an object can be called should affect their judgments of what the object *really* is.

Perhaps most importantly, radial theory would predict that participants may be willing to say that an artifact could belong *really* to both categories simultaneously. Neither essentialist theory allows for this. Whether participants think that an object can belong *really* to both categories cannot be evaluated by looking at main effects or interactions, however. Instead, this possibility will be examined by looking at response frequencies to see whether there was any consistent endorsement of both categories. This will be done descriptively; a correlation measure would only indicate response consistency (e.g. rejecting both the old and new and endorsing both at the same levels would produce correlations of 1.0 despite indicating vastly different opinions). Thus correlation would not be a valid indicator of participants' willingness to *endorse* both categories.

Method

Participants. Three hundred twenty-four Lehigh University undergraduates participated in exchange for course credit. Gender information was not recorded.

Design. The first independent variable was Entitlement, with three levels: Creator, Owner, and Shopper. The second independent variable was Numerosity, with two levels: One and Many. The third independent variable was Category, with two levels: Old and New. The final independent variable was Naming Method, with three levels: Discourse Judgments, Really Judgments, and Control ("What is it") Judgments. The dependent variable was the rating of name acceptability for each of the two potential categories in each of the three response measures.

Naming Method, Numerosity, Entitlement, and Item were between-subject variables, while Category was a within-subject variable. Participants read about only one item but rated it with respect to both the old and new categories. One-sixth of the participants read about the first item, one-sixth read about the second item, and so on.

Materials. In a pretest, feature listings were obtained for the items of interest. Fifteen Lehigh University undergraduates who did not take part in the larger study participated in exchange for candy. They were given 60 seconds to list the features of 18 artifacts (*end table, stool, tablecloth, picnic blanket, vase, pencil holder, candy dish, ashtray, umbrella, lampshade, cookie jar, planter, butter knife, letter opener, Frisbee, plastic plate, watering can, teapot*). Any features

listed by at least one-third of the participants for any item were included for consideration in constructing descriptions of the items, as suggested by Malt and Smith (1984). After these features were reviewed, I constructed descriptions of nine objects, with consideration to make each description plausible with respect to two potential categories. For instance, the first object was described in such as way as to appear plausible as both a stool and an end table. The goal was to describe the physical characteristics of objects that could plausibly be considered members of two different artifact categories. To do this, I used personal judgment to eliminate some item characteristics that were listed for one category of a pair but would negatively affect the plausibility with respect to the other category. For example, "having a handle" was a common feature listed for *umbrella*. But "having a handle" makes no sense if the object is being considered as a *lampshade*. So "having a handle" was not included in the item description for this object.

In a second pretest, ratings of typicality were obtained on a scale from 0 ("not a category member at all") to 7 ("highly typical") for the nine item descriptions to ensure that they would be perceived as sufficiently typical of both potential target categories. Thirty Lehigh University undergraduates who did not participate in the first pretest participated in exchange for candy. All 30 read the nine item descriptions. However, half rated each item for typicality with respect to one target category (e.g. *picnic blanket*), and the other half rated each for typicality with respect to the other target category (e.g. *tablecloth*). I decided beforehand

that any pair with a difference in means of 1.00 or greater would be eliminated, as would any pair which contained at least one category mean of 3.00 or below. The differences between mean typicality ratings for the old and new categories for *cookie jar/planter*, *vase/pencil holder*, and *plastic plate/Frisbee* were the largest differences in the nine pairs used and so were eliminated from use in the actual study. Of the six pairs accepted for use, the largest difference in means between old and new category was 0.74 (for *umbrella/lampshade*; *umbrella* received the higher rating). No means were below 5.30. The six test items are listed in Table 8 with their corresponding typicality ratings. The item pairs were: *End table/stool*, *candy dish/ashtray*, *tablecloth/picnic blanket*, *teapot/watering can*, *butter knife/letter opener*, and *umbrella/lampshade.*

The participants in the main study did not participate in either of the pretests. Stimuli were presented to participants in written form. Participants read descriptions of the artifacts being created for one purpose. Later, a character in the story (the creator, the owner, or a shopper) decides that it is not the original category but the new category. Rating scales were placed at the end of each story. Participants in the Discourse Naming Method rated the acceptability of using the old or new name in conversation for each story. Participants in the Really Naming Method rated the acceptability of thinking that the object is *really* the old or new category. Participants in the Control Naming Method rated the acceptability of thinking that the object is the old or new category (no *really* prompt). The rating scales contained two absolute extremes ("Absolutely not " and "It is a [stool] and

nothing else"), as well as a graded scale (ranging from "Barely acceptable" to "Very acceptable"), as implemented by Kalish (1995). See Appendix B for the descriptions of all stimuli, Appendix C for instructions in the Discourse Naming Method with a set of all stories and rating scales for one stimulus, and Appendix D for instructions in the Really Naming Method with a set of all stories and rating scales for one stimulus.

The order of questions was counterbalanced across all subjects so that half of the participants rated the old category first and half rated the new category first.

Procedure. Participants completed the experiment in groups or alone. For each scenario, they read the description of the artifact and the experimental story and then rated the acceptability of the old and new names. After rating each name, they explained their choices by writing a brief explanation under the rating scales. When the participants completed all the ratings, they were thanked and debriefed.

Results

Check for Item Consistency

A key assumption of this experiment was that all items were sufficiently typical with respect to each potential target category. However, because typicality cannot be controlled perfectly across all items and participants, Tukey's HSD was conducted as a post-hoc examination of the Item variable in an ANOVA that included all data for all Naming Methods. The *umbrella/lampshade* item ($\bar{x}$ = 4.86) was rated significantly lower than both *tablecloth/picnic blanket* ($\bar{x}$ = 5.94, $p < .001$) and *teapot/watering can* ($\bar{x}$ = 5.51, $p = .04$). Due to this finding,

umbrella/lampshade was removed from further analyses. No other item was rated significantly lower than any other item, so the remaining five items were included in the other analyses.

Effect of Really Prompt

To examine whether there were any systematic differences among responses in the Really Naming Method and the Control Naming Method, a 2 (Naming Method: Control vs. Really) × 2 (Numerosity: One vs. Many) × 3 (Entitlement: Creator vs. Owner vs. Shopper) × 2 (Category: Old vs. New) ANOVA was conducted. There were no significant differences between ratings in the Control Naming Method and the Really Naming Method. From this point on, the Control Naming Method data were not considered. The means for the Really Naming Method are presented in Table 9; those for the Control Naming Method are presented in Table 10.

Discourse Naming Method

All three theories predict no main effects in the Discourse Naming Method, but a three-way interaction of Category, Numerosity, and Entitlement. As the level of Entitlement progresses from creator to shopper, participants should place less weight on the entitled person's opinion in deciding how to refer to the object in discourse. Likewise, evidence of many people supporting a point of view from each of those levels should raise the relevance of using the associated name, as reflected in participants' ratings. The pattern of means, presented in Table 11, was not consistent with these predictions. There was a main effect of Category, F(1,

84) = 7.33, p = .008, η^2 = .08. Overall, ratings for the old category ($\bar{x}$ = 5.86) were higher than those for the new category ($\bar{x}$ = 5.22). The three-way interaction of Category, Numerosity, and Entitlement did not reach significance, F(2, 84) = 0.86, p = .43. There were no other significant main effects or interactions.

Really Naming Method

If strong essentialism is correct, people should judge the object to be "really" whatever it was originally intended to be, regardless of level of Entitlement or Numerosity. There should be a main effect of Category, with the old category preferred over the new category. If weak essentialism is correct, people should respect the decision of the creator or owner to change the category of the object. There should be an interaction of Entitlement and Category: Ratings should be high for the new category at the creator and owner levels of Entitlement, and low for the new category name at the shopper level, regardless of Numerosity. Conversely, ratings for the old category name should be low in the creator and owner levels, and high in the shopper level, regardless of Numerosity. If radial theory is correct, there should be no main effects, but the same three-way interaction that was predicted for Discourse Naming Method should obtain in the data.

The pattern of means, presented in Table 9, is most consistent with the predictions of radial theory. There were no significant main effects or interactions, and the predicted three-way interaction did not reach significance, F(2, 84) = 0.39, p = .68. The key predictions of strong essentialism (preference for

old category) and weak essentialism (preference for new category) also were not supported by the data. On the contrary, people were willing to endorse both the old ($\bar{x}$ = 5.42) and the new ($\bar{x}$ = 5.52) categories well above the midpoint of the scale; in fact, ratings for the new category were slightly higher than ratings for the old category. Only radial theory allows for this outcome.

Analysis of Really and Discourse Naming Methods Combined

To investigate any main effects or interactions at play between the two principal naming methods, I analyzed the data from the Really and Discourse naming methods in a 2 (Naming Method: Really vs. Discourse) × 2 (Numerosity: One vs. Many) × 3 (Entitlement: Creator vs. Owner vs. Shopper) × 2 (Category: Old vs. New) ANOVA. There were no significant main effects or interactions. The interaction of Category with Naming Method did reach marginal significance, however, $F(1, 168) = 3.60, p = .06, \eta^2 = .02$. In the Really Naming Method, ratings for the Old Category ($\bar{x}$ = 5.42) were slightly lower than those for the New Category ($\bar{x}$ = 5.52), but this trend was reversed in the Discourse Naming Method, where the ratings for the Old Category ($\bar{x}$ = 5.86) were somewhat higher than those for the New Category ($\bar{x}$ = 5.22).

Correlational Analyses

I calculated Kendall tau coefficients between ratings in the typicality pretest and in the main experiment to examine the strength of association between typicality ratings and ratings in the experiment. This was done to assess the possibility that participants in the main experiment were reacting primarily to each item's

typicality rather than the variables of interest. The typicality ratings for the Old categories were marginally significantly correlated with the ratings for the Old categories in the experiment, $\tau = .69$, $p = .056$. The typicality ratings for the New categories were not correlated with the ratings for New categories in the experiment, $\tau = -.14$, $p = .70$. Given these findings, it is possible that participants in the experiment tended to rate all the items favorably with respect to the Old categories because of the overall typicality of each item with respect to the Old categories. The lack of a significant correlation between typicality ratings for the New categories and ratings of the New categories in the experiment may have been due to the restricted range of responses in the typicality ratings, in which all items received mostly high ratings and the standard deviation for the New categories was .41.

Next, I checked for any correlations between ratings for the Old and New categories in the experiment. A Pearson correlation of 1.0 would indicate a tendency to rate the Old and New categories identically; a correlation of -1.0 would indicate a tendency to dichotomize responses, rating the Old Category at the opposite end of the scale from the New Category. I calculated Pearson correlations between the Old and New categories within each Naming Method separately. There were no significant correlations in the Discourse Naming Method, $r = .02$, $p = .87$ (see Figure 9), but there was a significant negative correlation between Old Category ratings and New Category ratings in the Really Naming Method, $r = -.28$, $p = .007$ (see Figure 10). These two correlations were

significantly different, $z = 2.03$, $p = .02$. On the surface, this suggests a response dichotomization among participants in the Really Naming Method: When the Old Category was rated high, the New Category was rated low. In the discussion section, I present an argument for why these correlations do not correspond to such a trend in the data.

Participant Justifications

In addition to rating the artifacts with respect to the Old and New categories, participants in Experiment 2 also explained their ratings in writing. This self-report measure was included to investigate which factors people considered when they made their ratings. To the extent that people have introspective access to the basis for their decisions (Ericsson & Simon, 1993), people should be able to report what influenced their ratings. For example, if participants considered the essence of an artifact to be its intended category, then they should have acknowledged this belief when explaining their category ratings for each artifact. The participants' justifications were coded into five categories, based on what they mentioned considering when making their decisions. The codings were based solely on the written justifications, not on the participants' ratings. The five categories were *radial*, *physical*, *functional*, *essentialist*, and *other*. Examples of each justification category are presented in Table 12.

Justifications were considered radial if they included more than one factor that contributed to the object's rating. For instance, many radial responses mentioned consideration of physical plausibility ("It has the features of a stool") and

functional plausibility ("It is being used as a stool, so it can be thought of as a stool"). Thus, people who responded in a radial manner did not consider any one factor, but indicated consideration of multiple factors, which is what radial theory predicts for artifact cognition. Justifications were considered physical if they only mentioned the physical plausibility of the object with respect to either category. Justifications were considered functional if they only mentioned the functional plausibility of the object with respect to either category. Justifications were considered essentialist if they mentioned only a creator's or owner's intention. Justifications coded as *other* did not fall into any of the other categories and tended to contain either incoherent or extraneous responses. For example, one participant's justification for *tablecloth/picnic blanket* cited having a similar tablecloth at home.

Once all justifications had been coded, I calculated the percentage of justifications for each of the five categories in each of the naming methods. As shown in Table 13, the most common type of justification was radial; more people indicated consideration of multiple factors than any single factor. Approximately half of the justifications were radial, roughly 40% were either solely functional or physical, and less than 5% were indicative of essentialist beliefs. In the Really Naming Method, 22 of the 50 radial justifications mentioned intention as a factor among others; 21 of these mentioned the creator and 1 mentioned the owner. In the Discourse Naming Method, 21 of the 53 radial justifications mentioned intention as a factor among others; 20 of these mentioned the creator and 1

mentioned the owner. Thus, although these participants did not indicate a strict essentialism in their justifications, they did indicate a consideration of creator's or owner's intention.

I was also interested to see whether the justifications contained any mention of the Entitlement variable, and if essentialist justifications were more likely than to contain references to Entitlement than the other justifications. This was not the case. In the Really Naming Method, 11 radial justifications mentioned either the owner's or shopper's entitlement to decide what an object is; only one essentialist justification mentioned Entitlement. In the Discourse Naming Method, eight radial justifications mentioned Entitlement; none of the essentialist justifications did so. Thus, most participants did not explicitly list Entitlement as a factor in their rating decisions, but even when they did list it, they tended to be listed as part of radial, not essentialist, justifications.

To check whether participants' justifications corresponded to their ratings for the items, I calculated mean ratings for the Old and New categories in each Naming Method as a function of Participant Justification. The means for the Really Naming Method are presented in Table 14, and those for the Discourse Naming Method are presented in Table 15. These means revealed several interesting trends. First, mean ratings by participants who made radial justifications tended to range from 5 to 6 for both the Old and New categories, regardless of Naming Method. These means were similar to the overall means in the main study; people who made radial justifications tended to view the items as

belonging plausibly to both categories. However, participants who gave radial justifications that included consideration of intentionality rated the Old Category higher than the New Category in both the Really Naming Method (Old Category $\bar{x} = 5.95$, New Category $\bar{x} = 4.86$) and the Discourse Naming Method (Old Category $\bar{x} = 6.05$, New Category $\bar{x} = 4.81$). These participants' ratings were consistent with their declared consideration of original intentionality: They preferred the Old Category over the New Category, although they did not consistently reject the New Category.

Another trend in both Naming Methods was for participants who gave functional justifications to rate *both* the Old and New categories comparably high on the scale. This was particularly pronounced in the Really Naming Method (see Table 14), where the means were greater than 6 for both categories for participants who gave functional justifications. Because item descriptions were constructed to appear plausible as both categories, it makes sense that participants who reported considering only functional plausibility would find the items to be highly acceptable as both categories.

Finally, participants who gave essentialist justifications showed an essentialist trend in their rating choices, but this trend was obscured by the small number of people who gave essentialist justifications and the varying directionality of their essentialism as expressed on the rating scales. In other words, some people were essentialists in favor of the Old Category, and some were essentialists in favor of the New Category, but these people only constituted 4% of the participants

overall. In the Discourse Naming Method (see Table 15), the mean rating for the Old Category for participants who gave essentialist justifications ($\bar{x} = 6.33$) was considerably higher than the mean rating for the New Category ($\bar{x} = 3.33$). This is somewhat consistent with essentialist predictions, yet a mean rating of 3.33 does not indicate wholesale rejection of the New Category. If anything, it might be argued that the participants who gave essentialist justifications behaved more like essentialists than the other participants when rating the items, but they neither consistently rejected the New Category, counter to the predictions of strong essentialism, nor consistently rejected the Old Category, counter to the predictions of weak essentialism.

Discussion

In Experiment 2, I found that the overwhelming majority of people do not engage in essentialist decision-making or explanation when thinking about what artifacts *really* are or how to talk about artifacts, even when the artifacts are being used for a purpose other than the originally intended one. Although some people either behaved as if they had essentialist beliefs or actually professed them, this only accounted for about 5% of the participants. The rest of the participants neither made essentialist ratings nor expressed strictly essentialist beliefs regarding artifact cognition and reference. These results do not support either strong essentialism, which predicts that people should consistently endorse only the old category, or weak essentialism, which predicts that people should consistently endorse only the new category when an intentional change has been

made by the object's owner or creator. Of the three competing hypotheses presented here, only radial theory's hypothesis included simultaneous endorsement of both categories as an acceptable outcome. According to radial theory, people may be willing to believe that an artifact can belong to multiple categories, can be called by multiple names, and can change categories; more generally, people may not have to believe that an object belongs *really* to any category. The results of Experiment 2 support radial theory, but not strong essentialism or weak essentialism.

This experiment also demonstrates that asking people what an artifact *really* is does not disclose an underlying essentialism that previous studies failed to reveal. Asking people what an artifact *really* is produced the same results as asking them simply what it is, or even asking them how they might refer to it. Thus, previous claims that people do not adhere to a strict essentialism for artifacts (Malt & Johnson, 1992; Malt & Sloman, 2007) were not misconstruing people's beliefs; Experiment 2's results suggest that there is no reliable difference between what people are willing to call something and what they think it *really* is. This claim is supported not only by the rating-scale data, but by people's written explanations of their ratings: More people in both the Really and Discourse naming methods accounted for their choices with reference to multiple factors than any one factor, and less than 5% of people accounted for their choices with reference to solely essentialist factors. Considering multiple factors is the hallmark of radial theory; people should use all the information available to them when thinking about what

an object is or how to refer to it. That is precisely what the participants did, without any encouragement or instruction to do so. Furthermore, the ratings for the Old Category were slightly higher, and the ratings for the New Category slightly lower, in the Discourse Naming Method than in the Really Naming Method, although these differences did not reach significance. Strong essentialism predicted just the opposite trend: People might be willing to *call* the object by the new category name, but they should not believe that that is what it *really* is. The actual results might have reflected a belief that it was slightly more appropriate to call it by the old name out of politeness or respect for the creator (who, in the vignettes, also had a goal to sell the artifact), but that the physical and functional plausibility of the artifact to belong to both categories necessitated an acknowledgement that it could *really* belong to both categories. Regardless of whether this trend was meaningful, people rated the old and new categories well above the midpoint of the scale for both the Discourse and Really naming methods.

Participants' responses to the *really* prompt revealed no universal, deep-seated essentialist reality to artifact cognition. The few people who behaved like essentialists did so with and without the *really* prompt. Thus, people who do subscribe to essentialism will do so in situations that offer little or no provocation of their essentialism. The *really* prompt had no effect on responses in Experiment 2 for essentialists and non-essentialists.

Some of the results of Experiment 2 were less cut-and-dry. The two variables introduced to build on the results of Experiment 1 – Numerosity and Entitlement – failed to produce significant main effects or interactions. Overall, people apparently did not take these variables into account when making their ratings or justifications. Coding the justifications offered some insight into this failure. Roughly 80% of participants offered explanations that referred to physical plausibility, functional plausibility, or a combination of factors (which always included either physical or functional plausibility or both). This may have been an outgrowth of my effort to control for physical and functional plausibility; it is possible (and, I think, likely) that functional plausibility was better controlled than physical plausibility with respect to both potential categories. When people offered functional justifications, they tended to point out that the artifacts were sufficiently functionally plausible with respect to both categories. As noted earlier, these participants rated both categories higher than any other participant group when grouped by justification type. But the physical justifications tended to point out physical shortcomings with respect to one category or another. The typicality ratings I obtained for the item descriptions in the second pretest revealed some discrepancies between typicality with respect to the first and second category for each item; no item was perfectly matched on both counts. It may have been the case that the notable differences were primarily physical, not functional, in nature. The item descriptions were also quite detailed with respect to physical properties, while Entitlement and Numerosity were mentioned more

subtly. People may have simply paid more attention to physical and functional features than the variables of interest.

If this reasoning is correct, then the items I used might have been *too* acceptable with respect to both categories to elicit any effects of Numerosity or Entitlement. In that case, it might be more constructive to use items slightly less typical with respect to either category. As a follow-up test of this possibility, I analyzed the data for the *umbrella/lampshade* item independently of the other items. For the main analysis, I had omitted this item because its ratings were significantly lower than those for some of the other items. But in hindsight it is possible that the relatively low typicality for this item could actually have elicited effects of Numerosity and Entitlement. The data for that item, however, revealed no effects of either variable. Future work aiming to find effects of Entitlement or Numerosity might benefit from employing items only *slightly* typical with respect to both categories; if the object is not strongly plausible as a member of either target category, then perhaps Entitlement and Numerosity could have a discernible effect, which might reflect how they have their effects in the real world. The *umbrella/lampshade* item had been rated more typical of an umbrella than a lampshade in the pretest; so it may have failed to elicit any interesting effects because people were responding to its physical plausibility as an umbrella and its functional plausibility as both an umbrella and a lampshade.

I am not aware of any other experiment that has manipulated Entitlement in an artifact domain. More careful investigation of the roles of a creator versus owner

versus shopper needs to be done to ascertain whether people do pay attention to Entitlement for artifacts the way they do for natural kinds (Braisby, 2001). Although some of the participants in Experiment 2 mentioned considering Entitlement in their justifications, this variable did not have a significant effect on the ratings overall, and there were not enough of these participants to warrant running an analysis of their responses as a function of the other variables.

The failure of this experiment to replicate Siegel and Callanan's (2007) Numerosity effect is more troubling. In their study, as in mine, Numerosity was a between-subjects variable; thus, the effect of Numerosity they obtained cannot be accounted for by response dichotomization by participants who noticed that one person made the decision in one story but many people made the decision in another story. One possibility for the different results is the different ways in which the two experiments operationalized Numerosity. Siegel and Callanan mentioned Numerosity several times in their vignettes; it was mentioned only once in mine. Siegel and Callanan's participants also read four stories while my participants only read one. Despite Numerosity being a between-subjects variable, Siegel and Callanan's participants might have picked up on the pattern of Numerosity being mentioned in each of the stories and used that information accordingly. Perhaps most importantly, Numerosity was more salient in Siegel and Callanan's vignettes than in mine. The stories I used simply included more details – about the object's physical properties and intentional history, and some anecdotal details about the characters in the story. Because fewer variables were

present to pay attention to in their study, Siegel and Callanan's participants may have been more affected by the Numerosity information than my participants. By extension, Numerosity may be a comparatively less important variable in artifact cognition than physical plausibility, functional plausibility, and intentional history. Nevertheless, in the absence of other competing variables, Numerosity may have an effect, and that may be what Siegel and Callanan revealed. It should also be noted that although Siegel and Callanan allowed participants to respond by choosing both categories, they did not suggest that possibility to their participants. Their principal question was "What is it really for?", which could have prompted a dichotomization of choices into "really for one function" and "not really for the other function," although not all of their participants responded in this way. Nevertheless, the nature of their response measure (free verbalization) was more amenable to a dichotomized response type than the one I employed, even if their participants did not *believe* that the object had to be really for one purpose only. Very few of my participants rejected either category, thus resulting in higher mean ratings for both categories and making it more difficult to find main effects or an interaction of Numerosity with Category, Naming Method, or Entitlement. This was further hampered by the failure of Naming Method and Entitlement to have any significant effects.

Another aspect of the results needing explanation is the correlation data. Correlation measures were taken to determine (1) whether participants were rating the items highly because of their typicality, and (2) whether participants were

dichotomizing their responses across Category. I found a marginally significant positive correlation between typicality ratings for the Old Category and experimental ratings for the Old Category, but no significant correlation was found for the New Category. This suggests that people may have been responding to typicality to some extent in Experiment 2, at least with respect to the Old Category. A significant negative correlation between Old and New categories was found in the Really Naming Method, but there was no significant correlation between Old and New categories in the Discourse Naming Method. This may reflect responses being somewhat more dichotomous in the Really Naming Method than in the Discourse Naming Method. However, a glance at Figure 10 reveals that there is more to this story. *Some* people dichotomized their responses, and this was more likely to happen in the Really Naming Method than in the Discourse Naming Method. But most ratings still fell in the midrange of the scale; a rating of 6 for one category was more often associated with 6's and 7's for the other category than 2's. But the presence of some dichotomies in the Really Naming Method was enough to provide an interpretable correlation statistic, and Figure 10 demonstrates this. But it also demonstrates that there is no simple correlation between the Old and New categories in Figure 10. The most accurate description would have to take into consideration the fact that most people were not dichotomizing their ratings; most people rated both categories well above the midpoint of the scale.

In addition, it may be the case that the correlation between the typicality ratings for the New categories and the main experiment ratings for the New categories failed to reach significance because the range of responses was greater in the experiment data than in the typicality data although the range of *possible* responses was identical. In the pretest, participants read a physical description and only considered typicality with respect to one category, without reading a vignette. This tended to elicit very high ratings for both categories. In the experiment, people chose lower ratings more often. This was likely related to the presence of more details in the experimental vignettes and the consideration of two potential categories rather than one. Therefore, the possibility should not be ruled out that the lack of effects in Experiment 2 was due to an effect of high typicality with respect to both categories, not just the Old Category, for the five items used in the analysis.

In conclusion, the results of Experiment 2 clearly demonstrate that strong essentialism and weak essentialism fail to account for the way people think about artifacts. Even when intentional history is made salient to people, which is not always the case in the real world, people do not necessarily use that information to guide their thinking about what an artifact is or how to talk about it. It is more likely that intentional history serves a role as one of many pragmatic factors in considering the best way to think or talk about artifacts, as suggested by radial theory. Although intentional history may be an important factor in artifact

cognition, the results of this study show that it is not as central to artifact cognition as strong essentialism and weak essentialism suggest.

General Discussion

In the experiments reviewed above, I tested predictions generated from three theories pertaining to artifact cognition. Below, I discuss implications my findings have for theories of how people think about and choose names for artifacts. Then I discuss how my findings might be interpreted in light of other evidence in the literature for and against essentialism. Finally, I discuss future directions for theoretical and empirical research into artifact cognition.

Is Essentialism for Artifacts Real?

The short answer to the question of whether psychological essentialism is real for artifacts is "no." For it to be real, it should be a pervasive phenomenon driving people's decisions about what an artifact really is. The goal of the experiments presented here was to investigate whether people possess an underlying essentialism for artifacts that manifests itself separately from their beliefs about how to talk about artifacts. Experiment 1 found reliable differences between how people were willing to talk about artifacts and what they thought they *really* were, but these differences were not of the sort predicted by essentialist theories. Strong and weak essentialist theories only differ in one respect: Weak essentialism allows for an essence to change. But neither essentialist theory can allow for an artifact to possess multiple essences, which would amount to admitting that artifacts do not have an essence at all. Experiment 1 presented stories about

artifact owners who decided to use their artifacts for a new purpose from now on. When thinking about how these characters would talk about these artifacts, participants reasonably rated the names associated with the new categories as more acceptable than those associated with the old categories. But when thinking about what the artifacts really were, participants rated both categories as acceptable, although they preferred the old one. The consistent endorsement of both categories above the midpoint of the scale presents a serious problem for any essentialist interpretation of the results: If people believe that an object's essence is determined by an intentional decision for the object to belong to a certain category, then they should either believe that it is absolutely the old category or absolutely the new category. The results of Experiment 1 do not support this essentialist prediction.

Furthermore, in Experiment 2, there were no coherent differences between what people were willing to call artifacts and what they thought the artifacts *really* were. Experiment 2's stories did not require participants to think about how a fictional character would talk or think about the object; participants had only to think for themselves. The strong preferences found in Experiment 1 for the new category in Discourse ratings and for the old category in Really ratings were reversed in Experiment 2: Participants rated the old categories slightly higher than the new categories in the Discourse ratings and the new categories slightly higher than the old categories in the Really ratings. This reversal may have reflected the differences between each experiment's Discourse Naming Method. In Experiment

1, participants rated both the old and new categories as potential ways for the object's owner to refer to the object in the story to another character who was aware of the object's intentional history. Considering that the object's owner had expressed a desire to use the object in its new use permanently, participants may have endorsed the new category more than the old category. In Experiment 2's Discourse Naming Method, participants were asked how sensible it would be to refer to the object using the old category name or the new category name. In endorsing the new category names in the Discourse Naming Method but the old category names in the Really Naming Method, participants may have been indicating a belief that it's more sensible to say that the object really belongs to the new category now, even if successful reference would be more likely with the old category name. In other words, although the object has really become the new category, it makes more sense to refer to it with the old category name due to its historical association with that name. This trend may have obtained in Experiment 1 had it not been for the pressure to take the character's perspective in deciding how to refer to the object.

Regardless of their etiology, these results are not consistent with strong essentialist predictions, in which the old category should be the *real* category containing the true essence. Although weak essentialism predicts endorsement of the new category in Really ratings, it cannot account for the simultaneous endorsement of the old category in the same ratings. Only radial theory can

account for participants' beliefs that artifacts can belong *really* to multiple categories.

Additional evidence against essentialism can be found in participants' justifications in Experiment 2. Less than 5% of participants accounted for their ratings with essentialist justifications; the overwhelming majority mentioned physical or functional justifications or a combination of these. Although the artifacts' intentional history was made clear to participants, they rarely reported using this information to guide their thinking about how to talk about the artifacts or what the artifacts really are. The small minority of people who did offer essentialist justifications made correspondingly essentialist ratings for the artifacts. It is therefore possible and likely that essentialists are real, but they represent only a small minority of the general population. The fact that their intuitions about artifacts are at odds with 95% of their peers suggests that they may be treating artifacts the way most people treat natural kinds – with a belief that they possess some kind of essence. Some psychologists have suggested that a failure to treat artifacts and natural kinds differently is a predictor of prejudice against stigmatized social categories (Rothbart & Taylor, 1992) and may reflect an underlying cognitive conservatism driven by a need for structure (Jost, Glaser, Kruglanski, & Sulloway, 2003). The few participants who behaved like essentialists may have been true essentialists. But their minority status precludes the seriousness of any possibility that essentialism for artifacts is universal. While

essentialists might be real, essentialism is not real in the sense that it does not drive artifact cognition for the overwhelming majority of people.

What About Other Evidence?

The introduction section of this thesis presented evidence from the literature for and against essentialism. The experiments presented here sought to disentangle artifact naming from artifact cognition by asking people several nuanced questions about artifacts that might belong to two categories: how they might talk about the artifact, how the object's owner might talk about it, what it is, what it *really* is, and what the artifact's owner thinks it *really* is. Participants were always given the option to endorse one category, both categories, or neither category. In addition, Experiment 2 allowed participants to explain their ratings. These response measures were employed to detect any potential effects or constraints that the question being asked could have upon people's responses. This was done in an attempt to overcome limitations of previous work which tended to use forced-choice response measures in which participants were led to believe that the experimenters required one, and only one, categorical response for what the artifact is (Kemler Nelson, Herron, & Morris, 2002; Diesendruck, Markson, & Bloom, 2003) and to avoid confusing how people will talk about artifacts with what they think the artifacts *really* are (Bloom, 2007). Furthermore, these response measures avoided any focus on function-based responding, in which participants might merely have been rating how well an artifact could fulfill the function of a particular category (Kemler Nelson, Herron, & Morris, 2002; Siegel

& Callanan, 2007). Finally, the instructions to participants in these studies did not include any demand characteristics that suggested they should dichotomize their responses, rating one category high and the other low.

In light of these considerations, some of the work that has been presented as supporting essentialism is called into question. Bloom and Markson (1998) found that children consider the category intended by an artist when choosing names for drawings. Diesendruck, Markson, and Bloom (2003) found that children rely on the category intended by the creator when choosing names for novel artifacts. Both of these studies presented children with unusual stimuli that had little information associated with them other than intended category membership. The physical properties of these stimuli were vague (Bloom & Markson, 1998) or bizarre (Diesendruck, Markson, & Bloom, 2003). Thus, in the paucity of other factors, children used intentional information to guide their naming decisions about these artifacts. But no information was collected on what the children believed the artifacts *really* were. Furthermore, these scenarios do not accurately reflect the processes in which children become familiar with artifacts in everyday life, in which there are pragmatic goals affecting how they might name and think about an artifact's features and functions. The first study used explicitly representational artistic material as stimuli, which necessarily abstract from the objects represented, and how children think about these representations should not be conflated with predictions about how children might think about the *objects* being represented. The second study used novel objects that bordered on the

representational (e.g. aluminum foil shaped like a three-leaf clover, a triangular shallow container made of clay), and is therefore open to the same criticism. I am aware of no study that has provided convincing evidence that children or adults consistently use intentional information to endorse one artifact category while simultaneously rejecting another artifact category because of its lack of intentional justification.

Kemler Nelson, Herron, and Morris (2002) provided evidence for weak essentialism: Their participants respected the new intended function when naming artifacts that appeared intentionally dysfunctional. In other words, their participants seemed to express an understanding that someone had intentionally altered the artifacts so as to make them belong to a new category. While these results are suggestive of weak essentialism, the study's design failed to demonstrate this fully. To be considered evidence for weak essentialism, it would have to be shown that people are consistently rejecting the old category and endorsing the new one for strictly intentional reasons. Their study did not achieve this demonstration; any potential gradation of category membership was not assessed because they did not allow participants to express anything other than an all-or-nothing endorsement of one, and only one, category, and the authors did not probe their participants for justifications of their responses. Their conclusions about a function-based artifact cognition relied on their coding of participants' name generations into discrete functional- and perceptual-based groups.

The evidence against essentialism has traditionally come from tasks allowing participants a greater range of responses, whether in their rating scale options (Kalish, 1995) or in the option of endorsing or rejecting multiple categories (Malt & Sloman, 2007). The experiments presented here employed both of these techniques and also allowed participants to explain their ratings. It is not surprising, therefore, that the results presented here paint a non-essentialist portrait of artifact cognition. When thinking about artifacts, people were affected by factors such as typicality, physical properties, and functionality more than intentional origins. Past evidence for essentialism can likely be summed up as follows: In the absence of other information (familiarity, functionality, physical properties, entrenchment), people may use intentional information to guide their decisions about how to talk about artifacts and what artifacts really are. If this information is collected in a way that precludes multiple responses or explanations of responses, then people's behavior may appear to be guided by an underlying essentialism. But when people are given richer information about artifacts and allowed to express a range of ideas about them, they do not behave as if they ascribe to any kind of essentialism. Because people are not constrained in how to talk or think about artifacts in everyday life, it follows that the designs used in these experiments more accurately reflect the way people really think about artifacts than experiments in which people are forced to choose one name for an artifact without further elaboration. The evidence presented here – that people are guided by factors other than intentionality when thinking about

artifacts and that they are willing to believe an artifact can belong *really* to multiple categories – suggests that previous evidence presented in support of essentialism failed to tell the whole story. There is no universal underlying essentialism for artifacts.

Future Directions

The difficulty of assessing artifact cognition without the influence of language is often cited as a problem in artifact research (Bloom, 1996, 2007). Although object-sorting experiments and studies using novel objects have sought to overcome this obstacle, it is never possible to conclude that any study obtained results that were free and clear of linguistic influence because language and thought are deeply intertwined. Even the sorting of novel objects may be guided by internal monologues that include linguistic mappings and descriptions of the objects and their features. In the studies presented here, I have taken the position that the connection of language to artifact cognition should be accepted rather than weeded out. I cannot make any conclusions about what my participants were thinking non-linguistically. But I *will* argue that the ways linguistic responses are elicited can determine the effects on how people think about their responses. For instance, in Experiment 1 there were reliable differences between how people thought an artifact's owner would talk about it and what they (the participants) believed the artifact *really* was. In Experiment 2, participants only considered how they would talk about the object (rather than answering for the artifact's owner). The differences between how people were willing to talk about the

artifacts and what they thought the artifacts *really* were disappeared. Future work in this domain should delve deeper into the differing effects of linguistic prompts on how people think about their linguistic responses (or any dependent measure) regarding artifacts, without making assumptions about non-linguistic cognition.

It also remains to be seen whether people hold similar beliefs about different types of artifacts, such as social categories, currency, written material, and art. The work presented here established that people do not subscribe to essentialism for somewhat mundane household artifacts such as tablecloths and end tables. But their beliefs about more complex artifacts may take on a different structure (Kalish, 1995). I have not concluded that essentialism is universally rejected for artifacts, but rather that it is not universally accepted for them. It is certainly possible that people may believe that certain types of artifacts have essences, much in the way that people appear to believe in essences for some (but not all) natural kinds.

Perhaps more broadly, investigation of the factors underlying artifact cognition is relevant because of its potential for revealing how people think about the world. If social prejudice can be traced to a person's mistaking of artifacts (social categories) for natural kinds (Keller, 2005), then understanding why people make these mistakes is of the utmost societal and international importance, with possible implications for how children should be taught to conceptualize things in the world as either artifacts or natural kinds. When people essentialize social categories, they tend to do so along several dimensions including immutability

(the degree to which a person can change category), discreteness (the degree to which categories are believed to have rigid boundaries), and naturalness (the degree to which category membership is perceived to be biologically determined) (Haslam, Rothschild, & Ernst, 2000, 2002). People who believe social categories have essences (immutable and natural characteristics creating discrete categories) are more likely, in general, to be prejudiced against certain social categories than people with non-essentialist beliefs (Bastian & Haslam, 2006). Although some aspects of social categories do seem to be somewhat immutable and natural (e.g. skin color), getting people to understand that these characteristics do not determine inherent character of category members should reduce prejudice (Gill & Andreychik, 2007). In this respect, social categories are artifacts in the sense that they are distinctions made by people, not given naturally by an inherent character or essence in their group members.

Evidence that people do not essentialize artifacts is hopeful: People should be less likely to essentialize, stigmatize, and discriminate against social categories if they learn to view those categories as artifacts rather than natural kinds. This may be achievable by teaching people to see the commonalities between complex artifacts such as social groups and simple artifacts such as tablecloths and end tables. For instance, school instructors teaching about ethnic groups could make clear how what it means to belong to an ethnic group is determined not by genetic history but by broad social, historical, political, and interpersonal factors for which family lineage cannot adequately account. By continuing to investigate

how people think about simple artifacts, a common structural isomorphism may be sought in which people have logical cause for viewing social groups as artifacts rather than natural kinds and for learning to think about them in the same non-essentialist ways with which they approach other artifacts.

Table 1

Stimuli Original and New Categories, Experiment 1.

Original Category	New Category
tablecloth	picnic blanket
cookie jar	planter
stool	end table
candy dish	ashtray

Table 2

Sample Story and Response Options for One Stimulus, Experiment 1.

The item shown here was sold as a tablecloth. Marla bought it several years ago and used it in the summer as a table cloth. Last year, Marla redecorated her dining room. She didn't think the tablecloth matched the room's new look very well. Last month, Marla was getting ready to go on a picnic. She needed to take a picnic blanket with her. When she checked in the linen closet, she realized that she needed a new picnic blanket. Marla noticed the old table cloth in the closet. She thought, "This will work. I don't want to use it as a tablecloth anymore." Marla decided that she could use it as a picnic blanket from now on. She took it with her and used it as a picnic blanket.

At the picnic, Marla met her friend Gary. Marla told Gary the story of how she bought this tablecloth but decided to start using it as a picnic blanket from now on. When the picnic was over and they were getting ready to leave, Marla said to Gary, "Let's get this stuff into the car; could you please ...

"... hand me the **tablecloth**?"

1 2 3 4 5 6 7

not acceptable somewhat acceptable very acceptable

"... hand me the **picnic blanket**?"

1 2 3 4 5 6 7

not acceptable somewhat acceptable very acceptable

It's **really** a tablecloth.

1 2 3 4 5 6 7

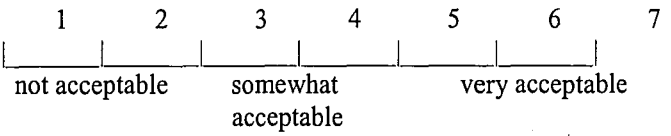
not acceptable somewhat acceptable very acceptable

It's **really** a picnic blanket.

1 2 3 4 5 6 7

not acceptable somewhat acceptable very acceptable

Marla thinks it's **really** a tablecloth.



Marla thinks it's **really** a picnic blanket.

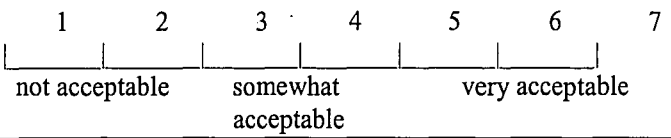


Table 3

Mean Ratings of Categories in Discourse Naming Method as a Function of Category and Typicality, Experiment 1.

	Typicality to New Category	
Category	Typical	Atypical
Old		
M.	3.75	4.52
S.D.	1.19	0.96
New		
M.	5.96	5.63
S.D.	0.74	0.77

Table 4

Mean Ratings of Categories in "Really" Naming Method as a Function of Category and Typicality, Experiment 1.

	Typicality to New Category	
Category	Typical	Atypical
Old		
M.	5.46	6.21
S.D.	0.91	0.61
New		
M.	3.94	3.79
S.D.	1.10	0.99

Table 5

Mean Ratings of Categories in "Character Really" Naming Method as a Function of Category and Typicality, Experiment 1.

	Typicality to New Category	
Category	Typical	Atypical
Old		
M.	5.13	5.94
S.D.	0.96	0.65
New		
M.	3.96	3.81
S.D.	1.36	0.82

Table 6

Mean Ratings of Old vs. New Categories as a Function of Naming Method, Experiment 1.

	Category	
Naming Method	Old	New
Discourse	4.14	5.79
Really	5.83	3.87
Character Really	5.53	3.89

Table 7

Sample Story and Response Options for Really and Discourse Naming Methods, Experiment 2.

Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, he decided to start selling them as stools. He has made and sold them as stools ever since. He never plans to make or sell them as end tables again. (One Creator)

It is really a stool.

☐ Absolutely not. 1 2 3 4 5 6 7 ☐ It's really a stool and nothing else.

barely acceptable — somewhat acceptable — very acceptable

It is really an end table.

☐ Absolutely not. 1 2 3 4 5 6 7 ☐ It's really an end table nothing else.

barely acceptable — somewhat acceptable — very acceptable

It is OK to call the object a stool.

☐ Absolutely not. 1 2 3 4 5 6 7 ☐ It's called a stool and nothing else.

barely acceptable — somewhat acceptable — very acceptable

It is OK to call the object an end table.

☐ Absolutely not. 1 2 3 4 5 6 7 ☐ It's called an end table and nothing else.

barely acceptable — somewhat acceptable — very acceptable

Table 8

Stimuli (Original and New Categories) and Mean Typicality Ratings, Experiment 2.

Original Category	Mean Typicality	New Category	Mean Typicality
end table	6.67	stool	6.33
candy dish	6.20	ashtray	6.33
tablecloth	6.47	picnic blanket	5.67
teapot	6.33	watering can	5.87
butter knife	6.33	letter opener	6.20
umbrella	6.07	lampshade	5.33

Table 9

Mean ratings of Old vs. New categories in Really Naming Method as a Function of Numerosity and Entitlement, Experiment 2.

NUMEROSITY: ONE		Entitlement	
Category	Creator	Owner	Shopper
Old			
M.	5.33	5.20	5.33
S.D.	1.84	2.21	1.54
New			
M.	5.73	5.60	5.33
S.D.	1.87	1.64	1.54

NUMEROSITY: MANY		Entitlement	
Category	Creator	Owner	Shopper
Old			
M.	5.40	5.40	5.73
S.D.	1.72	2.13	1.75
New			
M.	5.80	6.07	4.60
S.D.	1.47	1.91	2.20

Table 10

Mean ratings of Old vs. New categories in Control Naming Method as a Function of Numerosity and Entitlement, Experiment 2.

NUMEROSITY: ONE		Entitlement	
Category	Creator	Owner	Shopper
Old			
M.	5.20	6.20	5.73
S.D.	1.57	1.15	1.79
New			
M.	4.47	5.80	5.20
S.D.	1.92	1.86	2.08

NUMEROSITY: MANY		Entitlement	
Category	Creator	Owner	Shopper
Old			
M.	6.07	5.53	5.27
S.D.	1.44	1.88	2.28
New			
M.	4.53	6.20	5.00
S.D.	2.00	1.42	1.81

Table 11

Mean ratings of Old vs. New categories in Discourse Naming Method as a Function of Numerosity and Entitlement, Experiment 2.

NUMEROSITY: ONE	Entitlement		
Category	Creator	Owner	Shopper
Old			
M.	5.40	6.20	5.93
S.D.	1.35	1.01	1.16
New			
M.	4.93	5.33	5.27
S.D.	1.91	1.72	1.39

NUMEROSITY: MANY	Entitlement		
Category	Creator	Owner	Shopper
Old			
M.	5.07	6.20	6.33
S.D.	2.05	1.32	1.05
New			
M.	5.53	5.00	5.27
S.D.	1.55	2.36	1.62

Table 12

Examples of Five Participant Justification Categories, Experiment 2.

Radial	"The object is ambiguous and could be seen as either. Also, one could use the object in both ways."
Physical	"Letter openers do not have handles and are not slightly sharp."
Functional	"It can be used as both."
Essentialist	"If something is intended to be used as a butter knife and people are misusing it, it is the people's mistake."
Other	"My mom has a tablecloth like this that we used to paint on."

Table 13

Number and Percentage of Participant Justifications by Naming Method, Experiment 2.

	Radial	Physical	Functional	Essentialist	Other
Really					
Total #	50	30	14	4	10
Percentage	46	28	13	4	9
Discourse					
Total #	53	21	20	3	11
Percentage	49	19	19	3	10

Table 14

Mean Ratings of Old vs. New Categories in Really Naming Method as a Function of Justification Type, Experiment 2.

	Category	
Justification Type	Old	New
Radial (n = 50)	5.60	5.46
Radial, no essentialism (n = 29)	5.34	5.90
Radial, with essentialism (n = 21)	5.95	4.86
Physical	4.87	5.10
Functional	6.14	6.07
Essentialist	4.00	4.25
Other	5.90	5.10

Table 15

Mean Ratings of Old vs. New Categories in Discourse Naming Method as a Function of Justification Type, Experiment 2.

	Category	
Justification Type	Old	New
Radial (n = 50)	5.92	5.23
Radial, no essentialism (n = 29)	5.84	5.50
Radial, with essentialism (n = 21)	6.05	4.81
Physical	5.09	4.90
Functional	5.95	6.00
Essentialist	6.33	3.33
Other	5.18	5.18

Figure Captions

Figure 1. Typical (a) and atypical (b) picnic blankets, Experiment 1.

Figure 2. Predicted Experiment 1 trends for strong essentialism in (a) Discourse, (b) Really, and (c) Character Really naming methods.

Figure 3. Predicted Experiment 1 trends for weak essentialism in (a) Discourse, (b) Really, and (c) Character Really naming methods.

Figure 4. Predicted Experiment 1 trends for radial theory in (a) Discourse, (b) Really, and (c) Character Really naming methods.

Figure 5. Mean ratings of Old vs. New categories as a function of Naming Method, Experiment 1.

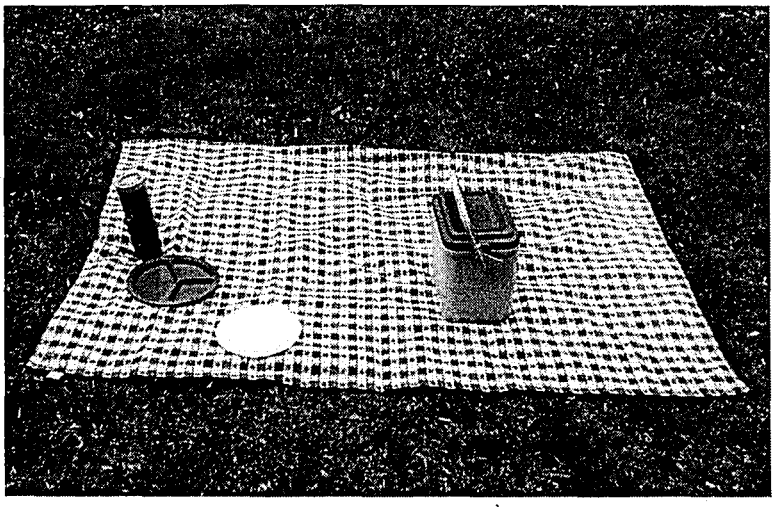
Figure 6. Predicted Experiment 2 trends for all theories in Discourse Naming Method for (a) old and (b) new categories.

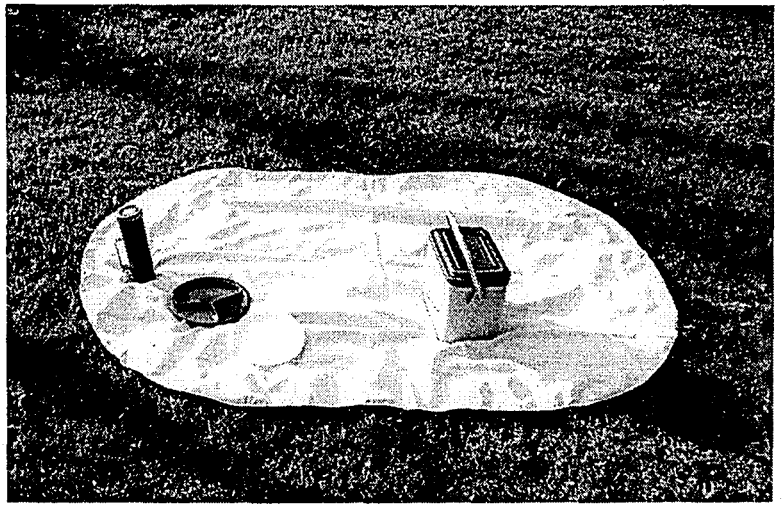
Figure 7. Predicted Experiment 2 trends for strong essentialism in Really Naming Method for (a) old and (b) new categories.

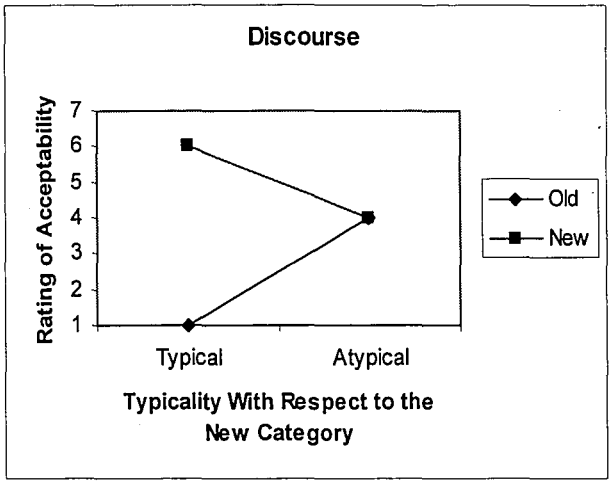
Figure 8. Predicted Experiment 2 trends for weak essentialism in Really Naming Method for (a) old and (b) new categories.

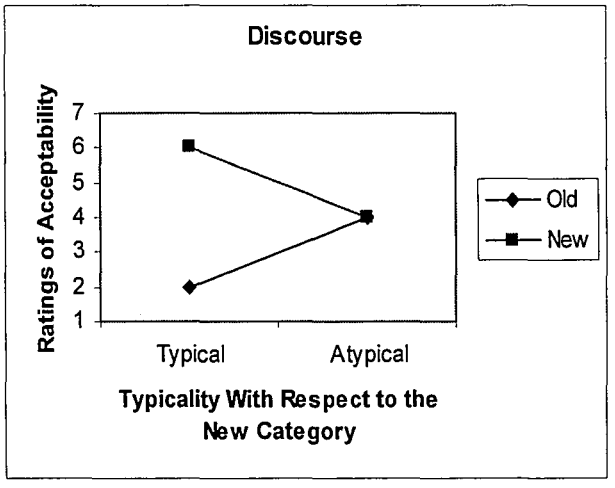
Figure 9. Scatterplot demonstrating lack of correlation of old and new categories in Discourse Naming Method, Experiment 2.

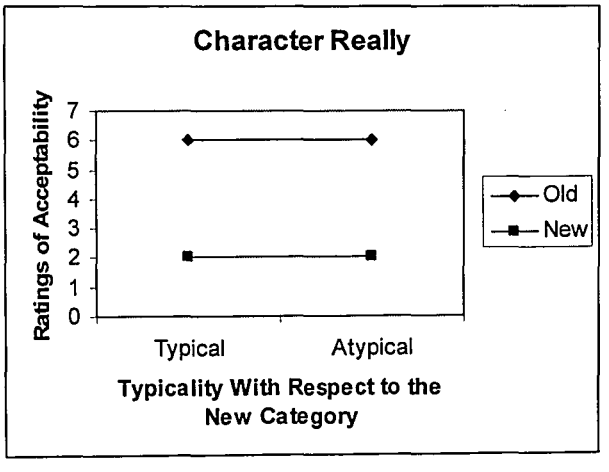
Figure 10. Scatterplot demonstrating negative correlation of old and new categories in Really Naming Method, Experiment 2.

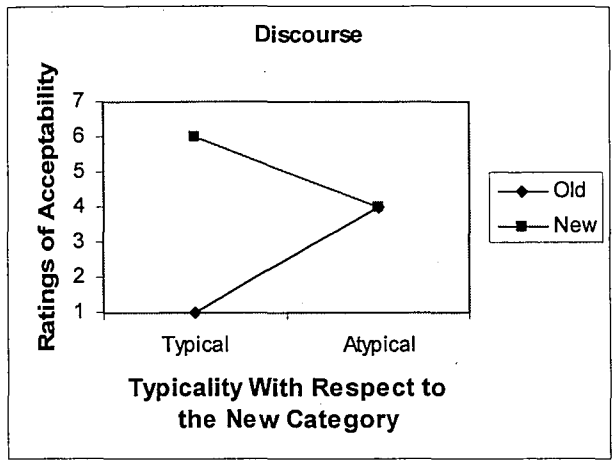


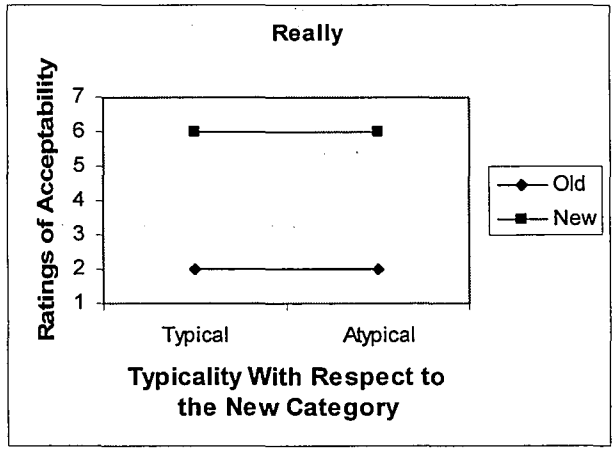


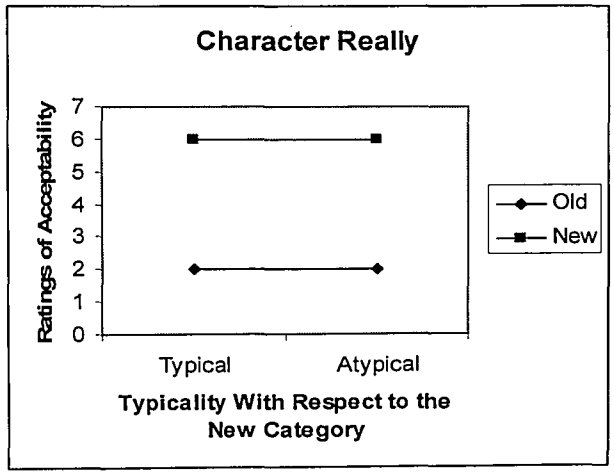

Discourse
Rating of Acceptability
7
6
5
4
3
2
1
Typical
Atypical
Old
New
Typicality With Respect to the New Category

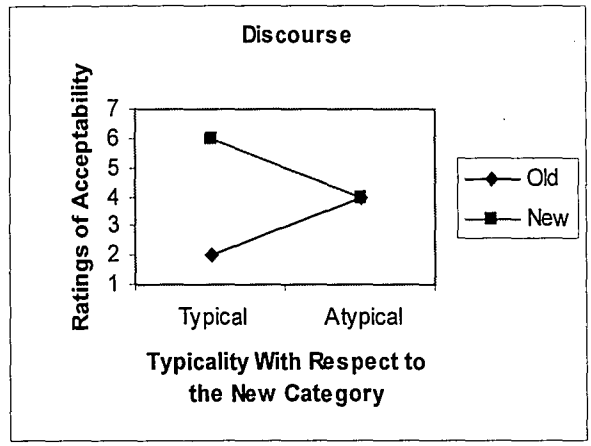

Discourse
Ratings of Acceptability
7
6
5
4
3
2
1
Typical
Atypical
Old
New
Typicality With Respect to the New Category

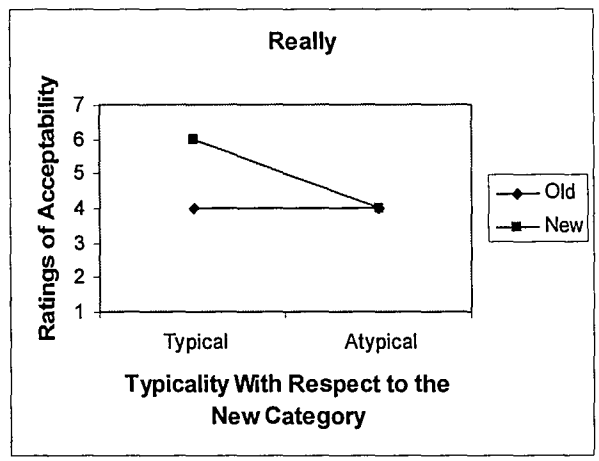

Character Really
Ratings of Acceptability
7
6
5
4
3
2
1
0
Typical
Atypical
Old
New
Typicality With Respect to the New Category

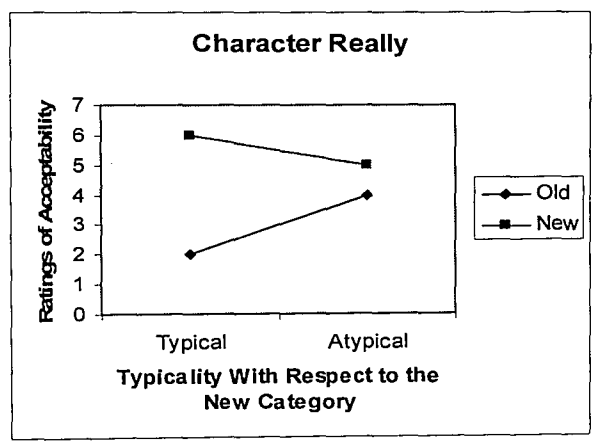

Discourse
Ratings of Acceptability
7
6
5
4
3
2
1
Typical
Atypical
Typicality With Respect to the New Category
Old
New

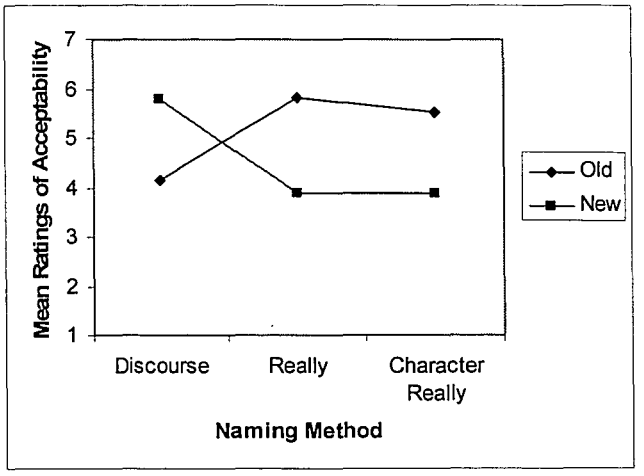

Really
Ratings of Acceptability
7
6
5
4
3
2
1
Typical
Atypical
Typicality With Respect to the New Category
Old
New

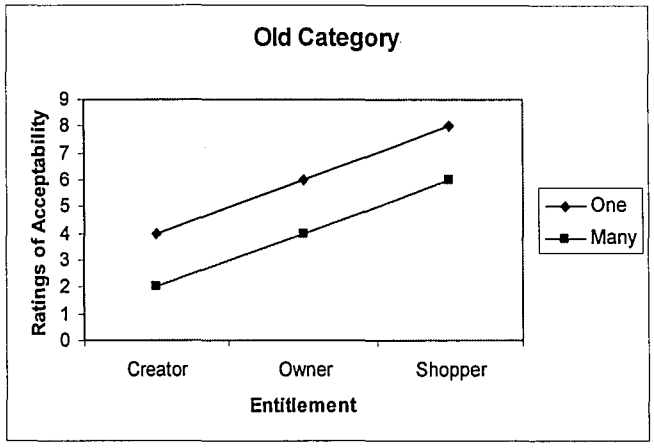

Character Really
Ratings of Acceptability
7
6
5
4
3
2
1
0
Typical
Atypical
Typicality With Respect to the New Category
Old
New

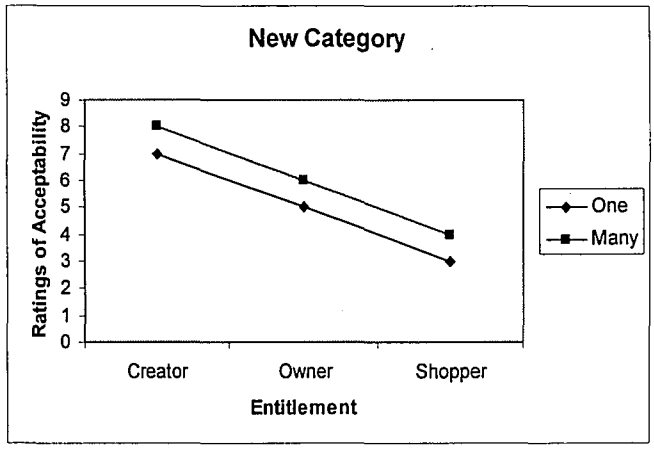

Discourse
Ratings of Acceptability
7
6
5
4
3
2
1
Typical
Atypical
Typicality With Respect to the New Category
Old
New

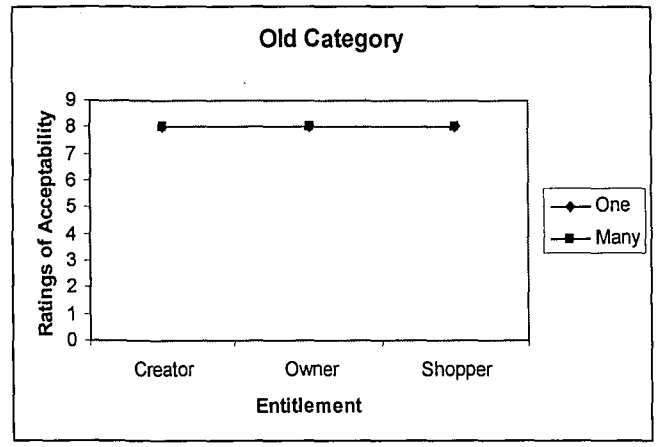

Really
Ratings of Acceptability
7
6
5
4
3
2
1
Typical
Atypical
Typicality With Respect to the New Category
Old
New

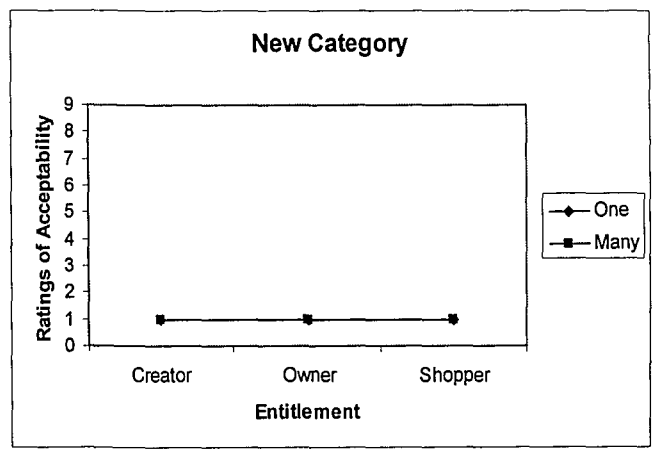

Character Really
Ratings of Acceptability
7
6
5
4
3
2
1
0
Typical
Atypical
Typicality With Respect to the New Category
Old
New

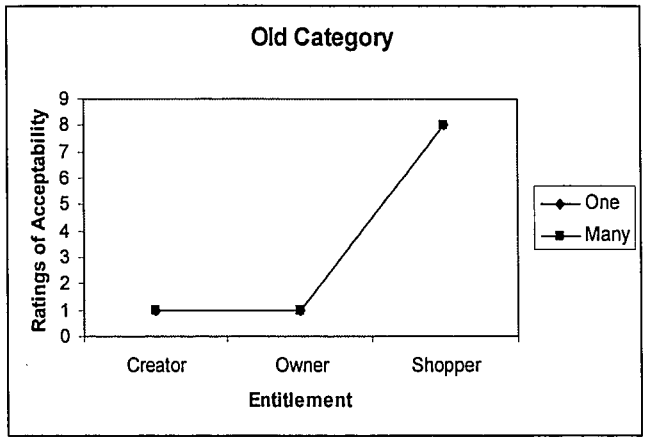

Mean Ratings of Acceptability
7
6
5
4
3
2
1
Discourse
Really
Character Really
Naming Method
Old
New

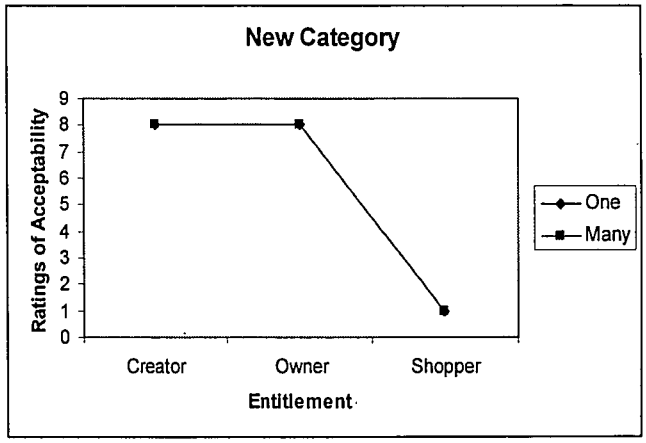

Old Category
Ratings of Acceptability
9
8
7
6
5
4
3
2
1
0
Creator
Owner
Shopper
Entitlement
One
Many

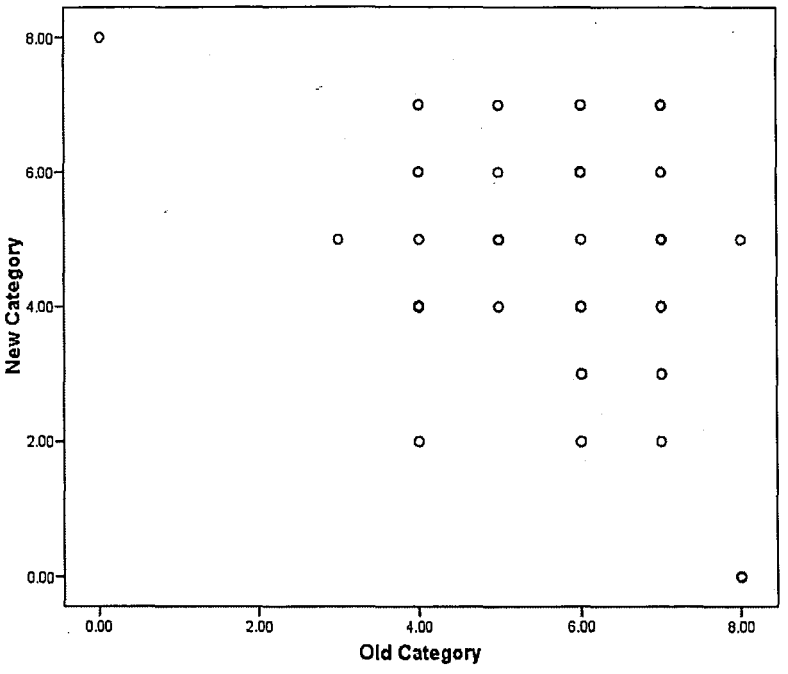

New Category
Ratings of Acceptability
9
8
7
6
5
4
3
2
1
0
Creator
Owner
Shopper
Entitlement
One
Many

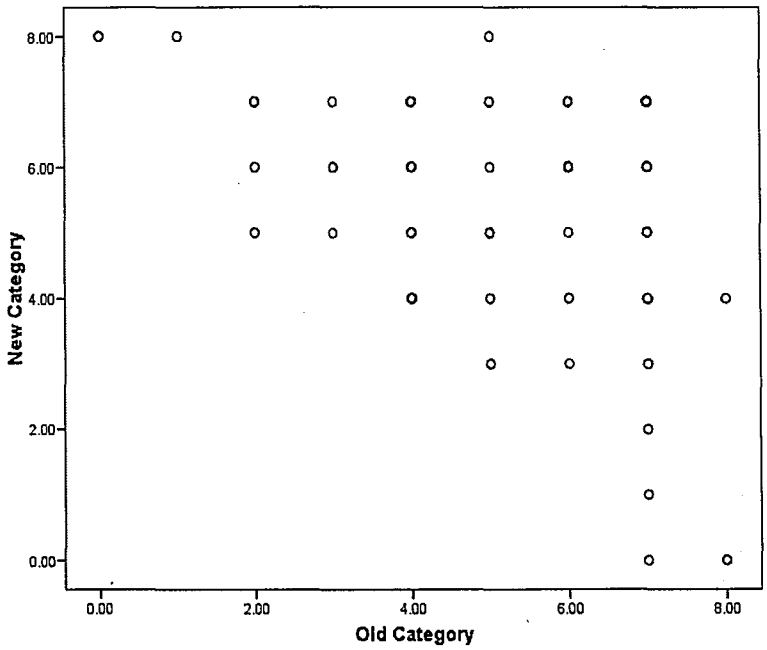

Old Category
Ratings of Acceptability
9
8
7
6
5
4
3
2
1
0
Creator
Owner
Shopper
Entitlement
One
Many


New Category
Ratings of Acceptability
9
8
7
6
5
4
3
2
1
0
Creator
Owner
Shopper
Entitlement
One
Many


Old Category
Ratings of Acceptability
9
8
7
6
5
4
3
2
1
0
Creator
Owner
Shopper
Entitlement
One
Many


New Category
Ratings of Acceptability
9
8
7
6
5
4
3
2
1
0
Creator
Owner
Shopper
Entitlement
One
Many


8.00
6.00
4.00
2.00
0.00
New Category
0.00
2.00
4.00
6.00
8.00
Old Category


8.00
6.00
4.00
2.00
0.00
New Category
0.00
2.00
4.00
6.00
8.00
Old Category

References

Bastian, B., & Haslam. N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology, 42*, 228-235.

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

Bloom, P. (1998). Theories of artifact categorization. *Cognition, 66*, 87-93.

Bloom, P. (2007). More than words: A reply to Malt and Sloman. *Cognition, 105,* 649-655.

Bloom, P., & Markson, L. (1998). Intention and analogy in children's naming of pictorial representations. *Psychological Science, 9*, 200-204.

Braisby, N. (2001). Deference in categorization: Evidence for essentialism? In Moore, J. D., & Stenning, K. (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum.

Braisby, N., Franks, B., & Hampton, J. (1996). Essentialism, word use, and concepts. *Cognition, 59*, 247-274.

Dennett, D. (1987). *The intentional stance*. Cambridge, MA: MIT Press.

Dennett, D. (1990). The interpretation of texts, people and other artifacts. *Philosophy and Phenomenological Research, 50,* 177-194.

Diesendruck, G., Markson, L., & Bloom, P. (2003). Children's reliance on creator's intent in extending names for artifacts. *Psychological Science, 14*, 164-168.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Estes, Z. (2004). Confidence and gradedness in semantic categorization: Definitely somewhat artifactual, maybe absolutely natural. *Psychonomic Bulletin and Review, 11*, 1041-1047.

Gelman, S. A., & Coley, J. D. (1991). Language and categorization: The acquisition of natural kind terms. In Gelman, S., & Byrnes, J. (Eds.), *Perspectives on language and thought: Interrelations in development.* Cambridge: Cambridge University Press.

Gill, M. J., & Andreychik, M. R. (2007). *Belief in human kinds versus inherent character: Distinct types of essentialism and their relation to prejudice.* Unpublished manuscript. Lehigh University: Bethlehem, PA.

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology, 39*, 113-127.

Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology, 41*, 87-100.

Hampton, J. A., Dubois, D., & Yeh, W. (2006). Effects of classification context on categorization in natural categories. *Memory & Cognition, 34*, 1431-1443.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129*, 339-375.

Kalish, C. W. (1995). Essentialism and graded membership in animal and artifact categories. *Memory & Cognition, 23*, 335-353.

Kalish, C. W. (2002). Essentialist to some degree: Beliefs about the structure of natural kind categories. *Memory & Cognition, 30*, 340-352.

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

Keleman, D. (1999). Function, goals, and intention: Children's teleological reasoning about objects. *Trends in Cognitive Science, 3*, 461-468.

Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology, 88*, 686-702.

Kemler Nelson, D. G., Herron, L., & Morris, C. (2002). How children and adults name broken objects: Inferences and reasoning about design intentions in the categorization of artifacts. *Journal of Cognition and Development, 3*, 301-332.

Kripke, S. (1980). *Naming and necessity.* Cambridge, MA: Harvard University Press.

Lakoff, G. (1987). *Women, fire, and dangerous things.* Chicago, IL: University of Chicago Press.

Malt, B. C. (1989). Features and beliefs in the mental representation of categories. *Journal of Memory and Language, 29*, 289-315.

Malt, B. C. (1994). Water is not H_2O. *Cognitive Psychology, 27*, 41-70.

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

Malt, B. C., & Johnson, E. C. (1998). Artifact category membership and the intentional-historical theory. *Cognition, 66*, 79-85.

Malt, B. C., & Sloman, S. A. (2007). Category essence or essentially pragmatic? Creator's intention in naming and what's really what. *Cognition, 105,* 615-648.

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.* Cambridge, MA: MIT Press.

Matan, A., & Carey, S. (2001). Developmental changes within the core of artifact concepts. *Cognition, 78*, 1-26.

Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In Vosniadou, S., & Ortony, A. (Eds.), *Similarity and analogical reasoning*. New York: Cambridge University Press.

Putnam, H. (1977). Meaning and reference. In S. P. Schwartz (Ed.), *Naming, necessity, and natural kinds*. Ithaca, NY: Cornell University Press.

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

Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In Semin, G. R., & Fiedler, K. (Eds.), *Language, interaction, and social cognition* (pp. 11-36). London: Sage.

Shiga, D. (2006). Pluto added to official "minor planet" list. *New Scientist.* Retrieved June 29, 2007, from the World Wide Web: http://space.newscientist.com/ article/dn10028-pluto-added-to-official-minor-planet-list.html.

Siegel, D. R., & Callanan, M. A. (2007). Artifacts as conventional objects. *Journal of Cognition and Development, 8*, 183-203.

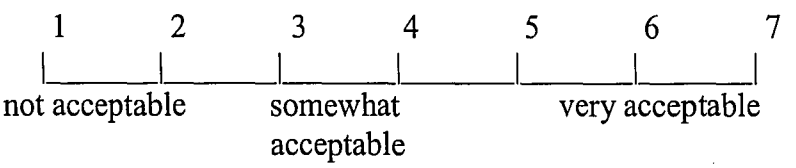
Webster, A. C., & Stewart, R. A. C. (1973). Theological conservatism. In Wilson, G. D. (Ed.), *The psychology of conservatism* (pp. 129-147). London: Academic Press.

Item 1: Please look at Picture #1.

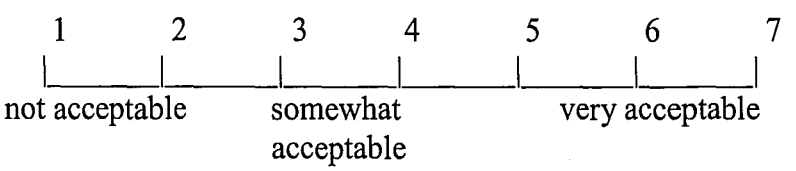
The item shown here was sold as a tablecloth. Marla bought it several years ago and used it in the summer as a table cloth. Last year, Marla redecorated her dining room. She didn't think the tablecloth matched the room's new look very well. Last month, Marla was getting ready to go on a picnic. She needed to take a picnic blanket with her. When she checked in the linen closet, she realized that she needed a new picnic blanket. Marla noticed the old table cloth in the closet. She thought, "This will work. I don't want to use it as a tablecloth anymore." Marla decided that she could use it as a picnic blanket from now on. She took it with her and used it as a picnic blanket.

At the picnic, Marla met her friend Gary. Marla told Gary the story of how she bought this tablecloth but decided to start using it as a picnic blanket from now on. When the picnic was over and they were getting ready to leave, Marla said to Gary, "Let's get this stuff into the car; could you please …

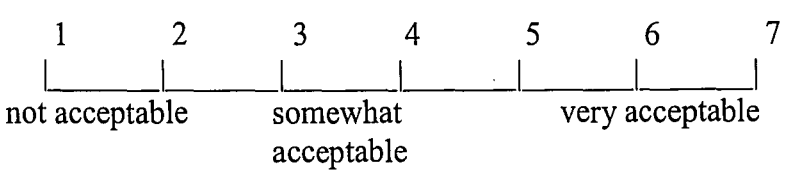
"… hand me the **tablecloth**?"



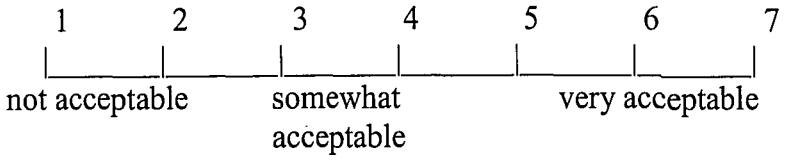
"… hand me the **picnic blanket**?"



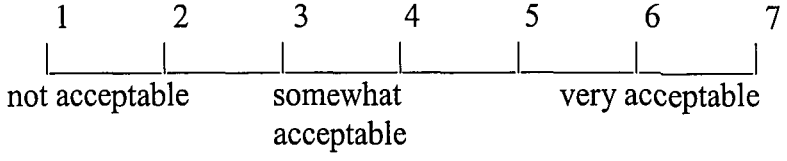
It's **really** a tablecloth.



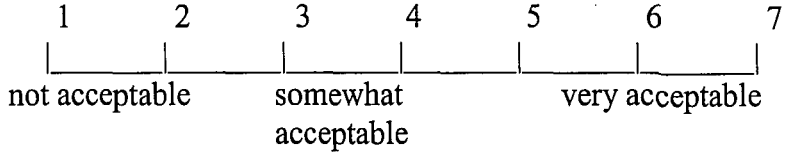
It's **really** a picnic blanket.



Marla thinks it's **really** a tablecloth.



Marla thinks it's **really** a picnic blanket.

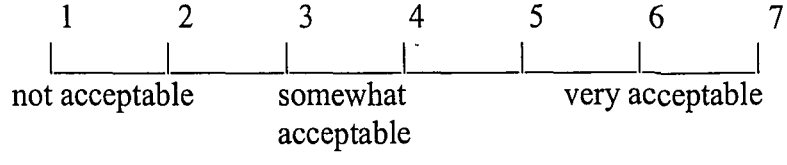


Item 2: Please look at Picture #2.

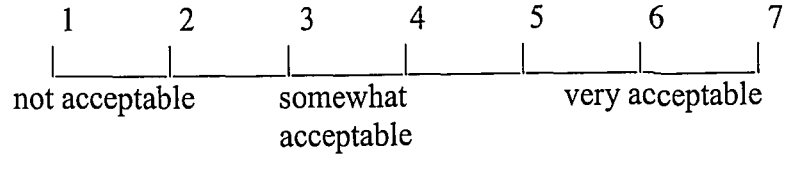
The item shown here was sold as a ceramic cookie jar. John bought it several months ago. Last month he finished the last of the cookies. He bought some bulbs with his friend Linda that day, and they needed a planter in which to place them. He remembered the cookie jar and said to Linda, "This will work. I don't need it for cookies anymore." He filled it with dirt and placed the bulbs inside. He liked it that way. Now it's filled with flowers.

Linda is visiting John again today. Linda asked if how the flowers were doing. John said, ...

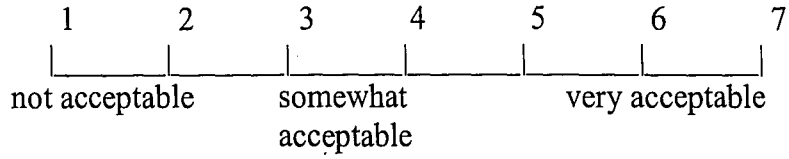
"They're over there in the **cookie jar**."



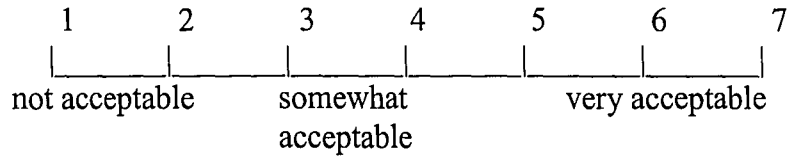
"They're over there in the **planter**."



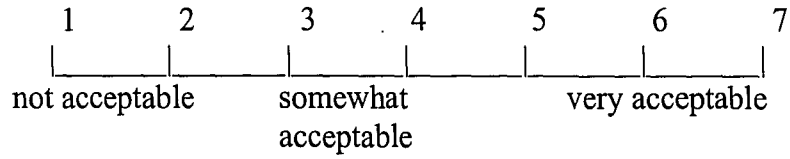
It's **really** a cookie jar.



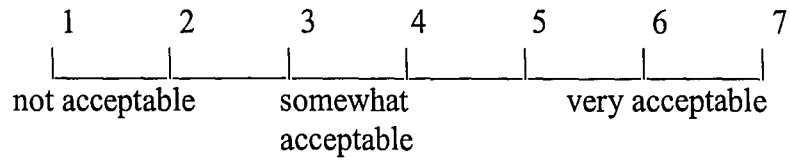
It's **really** a planter.



John thinks it's **really** a cookie jar.



John thinks it's **really** a planter.

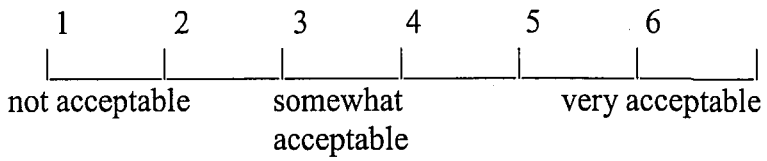


Item 3: Please look at Picture #3.

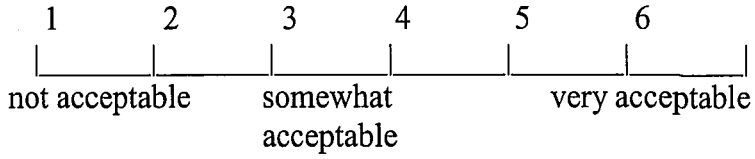
The item shown here was sold as a stool. Rita bought it several years ago. She used it as a stool for her kitchen counter. But when Rita had her kitchen renovated, she bought a better stool that was elevated higher off the ground. Rita needed an end table in her living room, and she remembered the old stool. She thought, "This will work. It's the right height for an end table." Rita placed it in the living room and has been using it as an end table ever since.

Now Rita's friend Frank is visiting her. Rita just finished telling Frank the story of how she started using this object for a new purpose. Frank asked Rita where he could put his drink. Rita said,

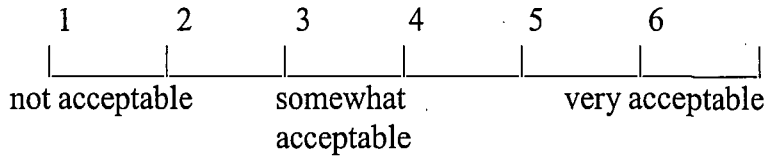
"You can put it on the **stool**."



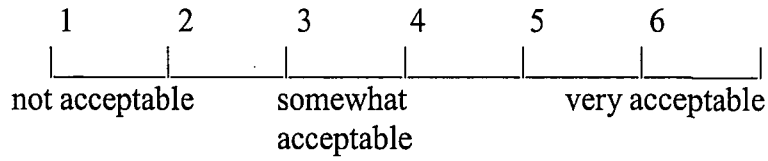
"You can put it on the **end table**."



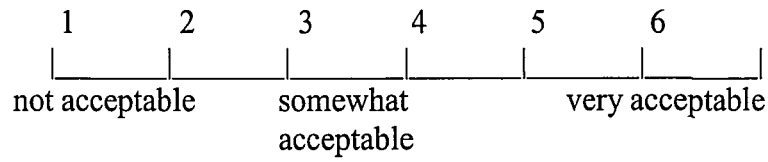
It's **really** a stool.



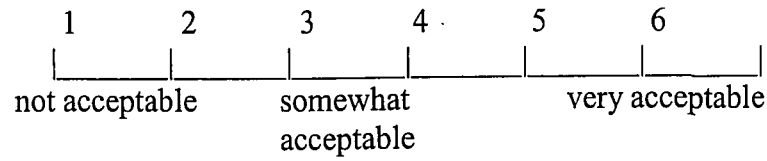
It's **really** an end table.



Rita thinks it's **really** a stool.



Rita thinks it's **really** an end table.

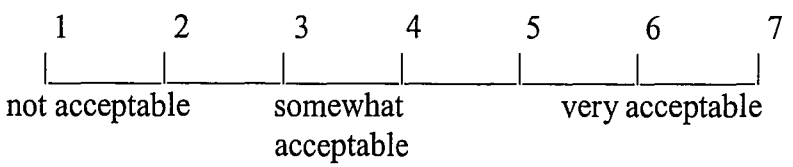


Item 4: Please look at Picture #4.

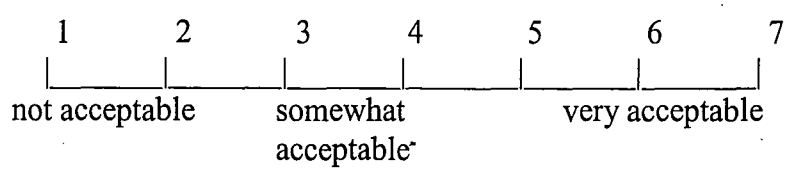
The item shown here was sold as a desk. Jamie bought it several years ago to use for his office laptop. A few months ago, Jamie bought a new desktop computer, and the desk he had was not big enough to hold it. So he bought a new desk to work at. But he wanted a cart to hold his new printer, so that he could share it with his friend Esther who worked next door. He remembered the old desk. Jamie thought, "This will work. It has wheels and is small enough to be a cart." He told Esther the idea, how he had been using the object as a desk but wanted to use it from now on as a cart. Esther agreed with Jamie. Jamie has been using it as a cart ever since.

Today, Esther came into Jamie's office to ask where he kept the extra toner for the printer. Jamie said,

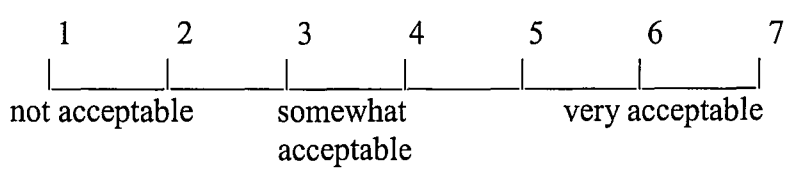
"I put it over on the **desk**."



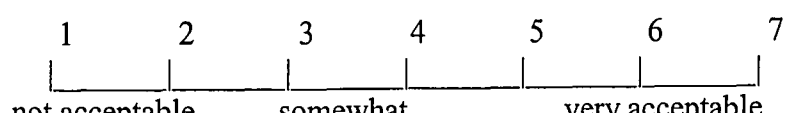
"I put it over on the **cart**."



It's **really** a desk.

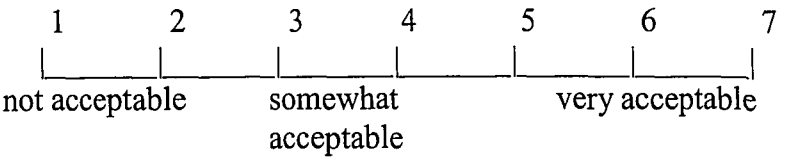


It's **really** a cart.

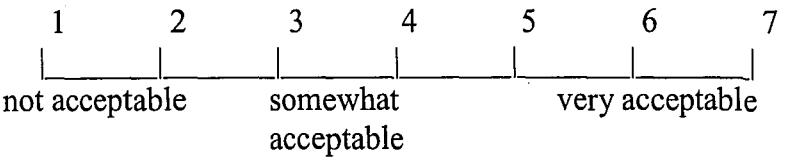


acceptable

Jamie thinks it's **really** a desk.



Jamie thinks it's **really** a cart.

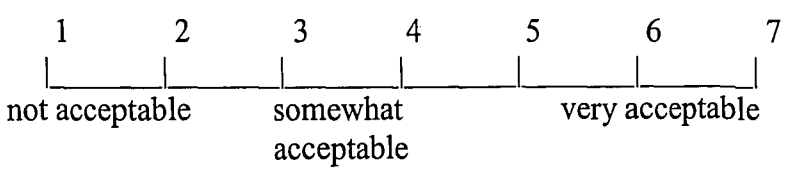


Item 5: Please look at Picture #5.

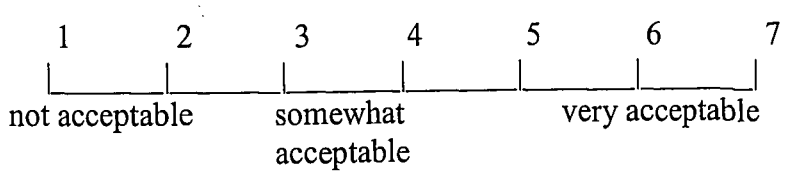
The item shown here was sold as a candy dish. Wilbur bought it several years ago to keep in his living room and hold candy. A few months ago, Wilbur's roommate Julie returned from Russia and told Wilbur that she started smoking in Russia. When Julie lit up a cigarette, she asked if Wilbur had an ashtray. Wilbur remembered the candy dish and thought, "This will work. I shouldn't eat candy anyway." He gave it to Julie to use as an ashtray. He said, "Remember we used this to hold candy? You can use it for your ashes from now on." Julie has used it as an ashtray ever since.

Today Wilbur and Julie were having lunch. Julie lit up a cigarette. Wilbur noticed that she didn't have anything to place her ashes in, so he said,

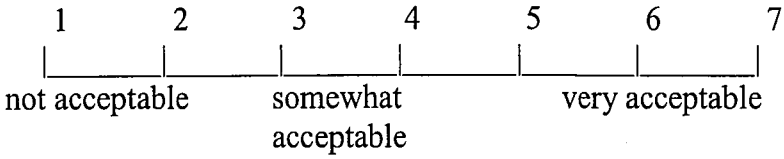
"Hold on, let me get you the **candy dish**."



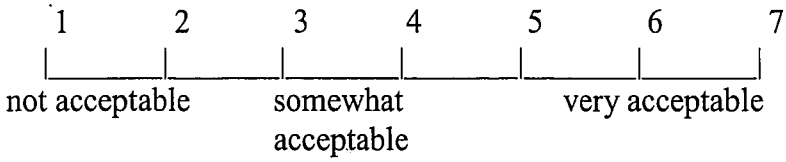
"Hold on, let me get you the **ashtray**."



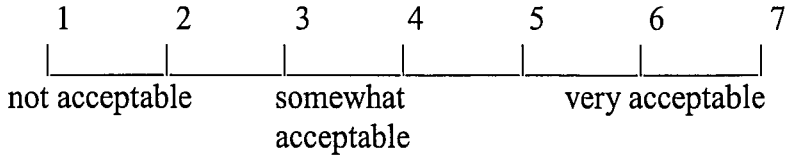
It's **really** a candy dish.



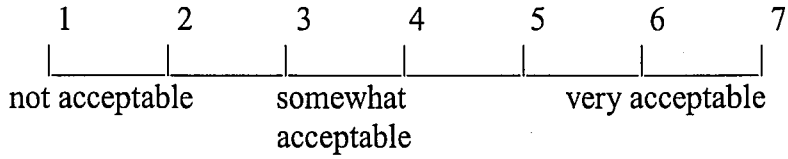
It's **really** an ashtray.



Wilbur thinks it's **really** a candy dish.



Wilbur thinks it's **really** an ashtray.

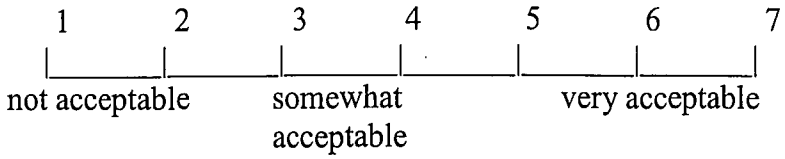


Item 6: Please look at Picture #6.

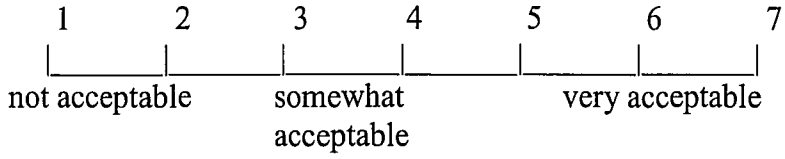
The item shown here was sold as a beer stein. Craig bought it several years ago at a brewery. He used it to drink beer for a few years. Then a few months ago, Craig broke the pencil holder on his desk. He thought of buying a new one, but then he remembered the beer stein. He said to his wife, Maude, "This will work. It can hold as many pencils as my old pencil holder. I have plenty of other beer steins, so it won't be a problem for me to use this one to hold pencils from now on." He put it on his desk and has used it as a pencil holder ever since.

Today, Maude came home from work. She said to Craig, "Look, I got a bunch of free pencils today at the office. Where should I put them?" Craig said,

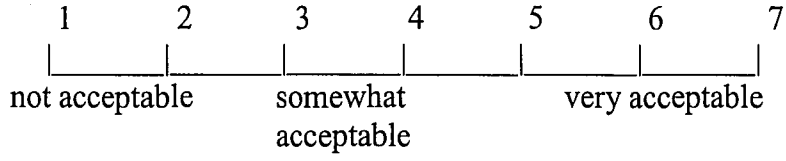
"Put them in the **beer stein**."



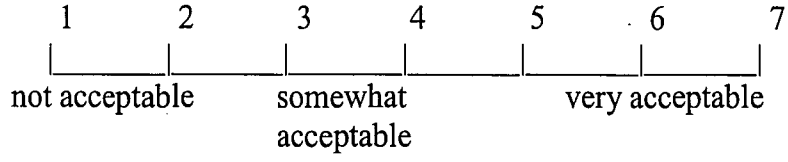
"Put them in the **pencil holder**."



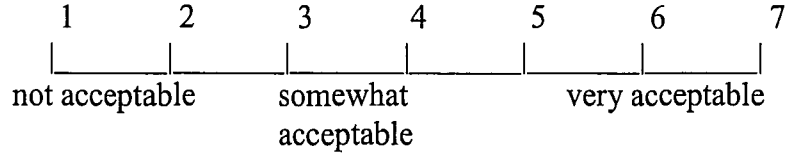
It's **really** a beer stein.



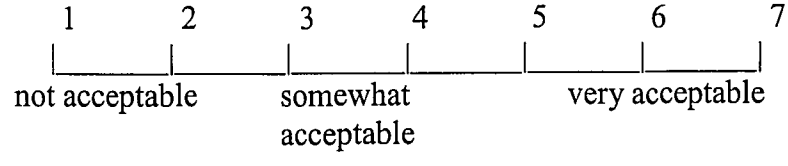
It's **really** a pencil holder.



Craig thinks it's **really** a beer stein.



Craig thinks it's **really** a pencil holder.



Appendix B

Item 1. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

Item 2. Picture a circular object made of smooth glass, 5 inches in diameter, with a flat bottom and a raised, 1.5-inch edge around the circumference. There are small grooves around the edge.

Item 3. Picture a flat, rectangular object made of foldable cloth, with a thin plastic veneer on one side. It is 7 feet long and 5 feet wide, with a red-and-white checkered pattern on one side.

Item 4. Picture an object made of metal, about 10 inches in height, shaped like a dome with a round base and able to hold 1.5 quarts of liquid. It has a handle, spout, and removable lid. The base of the container is 8 inches in diameter.

Item 5. Picture a thin silver object 8 inches long, smooth and sleek. It has a handle at one end and an edge at the other which is slightly sharp but not dangerous. The handle is thicker than the edge. It is no more than an eighth of an inch thick.

Item 6. Picture a conical, canopy-like object made of dull-colored cloth, with a slick plastic finish, 2.5 feet in diameter at its base. It is attached to a 2-foot-long pole by means of a thin metal frame.

INSTRUCTIONS (Discourse)

On the other side of this paper is a short story about a humanmade object. After you read the story, you'll be asked to rate two statements about the object for their acceptability. The statements will be regarding ways of talking about the object.

For instance, suppose the first statement is "It is OK to call the object a hammer." If you think calling the object a hammer is absolutely wrong, and there is no reasonable cause for calling it a hammer, then mark the box for "*Absolutely not.*" If you think it could reasonably be called a hammer, then mark one of the numbers on the scale. A "1" means that you think it's barely acceptable to call it a hammer, and a "4" means that it's somewhat acceptable. The other end of the scale, a "7," indicates that you think calling the object a hammer is very acceptable. If you think calling the object a hammer is *absolutely correct and nothing else will do*, then check the box at the far right end of the scale.

Then look at the next statement for that item. It will ask you to think about a second name that the item might be called, using the same scale. Please mark the scale accordingly.

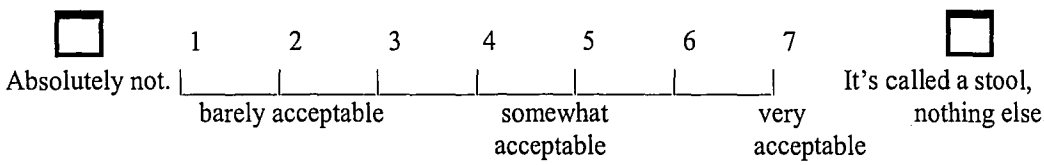
In making your two judgments, you might think that an item should be called by only one of the two names, or you might think that it can be called by both. You might think it should be called neither. The rating scales have been designed to let you express all of these opinions and more.

After you've made your mark on each scale, write a brief explanation under the scales about why you responded as you did for both statements together.

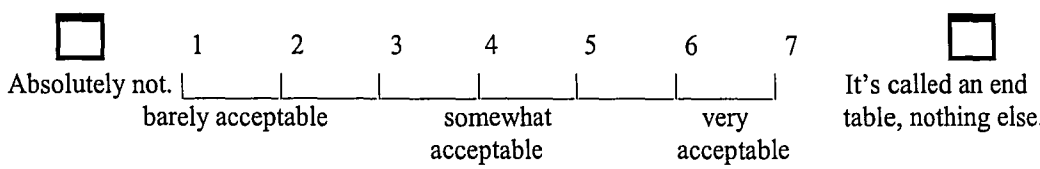
1. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, he decided to start selling them as stools. He has made and sold them as stools ever since. He never plans to make or sell them as end tables again. (One Creator)

It is OK to call the object a stool.



It is OK to call the object an end table.

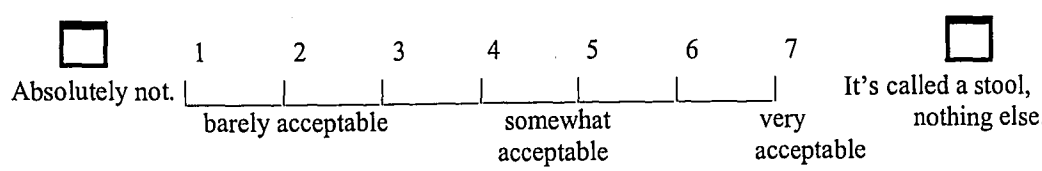


Please explain your choices in a few sentences.

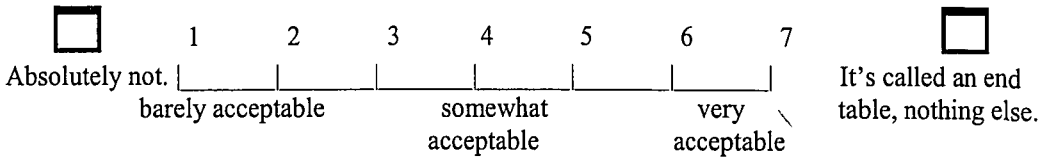
2. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, he sold one at his shop to Marla, but Marla didn't buy it to use as an end table. Although she knew it was intended to be used as an end table, Marla bought it to use as a stool and has been using it as a stool ever since. (One Owner)

It is OK to call the object a stool.



It is OK to call the object an end table.

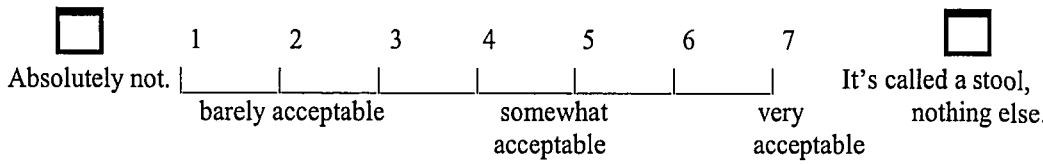


Please explain your choices in a few sentences.

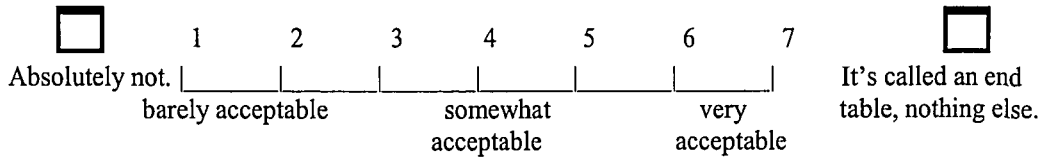
3. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, Marla was browsing in the shop and noticed one. She thought, "If I could afford that, I'd use it as a stool, not an end table." Marla has often come to the shop over the years and admired the object, but she hasn't been able to afford it. (One Shopper)

It is OK to call the object a stool.



It is OK to call the object an end table.



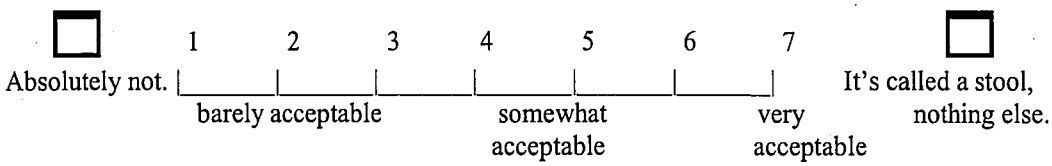
Please explain your choices in a few sentences.

10. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

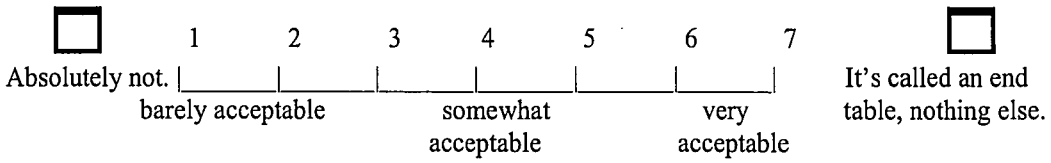
Carpenters in America have been creating objects like this one to be sold as end tables. They've been for sale in their shops as end tables for several years. But a few years ago, the carpenters decided as a group that they would start selling them

as stools. They've been making and selling the objects as stools ever since. They never plan to make or sell them as end tables again. (Many Creators)

It is OK to call the object a stool.



It is OK to call the object an end table.

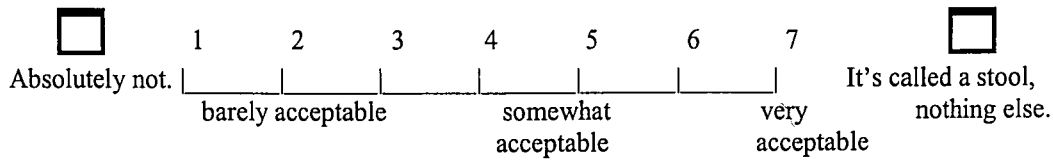


Please explain your choices in a few sentences.

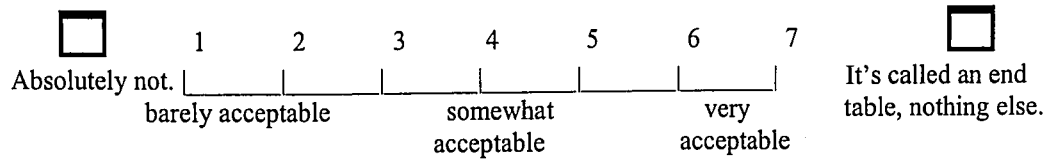
11. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. All the people who have been buying these objects have been buying them to use as stools, not end tables. Although they knew the objects were intended to be used as end tables, they've been using them as stools ever since they bought them. (Many Owners)

It is OK to call the object a stool.



It is OK to call the object an end table.



Please explain your choices in a few sentences.

12. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. John has noticed over the years that his customers would prefer to use the objects as stools; "If I could afford one of those, I'd use it as a stool, not an end table," they often say. (Many Shoppers)

It is OK to call the object a stool.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. | | | | | | | It's called a stool, nothing else.

barely acceptable — somewhat acceptable — very acceptable

It is OK to call the object an end table.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. | | | | | | | It's called an end table, nothing else.

barely acceptable — somewhat acceptable — very acceptable

Please explain your choices in a few sentences.

Appendix D

INSTRUCTIONS ("Really")

On the other side of this paper is a short story about a humanmade object. After you read the story, you'll be asked to rate two statements about the object for their acceptability. The statements will be regarding what the object ***really*** is.

For instance, suppose the first statement is "It is really a hammer." If you think it is ***really not*** a hammer, then mark the box for "*Absolutely not.*" If you think it could reasonably be thought of as ***really*** a hammer, then mark one of the numbers on the scale. A "1" means that you think it's barely acceptable to think of it as ***really*** a hammer, and a "4" means that it's somewhat acceptable. The other end of the scale, a "7," indicates that thinking that the object is ***really*** a hammer is very acceptable. If you think the object is ***really*** a hammer *and nothing else*, then check the box at the far right end of the scale.

Then look at the next statement for that item. It will ask you to think about a second thing that the item might ***really*** be, using the same scale. Please mark the scale accordingly.

In making your two judgments, you might think that an item is ***really*** only one of the two things, or you might think that it is ***really*** both. You might think it is ***really*** neither. The rating scales have been designed to let you express all of these opinions and more.

After you've made your mark on each scale, write a brief explanation under the scales about why you responded as you did for both statements together.

4. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, he decided to start selling them as stools. He has made and sold them as stools ever since. He never plans to make or sell them as end tables again. (One Creator)

It is really a stool.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. barely acceptable somewhat acceptable very acceptable It's really a stool, nothing else.

It is really an end table.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. barely acceptable somewhat acceptable very acceptable It's really an end table, nothing else

Please explain your choices in a few sentences.

5. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

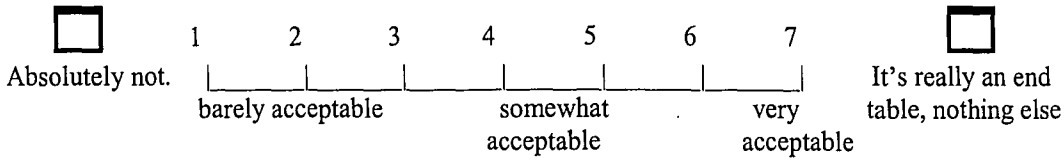
John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, he sold one at his shop to Marla, but Marla didn't buy it to use as an end table. Although she knew it was intended to be used as an end table, Marla bought it to use as a stool and has been using it as a stool ever since. (One Owner)

It is really a stool.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. barely acceptable somewhat acceptable very acceptable It's really a stool, nothing else.

It is really an end table.

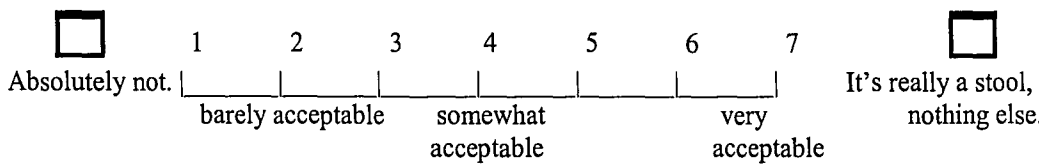


Please explain your choices in a few sentences.

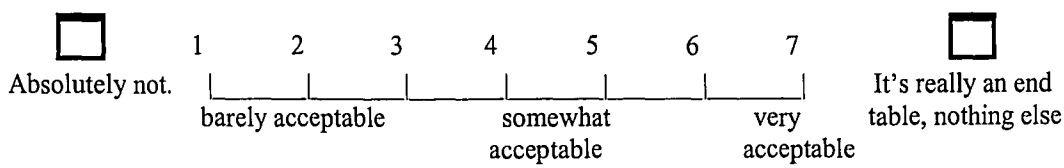
6. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. Then, one day a few years ago, Marla was browsing in the shop and noticed one. She thought, "If I could afford that, I'd use it as a stool, not an end table." Marla has often come to the shop over the years and admired the object, but she hasn't been able to afford it. (One Shopper)

It is really a stool.



It is really an end table.



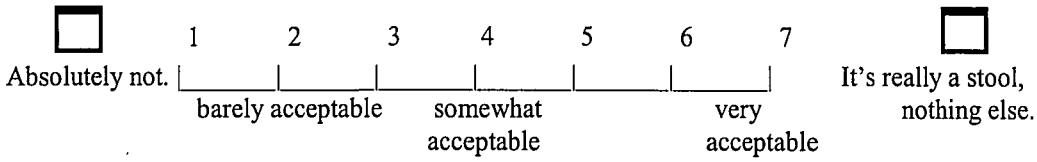
Please explain your choices in a few sentences.

13. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

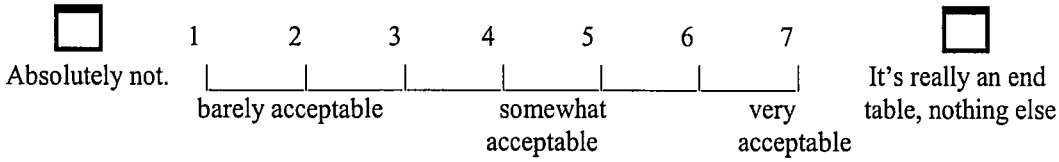
Carpenters in America have been creating objects like this one to be sold as end tables. They've been for sale in their shops as end tables for several years. But a few years ago, the carpenters decided as a group that they would start selling them

as stools. They've been making and selling the objects as stools ever since. They never plan to make or sell them as end tables again. (Many Creators)

It is really a stool.



It is really an end table.

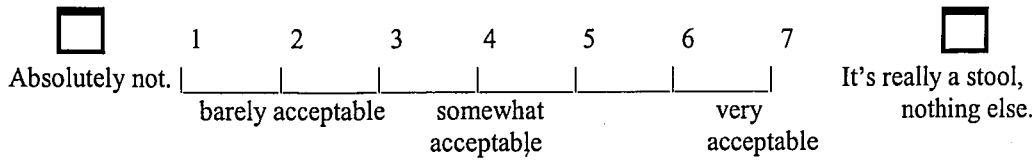


Please explain your choices in a few sentences.

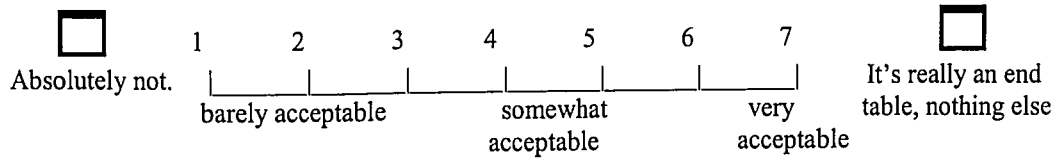
14. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. All the people who have been buying these objects have been buying them to use as stools, not end tables. Although they knew the objects were intended to be used as end tables, they've been using them as stools ever since they bought them. (Many Owners)

It is really a stool.



It is really an end table.



Please explain your choices in a few sentences.

15. Picture an object made of dark varnished wood, about 2.5 feet high, with four cylindrical legs and a flat, 1.5-ft.-by-1.5-ft. top.

John is a carpenter. He created objects like this one to be sold as end tables. They've been for sale in his shop for several years. John has noticed over the years that his customers would prefer to use the objects as stools; "If I could afford one of those, I'd use it as a stool, not an end table," they often say. (Many Shoppers)

It is really a stool.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. It's really a stool, nothing else.

barely acceptable somewhat acceptable very acceptable

It is really an end table.

☐ 1 2 3 4 5 6 7 ☐

Absolutely not. It's really an end table, nothing else

barely acceptable somewhat acceptable very acceptable

Please explain your choices in a few sentences.

Curriculum Vitae

Michael R. Paquet

Lehigh University	1643 Exeter Circle
Psychology Department	Downingtown, PA 19335
17 Memorial Drive East	Cell: (484) 437-6968
Bethlehem, PA 18015	

Personal Information

Date of Birth: April 10, 1985

Place of Birth: Upper Darby, Pennsylvania

Parents: Michael and Barbara Paquet

Education

- M.S., Psychology, 2008, Lehigh University, Bethlehem, PA (GPA 3.96)
- B.A., Psychology, 2006, Lehigh University, Bethlehem, PA (GPA 3.97)
    - Philosophy minor and Journalism practicum
    - Dean's List; Summa Cum Laude

Professional Experience

Psychometrician, Neuropsychology, Allentown, PA

April 2006 to present.

Test patients and score tests for neuropsychologist David Glosser, Sc.D.

Cognitive Psychology Graduate Student, Lehigh University, Bethlehem, PA

August 2006 to May 2008.

Conduct experiments, analyze data, write reports.

Teaching Assistant, Lehigh University Department of Psychology

August 2006 to May 2008.

Prepare and present lectures; grade exams and assignments; hold office hours; organize review sessions; tutor students.

Research Assistant, Lehigh University

January 2006 to May 2006.

Under supervision of Laura Gonnerman, Ph.D., help enter and code data for language development study.

Student Intern, Good Shepherd Hospital Pain Rehab Clinic, Office of Psychology

August 2005 to December 2005.

Observation of clinical psychologist, Timothy Lomauro, PhD, in out-patient interviews, in-patient consultations and psychological assessment.

Honors, Awards

- Presidential Scholarship, Lehigh University, 2006-2007
- First Place, Williams Senior Prize for Philosophy, Lehigh University, May 2006
- Third Place, Williams Senior Prize for Science Writing in Psychology/Sociology/ Anthropology, Lehigh University, May 2006
- Elisha Wilbur Prize, Lehigh University, October 2005

• Dean's List, Lehigh University (Fall 2003, Spring 2004, Fall 2004, Spring 2005, Fall 2005, Spring 2006)

• Asa Packer Scholarship, Lehigh University, 2003-2006

• First Place, Williams Prize for Sophomore English Composition, Lehigh University, May 2004

• Diocesan Scholarship, Archdiocese of Philadelphia, 2002-2003

• Neumann Scholarship, The Connelly Foundation, 1999-2003

• School representative to the Hugh O'Brien Youth Leadership Conference (HOBY), King of Prussia, Pa., 2001

• Humanities and Social Sciences Award of Excellence, University of Rochester, 2002

• Honor Diploma, Junior State of America Summer School, Princeton University, July 2001

• 1st Place, Delaware County Law Day Essay Contest, May 2000

Association Memberships

• Psi Chi National Honor Fraternity

• Phi Beta Kappa

Teaching Experience

• Intro to Cognitive Neuroscience, Teaching Assistant, Lehigh University

• Abnormal Psychology, Teaching Assistant, Lehigh University

• Health Psychology, Teaching Assistant, Lehigh University

• Intro to Social Psychology, Teaching Assistant, Lehigh University

END OF TITLE