On the Psychological Basis for Rigid Designation

Nick Braisby

Department of Psychology
London Guildhall University
Old Castle Street
London E1 7NT, UK
braisby@unixa.lgu.ac.uk

Bradley Franks

Department of Psychology
London School of Economics
Houghton Street
London WC2A 2AE, UK
B.Franks@vax.lse.ac.uk

James Hampton

Department of Psychology
City University
Northampton Square
London EC1 0HB, UK
sc318@city.ac.uk

Abstract

Kripke (1972) and Putnam (1975a; 1975b) have argued forcefully for the philosophical view of word meaning known as rigid designation. While certain psychological studies have appeared to offer this view support (Keil, 1986; Rips, 1989), we argue that these have not provided an exhaustive evaluation. In particular, the original discussions of Kripke and Putnam reveal that their view rests on an explicit appeal to intuition concerning word use in a range of different scenarios. The study reported here investigates word use under three such types of scenarios, for a variety of natural kind terms, by investigating subjects' judgements of truth or falsity for a range of statement types. We argue that the results obtained indicate that the intuition on which rigid designation rests is not one which is generally true of agents' language use. Further, we obtain patterns of apparent contradiction which appear strictly inconsistent with rigid designation and which require an account of word meaning which allows that the sense of words may vary systematically with context (Franks & Braisby, 1990).

Introduction

This paper focuses on the validity of an intuition concerning language use which forms the foundation of the rigid designation view of word meaning (Kripke 1972; Putnam 1975a; 1975b). In the next section, we outline the arguments of Kripke and Putnam in favour of rigid designation and indicate how they crucially depend on an intuition concerning the behaviour of language users in certain possible, though non-actual, circumstances. These are circumstances concerning discoveries, some pertaining to individual members of a category, and others to entire categories. We then outline an experiment designed to examine whether language users characterise their behaviour in the way that Kripke and Putnam intuit. The results suggest that the intuition on which rigid designation is based is faulty and that rigid designation may be too rigid a theory of word meaning for psychological purposes.

Rigid Designation

According to rigid designation, natural kind terms refer to underlying natural kinds, each instance of which shares the same essential properties. As stated, the position has the following entailments concerning the nature of word meaning.

1 Essential properties determine reference.
2 Non-essential (or contingent) properties do not determine reference.
3 Reference is determined independently of people's beliefs about which properties determine reference.

This last entailment gives rise to Putnam's assertion that "'meanings' just ain't in the *head*!" (Putnam, 1975b: 227).

The claim arises because of the ever present possibility that agents may be mistaken in what they assume are the essential properties associated with a word. Fallibility of knowledge does not, however, imply fallibility of reference, for reference, according to 1 and 2, is determined by essential properties, not by properties mistakenly assumed to be essential. For example, though we assume that all cats are mammals, it is conceivable that one could discover that all cats are in fact robots controlled from Mars. That is, one might *assume* the essential properties for "cat" to include "being a mammal", whereas the essential properties for "cat" *actually* include "being a robot". Then, according to rigid designation, the word "cat" never referred to a subset of mammals even though concept possessors might have been under the impression that it did. Instead, the word "cat" always referred to the same category of objects: it is merely that at one time these objects were assumed to be mammals and, at a later time, they were known to be robots controlled from Mars.

The discussions of Putnam and Kripke presuppose that the classification of an individual as a member of a category depends crucially on whether that individual possesses the essential properties for that category. Indeed, a number of studies apparently demonstrate that classification of individuals depends upon their possessing the presumed essential properties of a category (Keil, 1986; Rips, 1989). Despite this apparent support for rigid designation views, the range of contexts utilised in these and other studies do not provide an exhaustive evaluation of rigid designation. We believe this claim is supported by a closer examination of the original discussions of Kripke and Putnam. In particular, these indicate that other contexts are equally important for defining rigid designation. Putnam considers a scenario in which it is discovered that *all* cats are robots controlled from Mars. Nonetheless, he states that, after the

discovery, ``we will still *call* them `cats''' and, further, ``not only will we still *call* them `cats,' they are cats'' (Putnam, 1975a: 142). It is merely that cats are not as they had previously been thought. Kripke (1972) offers a similar example where it is discovered that it is an optical illusion that gold is yellow and that, in fact, it is blue. According to Kripke, this would not entail that people would say "gold does not exist", nor that "there is no such thing as gold". He asks, ``would there on this basis be an announcement in the newspapers: ``it has turned out that there is no gold. Gold does not exist. What we took to be gold is in fact not gold."?'' and answers ``it seems to me there would be no such announcement'' (p. 316).

The scenarios considered by Kripke and Putnam thus contrast with the contexts used by Keil and Rips, in that they involve discoveries pertaining to the presumed essential and non-essential properties of the *entire category*, and not merely pertaining to the essential properties as possessed by a single individual. It is also clear that Kripke and Putnam make an explicit appeal to our intuitions concerning the way in which words are used under certain counter-factual circumstances. It is the validity of this intuition that we seek to address. Our study used the following kinds of scenario, with specific examples given for "cat":

Non-Essential Category: *Non-essential* properties associated with a category are discovered not to be true of that *Category* (from entailment 2 above, word use should not change). E.g., You have a female pet cat named Tibby whom you believe to be able to miaow. Indeed, she seems to miaow quite loudly. However, in spite of the fact that most people assume that cats can miaow, scientists have discovered that the noise of miaowing is actually created by small parasites which live in their fur and that, in fact, *no* cats can actually miaow. Indeed, Tibby, once thoroughly cleared of parasites, cannot miaow.

Essential Individual: *Essential* properties of a category are discovered not to be true of a presumed *Individual* category member (from 1, word use should change with respect to the individual). E.g., You have a female pet cat named Tibby who has been rather unwell of late. Although cats are known to be mammals, the vet, on examining Tibby carefully, finds that she is, in fact, a robot controlled from Mars.

Essential Category: Presumed *Essential* properties of category are discovered not to be the actual essential properties for that *Category* (from 1 and 3, word use should not change). E.g., You have a female pet cat named Tibby. For many years people have assumed cats to be mammals. However, scientists have recently discovered that they are *all*, in fact, robots controlled from Mars. Upon close examination, you discover that Tibby too is a robot, just as the scientists suggest.

To the authors' knowledge, no study has attempted to examine the intuitions behind rigid designation under thisrange of scenarios. Indeed, it is particularly surprising that scenarios of type Essential Category have not been employed, given the emphasis on scenarios of this type in the arguments of Kripke and Putnam. The possibility that rigid designation will not be supported under scenarios of this type forms the focus of this study.

Method

Subjects

Subjects were 37 undergraduate students from London School of Economics and City University London. They were unpaid volunteers.

Design

The experiment investigated rigid designation under the three scenario types: Non-Essential Category, Essential Individual and Essential Category. We selected 7 natural kind categories of the sort which figure prominently in the argumentation of Kripke and Putnam: "cat", "water", "tiger", "gold", "bronze", "lemon" and "oak". The scenarios used were similar to those given earlier. In order to investigate the use of the natural kind terms, we obtained "true or false" judgements about three types of statement: Existential, Qualified and Membership. A positive (+) and negative (-) variant of each type was used. Below are the statements for "cat".

Existential (+): Cats do exist
Existential (-): Cats do not exist
Qualified (+): Cats do exist, and people's beliefs concerning cats have changed
Qualified (-): There are no such things as cats, only robots controlled from Mars
Membership (+): Tibby is a cat, though we were wrong about her being a mammal
Membership (-): Tibby is not a cat, though she is a robot controlled from Mars

Existential statements exactly reflect the form of statement used by Kripke and Putnam in their arguments. However, we believed that the Existentials could be rated *in accordance with* rigid designation without actually entailing it. For example, consider scenario Essential Category and the Existential (-) statement "cats do not exist". Under an "extensional" reading of "cat" (referring to the same set of objects described in the scenario) then the statement is false: the set of particular objects does exist. However, this response of False need not reflect rigid designation. According to rigid designation, not only do the objects exist but they are members of the particular

Table 1 Responses obtained and predictions made by rigid designation (RD)

| Statement | Scenario Type | | | | | |
| | Non-Essential Category | | Essential Individual | | Essential Category | |
	RD	Actual	RD	Actual	RD	Actual
Existential (+)	7	6.89	7	6.41	7	6.19
Existential (-)	0	0.08	0	0.70	0	0.70
Qualified (+)	7	6.35	0	3.24	7	5.84
Qualified (-)	0	0.22	0	1.00	0	1.30
Membership (+)	7	6.57	0	3.57	7	6.30
Membership (-)	0	0.49	7	4.49	0	1.62

natural kind in question. To circumvent this possible interpretation of the natural kind terms, we included Qualified statements which make explicit reference to the set of objects referred to in the scenario. Hence, one should not be able to judge "there are no such things as cats, only robots controlled from Mars" false simply by taking an extensional interpretation of "cat", since the sentence itself militates against such an interpretation. Finally, we included Membership statements to reflect our view that rigid designation has little to do with statements concerning the existence of entire categories. Rather, the issue is, quite simply, whether or not the specified entities lie in the same category both before and after a discovery. Thus, we consider that rigid designation may be best measured by considering Membership statements, these being better indicators than Qualified, and these being better than Existentials.

Hypotheses
We predicted agreement with rigid designation in the Non-Essential Category and Essential Individual conditions, but significantly less agreement in the Essential Category condition. We also predicted that Existential statements would show significantly greater (apparent) agreement with rigid designation than Membership statements, with Qualified statements intermediate.

Results

Responses were coded 1 for True, 0 for False. Table 1 contains the frequency of positive responses obtained together with the predictions of rigid designation and Table 2 contains these responses recoded so as to indicate agreement with rigid designation (maximum 7). All analysis was performed on these recoded scores, so significant differences indicate significant differences in terms of agreement with rigid designation.

As can be seen from Tables 1 and 2, rigid designation appears vindicated only under certain combinations of scenario type and statement type. In particular, rigid designation fares poorly in scenario types Essential Category and Essential Individual under Membership statements. There was no effect of natural kind (Friedman, χ^2=4.70, d.f.=6, p=0.58). All other effects were as predicted. That is, we obtained the expected effect of scenario type (Friedman χ^2=33.95, d.f.=2; p<0.0001): multiple comparisons indicated that the Non-Essential

Category scenario type showed significantly greater agreement with rigid designation than the Essential Category scenario type, which showed greater agreement than Essential Individual (p<0.025: Essential Category mean rank = 2.05, Essential Individual = 1.30, Non-Essential Category = 2.65). There was also a significant effect of statement type (Friedman χ^2=73.29, d.f.=5; p<0.0001); multiple comparisons reveal an over-all significant difference between Existential, and Qualified and Membership statements in their degree of agreement with rigid designation (p<0.025): Existential statements showed greatest agreement, Membership showed least, and Qualified were intermediate (Membership mean rank = 2.42, Qualified = 3.27, Existential = 4.81).

Table 2 Responses recoded to indicate agreement with RD

| Statement | Scenario Type | | |
	Non-Essential Category	Essential Individual	Essential Category
Existential (+)	6.89	6.41	6.19
Existential (-)	6.92	6.30	6.30
Qualified (+)	6.35	6.76	5.84
Qualified (-)	6.78	6.00	5.70
Membership (+)	6.57	3.43	6.30
Membership (-)	6.51	4.49	5.38

In order to understand the apparent disagreement between subjects in these data we also analysed the degree of within-subject inconsistency. An apparent contradiction was said to occur when a S gave either both False responses or both True responses to the positive and negative forms of a statement type, within the same scenario. For example, a subject responding False to "Water does exist" and to "Water does not exist", within the same scenario type, would be considered to be expressing an apparent contradiction. In such cases, it was considered that subjects would be expressing a logical contradiction, unless the natural kind term in question had different senses in the positive and negative forms of the statement. In the following analysis we omitted the Qualified statements in Essential Individual scenario types since it was, in fact, logically possible for subjects to respond False to both the positive and negative forms of this statement type, without this implying different senses for the natural kind term. Since this cell contains the largest number of contradictions, we believe our analysis is erring on the side of rigid

designation, since this latter cannot readily account for contradictions of any kind.

Table 3 shows that the majority of contradictions were obtained when Membership statements were used under Essential Individual and Essential Category scenario types and when Qualified statements were used under Essential

Table 3 Number and proportion of apparent contradictions

Statement	Scenario Type		
	Non-Essential Category	Essential Individual	Essential Category
Existential	3 (1%)	10 (4%)	17 (7%)
Qualified	30 (12%)	NA	49 (19%)
Membership	24 (9%)	51 (20%)	40 (15%)

Category scenario types. An analysis of variance on the number of apparent self-contradictions revealed a main effect of scenario type ($F_{2,72}$=227.80; p<0.0001) with the following ordering (Non-Essential Category < Essential Individual < Essential Category) with mean number of contradictions, in that order, being, 0.51, 0.82 and 0.91). There was also a main effect of statement type ($F_{2,72}$=183.15; p<0.0001) with the Existential statements showing fewer contradictions than both the Qualified and Membership statements, these producing marginally fewer contradictions than the Qualified statements (mean number of contradictions, in that order, were, 0.23, 1.07 and 1.04). Additionally, there was an interaction between scenario and statement type ($F_{2,73\ (adj.)}$=258.31; p<0.0001). Again, it appears that only certain combinations of scenario and statement types result in substantial levels of contradictions.

Discussion

Existential statements show agreement ranging from 87% to 98% with the rigid designation view. While this may appear to support the prediction of the view reasonably well, the pattern for other statement types was not so clear. It was argued that the strictest test of the theory would be found with the Membership statements, which concerned the effect of the discovery on category membership of an individual case. Rigid designation fared much worse in the Essential Individual and Essential Category scenarios when Membership statements were given (agreement with rigid designation ranged from 47% to 76%) or when Qualified statements were given, (46% to 73%). According to rigid designation word use should follow *exactly* the predictions indicated in Table 1. This is for the simple reason that if rigid designation is a true account of word meaning then referential word use is *not* a matter of degree, and does *not* vary from one context to another. Thus the results reported here can not be taken as providing much comfort for the thesis.

It may be considered unreasonable to expect subjects to respond consistently on all occasions in our task. However the difficulties for rigid designation cannot easily be explained by simply claiming that subjects did not understand their task or that they found it too diffcult to respond consistently. First, when subjects were asked to

rate the difficulty of their task, the mean rating showed that subjects viewed the task as neither easy nor difficult. Second, if in general the task *had* been poorly understood then responding should have been more nearly random in nature. The fact that in certain combinations of scenario and question type (particularly the Non-Essential Category scenario) subjects responded consistently (and in agreement with rigid designation) may rule out this possibility. Most importantly, since the range of scenarios employed is exactly that used by Kripke and Putnam, the failure of educated lay people to share their intuitions must constitute a serious challenge to their argument.

Our pattern of results contrasts with earlier research by Keil (1986) and Rips (1989). In particular, the results for the Essential Individual scenario with both Qualified and Membership statements appear inconsistent with their findings. Keil used an individual discovery scenario like ours, with both natural kinds and artifacts. His adult group judged that discoveries about essential features of individuals could change the categorization of natural kinds but not of artifacts. By contrast, at least 50% of our subjects in the Essential Individual scenario considered that the category membership was unaffected. A possible source of explanation here, could be the framing of the experiment itself. Keil's work was done as an adult 'control' for a primarily developmental study. It is possible that his adult subjects were therefore inclined to treat the task in a different way.

Rips (1989) introduced scenarios in which individuals were transformed (for example by exposure to radiation) from one appearance to another, and showed that people are inclined to say that category membership is unchanged by such transformations. One major difference between our study and that of Rips was that our scenarios involved discoveries rather than transformations.

A difference in procedure between our study and both earlier studies was that we provided subjects with the possibility of agreeing with both positive and negative categorization statements. A surprising number of people availed themselves of this opportunity to be illogical, which suggests to us the importance of this procedural variation.

Our over-all pattern of results is however consistent with a study by Malt (1991). Her findings revealed discrepancies between the predictions of essentialism and the extensions of terms, as defined by data concerning the kinds of object to which noun terms might be used to refer. However, Malt's strategy differed from ours in that her concern was how language is used to refer to objects in everyday parlance. As such, these uses are likely to include some which are referential rather than attributive (Donnellan, 1969). Referential uses are those which use a term to pick out an individual, without the intention that it is an accurate description of that individual (for example in the Method section we described how the term "Cats" in the statement "Cats do not exist" could be interpeted extensionally -- that is as referring to the particular class of individuals which had hitherto been called cats.) Malt's approach is also more correlational and concerned with the relation between language use and the structure of the lexicon. By contrast, in order to make an appropriate test of the rigid designation

thesis we have focussed exclusively on attributive uses of natural kind terms and employed an experimental design in which scenarios are presented in a controlled way.

One possible reason for the discrepancy between our results and the predictions of the rigid designation thesis may be that subjects respond in a contradictory fashion, a possibility that the positive and negative pairs were designed to evaluate. Franks & Braisby (1990) proposed an account of the relation between concepts and word use whereby the same word, even when used attributively may be used with different conceptual content in different contexts. A consequence is that the same entity may be classified as both a member and a non-member of the extension of a term without contradiction. Classification is thus treated as relative to the perspective adopted. This view echoes other treatments of the perspectival nature of communication (e.g., Clark, 1992). On this view of classification, the patterns of word use obtained in the present experiment, including apparent contradictions, could be explained by assuming that subjects are tempted into adopting different perspectives in responding to positive and negative versions of the same statement, and are thereby using the same natural kind term with different conceptual content. Of course, it was not the intention of the experiment to explicitly manipulate perspectival interpretation, and so our account remains strongly post hoc.

Similar apparent inconsistencies in categorization have been reported by Hampton (1982, 1988a, 1988b). Hampton (1982) found that people were willing to accept the truth of two apparently contradictory statements such as "Chairs are a kind of furniture" and "Not all chairs are furniture". In Hampton (1988a, 1988b) the extensions of conjunctions and disjunctions of natural categories were compared with the extensions of their two constituents, and a degree of inconsistency was identified. It appeared that the interpretation of a category could be influenced by placing it in disjunction with another closely related category. The possibility exists that a perspectival shift account may also provide an explanation of these other phenomena.

In conclusion, we have demonstrated that when faced with scenarios of the kind made famous by Kripke and Putnam, people are divided in how to respond, and may even make apparently inconsistent responses. Since their argument depends on the generality of their intuitions about language use, severe doubt has been cast on the case for their thesis. For psychological purposes, it would appear, rigid designation, with its essentialist leanings, may be just too rigid.

References

Clark, H.H. (1992) *Arenas of Language Use.* Chicago: Chicago University Press.

Donellan, K. (1966) Reference and definite description. *Philosophical Review, 75,* 281-304.

Franks, B., & Braisby, N. R. (1990). Sense generation or how to make a mental lexicon flexible. In *Proceedings of the 12th annual conference of the cognitive science society*. MIT, July 1990.

Hampton, J.A. (1982) A demonstration of intransitivity in natural categories. *Cognition, 12,* 151-164.

Hampton, J.A. (1988a) Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion. *Journal of Experimental Psychology: Learning Memory and Cognition, 14,* 12-32.

Hampton, J.A. (1988b) Disjunction of Natural Concepts. *Memory & Cognition*, 16, 579-591.

Keil F. (1986). Conceptual Development and Category Structure. In Neisser U. (ed.), *Concepts and Conceptual Development*. Cambridge: CUP.

Kripke S. A. (1972). Naming and Necessity. In Davidson D. & G. Harman (eds.). *Semantics of Natural Languages.* Dordrecht: Reidel.

Malt, B. (1991) Word meaning and word use. In P.Schwanenflugel (Ed.) *The psychology of word meanings.* Hillsdale NJ: Erlbaum.

Putnam H. (1975a). Is semantics possible? Chapter 8 in *Mind, Language and Reality, Volume 2: Philosophical Papers*. Cambridge: CUP.

Putnam H. (1975b). The Meaning of "Meaning". Chapter 12 in *Mind, Language and Reality, Volume 2: Philosophical Papers*. Cambridge: CUP.

Rips L. (1989). Similarity, Typicality and Categorisation. In Vosniadou S. & Ortony A. (eds.), *Similarity and Analogical Reasoning*. Cambridge: CUP.