Why is Language Vague?[1]

Barton L. Lipman

Department of Economics
University of Wisconsin
1180 Observatory Drive
Madison, WI 53706

email: blipman@ssc.wisc.edu

First Draft

October 2001

[1]I thank Eddie Dekel and Marie–Odile Yanelle for helpful conversations.

Abstract

I don't know.

Warning: I really don't know. This is more a research proposal than a research paper, even if it does apply the word "theorem" to some statements.

"Is he — is he a tall man?"

"Who shall answer that question?" cried Emma. "My father would say, 'Yes'; Mr. Knightly, 'No'; and Miss Bates and I, that he is just the happy medium."

— From *Emma* by Jane Austen.

1 Introduction

When one thinks about language as spoken by people on a day–to–day basis, it is hard to ignore the fact that much of what is said is vague. Consider, for example, the word "tall." There is no precise, known height which defines the line between a person who is tall and a person who is not. Why do we use a language in which such terms are so prevalent? Why don't we simply adopt as a definition that "tall" will mean above, say, 6 foot 2? We could even adopt a context–specific definition, saying for example that "tall" for a newborn means above 15 inches, while "tall" for a professional basketball player means above 6 foot 10.

In this paper, I will argue that we cannot explain the prevalence of vague terms in natural language without a model of bounded rationality which is significantly different from anything (that I know) in the existing literature. In a nutshell, the argument is that any model along existing lines will imply that a precise language like the one described above would Pareto dominant the vague language we see in every society in history. Of course, it seems rather far–fetched to conclude that we have simply tolerated a world–wide, several–thousand–year efficiency loss. Further, even a moment's reflection will suggest that it is easier to speak when one is allowed to use vague language than it would be if such language were banned. Hence this dominance surely tells us that there is something wrong with the model, not the world.

First, let me try to be more precise about what I mean by vagueness. Following Sainsbury [1990],[1] I will say that a word is *precise* if it describes a well–defined set of objects. By contrast, a word is *vague* if it is not precise. Hence the alternative definition given above for "tall" would make this term precise, whereas it is vague in its current usage.

I emphasize that vagueness (as I use the term) is *not* the same thing as less than full information. To say that a person's height is above six feet is precise in the sense that it

[1] All papers cited here but not listed in the references are contained in Keefe and Smith [1996], an excellent introduction to the philosophy literature on this topic.

defines a set of people unambiguously. It is less informative than saying that a person's height is 6 foot 2, but it is *not* vague as I use the term.

The classic illustration of vagueness is what is referred to in the philosophy literature as the *sorites paradox*. For one version of this paradox, note that two facts seem clear about the way people use the word "tall." First, anyone whose height is 10 feet is tall. Second, if one person is tall and a second person's height is within 1/1000 of an inch of the first, then the second person is tall as well. But then working backward from 10 feet, we eventually reach the absurd conclusion that a person whose height is 1 inch is tall. Of course, the source of the difficulty is the vagueness of the word "tall." Because there is no fixed set of heights which "tall" corresponds to, no line which separates "tall" from not, one has the feeling that a small change should not matter. Of course, many small changes add up to a big change, so this is not consistent.

Many other words are vague in this sense. Some words which can easily yield a sorites paradox are "bald," "red" (imagine a sequence of objects moving continuously from red to orange to yellow), "thin," "child" (imagine a sequence of people, each one second older than the previous), "many," and "probably." Some philosophers have constructed less obvious sorites paradoxes for "tadpole," "chair," and other seemingly clear–cut terms.

Many of these terms would be difficult to redefine in a precise way. On the other hand, many could be given such redefinitions. As noted above, "tall" could be defined to be over a specific height, "bald" could be defined by a specific fraction of the skull covered by hair, etc.[2] In such cases, why is the precise redefinition not adopted?

While it seems clear that it would be more difficult to converse in such a precise language, it is not obvious why. Some philosophers of language have tried to address this point. For example, Wright [1976, page 154] observes that "the utility and point of the classifications expressed by many vague predicates would be frustrated if they were supplied with sharp boundaries." Sainsbury [1990, page 251] more colorfully asks of vagueness: "is it so obvious that it is a Bad Thing, given the extent to which the throbbing centres of our lives appear to be describable only in vague terms?" With relatively few exceptions, however, they have not explained what exactly it is about vague terms that make them preferred to their precise analogues.

In the next section, I give a simple but reasonably general model and show that vague terms are Pareto dominated by precise ones in a certain sense. As I will argue, this model, while simple, illustrates that many seemingly obvious reasons why language is vague cannot be made a part of any reasonably standard model. I will argue that in

[2] Of course, sometimes we do develop such definitions for legal purposes, most notably in the case of the word "child." On the other hand, it is clear that common usage of this word outside the legal context is not based on such a precise definition.

some of these cases, the fault is with the model, not the argument for the optimality of vagueness. In Section 3, I argue that explaining the value of vagueness requires a new kind of model of bounded rationality. I suggest that such a model can potentially formalize some of the arguments which do not fit into the model in Section 2.

2 The Suboptimality of Vagueness

In this section, I give a simple model and use this to show why many seemingly obvious advantages of vague terms do not fit into any standard model.

So consider the following two player game, a version of the standard Crawford–Sobel [1982] sender–receiver game.[3] Player 1 observes h, a random draw from the set H with distribution function F. For example, he may observe the height of a person (which the letter h is intended to suggest). He then chooses a message m from a set M. 2 observes this message but not the value of h and then chooses an action a from a set A. I assume that messages themselves are costless (think of a message as 1 making a verbal statement to 2). Also, it is clear that if the objective functions of the agents differ, then 1 may wish to hide information from 2 and hence may wish to be vague in some sense. This does not seem to be the real reason language is vague, so I rule this out by assuming that the two agents have the same utility function u over (h, a) pairs.

A pure strategy for player 1 is a function $s_1 : H \to M$. A mixed strategy for 1 is a probability distribution σ_1 on the set of pure strategies.[4] A pure strategy for player 2 is a function $s_2 : M \to A$, while a mixed strategy is a probability distribution σ_2 over the set of 2's pure strategies.

Clearly, if there are as many messages as there are possible values of h, the best equilibrium for 1 and 2 is for 1 to tell 2 the exact value of h. That is, 1's strategy should be invertible. Trivially, vagueness is suboptimal in such a model.

While this assumption is often viewed as natural, I think it is not particularly realistic as a model of language. Consider the case where h is the height of some individual. The set of possible heights presumably is uncountable, but the set of words we can use to state a height (in, say, less than a century) is finite. Even if one does not accept this view, I will argue below that models of bounded rationality in the literature can only be used to explain vagueness if vagueness is an implication of M being too small. Hence demonstrating that vagueness is suboptimal even when M is small is both less trivial

[3] One can also view this game as a generalization of Dow [1991].

[4] It is more common to work with behavior strategies. Of course, it is equivalent to work with mixed strategies and this is more convenient for my purposes.

and more important to the argument I wish to make. I make no assumptions about the size of M, but consider the interesting case to be where M is finite.

What would it mean to have a vague language in equilibrium? Note that any pure strategy for player 1 would give a precise language, not a vague one. A pure strategy specifies a message as a function of h. Given any message, then, 2 knows which set of h's lead to this message. In this sense, the message has a precise meaning: it corresponds to a particular set of h's. (I discuss an alternative formulation below.)

The obvious way to try to obtain vagueness, then, would be for player 1 to use a (nondegenerate) mixed strategy. With such a randomization, player 2's interpretation of a given message would be that 1 is more likely to use this message to describe some h's than others. In this sense, 1's message is vague.

However, the following theorem demonstrates that this is suboptimal under very weak conditions. First, some notation: For any pair of mixed strategies, (σ_1, σ_2), let $V(\sigma_1, \sigma_2)$ be the expected utility of player 1 (which equals the expected utility of 2) if these strategies are played. I will write $V(\sigma_1, \sigma_2) = V(s_1, \sigma_2)$ when σ_1 is a degenerate mixed strategy with probability 1 on the pure strategy s_1 and similarly for strategies of player 2. Let

$$V^* = \sup_{(\sigma_1, \sigma_2)} V(\sigma_1, \sigma_2).$$

Theorem 1 *If there is a pair of strategies (σ_1, σ_2) such that the supremum is attained, then every pair of pure strategies (s_1, s_2) in the support of (σ_1, σ_2) is a pure Nash equilibrium in which V^* is the expected payoff.*

Put differently, vagueness *cannot* have an advantage over specificity and, except in unusual cases, will be strictly worse.

The proof is almost trivial. (Some readers may prefer to drop the word "almost" from the preceding sentence.) Let (σ_1^*, σ_2^*) be a pair of strategies satisfying $V(\sigma_1^*, \sigma_2^*) = V^*$. Obviously, (σ_1^*, σ_2^*) is a Nash equilibrium — any deviation by either player (weakly) reduces V and hence the payoff to both players. Since σ_1^* is an optimal strategy for player 1, it must yield the same payoff as any[5] pure strategy in its support. Hence there is a pure strategy s_1 such that $V^* = V(s_1, \sigma_2^*)$. Again, this must be a Nash equilibrium since any deviation by either player leads to a weakly lower payoff for both. Again, since σ_2^* is optimal for 2, then for any pure strategy s_2 in its support, we have $V^* = V(s_1, s_2)$. By the same argument as before, (s_1, s_2) is a pure Nash equilibrium. ▮

[5] Or almost any if the distribution is not discrete.

Remark 1 This result is not much more than Corollary 2.2 of Monderer and Shapley [1996].

Remark 2 One could give a similar result for the case where the supremum is not achievable. It is not hard to see that for any pair of strategies (σ_1, σ_2) achieving a payoff within ε of V^*, there is (s_1, s_2) in the support of (σ_1, σ_2) which is a pure ε–Nash equilibrium with a payoff within ε of V^*. A proof: Fix (σ_1, σ_2) with payoff within ε of V^*. Obviously, this must be a ε–Nash equilibrium. Player 1's payoff is, of course, a convex combination of the payoffs he gets from his various pure strategies. Hence there must be a pure strategy s_1 such that $V(s_1, \sigma_2) \geq V(\sigma_1, \sigma_2)$. Clearly, this yields both players payoffs within ε of V^* and so is a ε–Nash equilibrium. A similar argument for player 2 completes the proof.

Remark 3 These games often have no (nontrivial) mixed equilibria, even if we drop the assumption that the players have the same preferences over (h, a) pairs. For example, the original Crawford–Sobel [1982] model assumes that player 2's payoff function is strictly concave in his own action and that the set of actions is convex. Given this, 2's best reply is always pure. The model adds sufficient structure on 1's payoffs to ensure that the set of types who are indifferent between inducing two different actions is of measure zero, implying that 1 (essentially) does not randomize.

One might object to my identification of vagueness with mixed strategies. In particular, it is possible that 1 can use a pure strategy and 2 does not know the "true meaning" of 1's messages. To see the point, suppose $H = \{L, M, R\}$ and $M = \{m_1, m_2\}$. Suppose that L and M are "the same" in the sense that they induce the same preferences over actions. That is, letting u denote the utility function for 1 and 2 over (h, a) pairs, we have $u(L, a) > u(L, a')$ if and only if $u(M, a) > u(M, a')$. Suppose we have an equilibrium in which 1's strategy is $m(L) = m(R) = m_1$ and $m(M) = m_2$. What does player 2 learn when he receives message m_1? In one sense, his information is precise: the signal received by 1 was either L or R. On the other hand, one could interpret this differently. One could say that what is really relevant is whether the signal is R or in $\{L, M\}$. In this sense, m_1 conveys noisy information. It lowers the probability of $\{L, M\}$ but does not refute it.

This kind of vagueness cannot be optimal either. To state this more precisely, for each h, define a preference relation over A by $a \succ_h a'$ if and only if $u(h, a) > u(h, a')$. Say that h and h' are equivalent, denoted $h \sim^* h'$ if they induce the same preferences over actions or $\succ_h \; = \; \succ_{h'}$. If h and h' are equivalent, then, neither player 1 nor player 2 care about which of these two signals is received.[6] In light of this, I will say that strategy σ_1 *induces*

[6] One might object to this statement, since I have defined h and h' to be equivalent if they induce the

a precise language if it is pure and $\sigma_1(h) = \sigma_1(h')$ whenever $h \sim^* h'$. Otherwise, I will say that σ_1 induces a vague language.

Theorem 2 *Suppose there is a pair of strategies (σ_1, σ_2) such that $V(\sigma_1, \sigma_2) = V^*$. Then there must be an equilibrium (σ_1', σ_2') which also achieves payoff V^* with the property that σ_1' induces a precise language.*

Proof. From the previous theorem, we know that given the condition stated, there is a pure strategy equilibrium achieving V^*. So fix such an equilibrium, say (s_1, s_2). If s_1 does not induce a precise language, then, by definition, we must have some h and h' with $s_1(h) \neq s_1(h')$ even though $h \sim^* h'$. Note that for any such h and h', it must be true that $u(h, s_2(s_1(h))) = u(h, s_2(s_1(h')))$. That is, player 1 must be indifferent between sending the message $s_1(h)$ or $s_1(h')$ when the signal is either h or h' — otherwise, he'd use the better of the two in response to either signal.

In light of this, construct a new strategy $\hat{s}_1$ as follows. Fix a subset of H, say H^*, with the property that for every $h \in H$, there is exactly one $h^* \in H^*$ with $h \sim^* h^*$. That is, H^* takes one representative from each equivalence class under $\sim^*$. Let $\hat{s}_1(h^*) = s_1(h^*)$ for all $h^* \in H^*$. For any other h, define $\hat{s}_1(h)$ to equal $\hat{s}_1(h^*)$ for that $h^* \in H^*$ with $h \sim^* h^*$. Clearly, $\hat{s}_1$ induces a precise language. (If s_1 already induced a precise language, $\hat{s}_1 = s_1$.)

Note also that $V(\hat{s}_1, s_2) = V(s_1, s_2) = V^*$ since $\hat{s}_1$ and s_1 differ only in which of several equally good messages to send. Hence, by definition of V^*, there cannot be any s_2' with $V(\hat{s}_1, s_2') > V(\hat{s}_1, s_2)$. That is, $(\hat{s}_1, s_2)$ is a Nash equilibrium. ▌

An Example.

1 must describe Mr. X to 2 who must pick him up at the airport. 1 knows X's height; 2 doesn't. Height is continuously distributed on the interval [0,1] with density function f and distribution function F, independently across people. Let μ denote the median height. There will be n people at the airport in addition to Mr. X. 2 must pick one of these $n + 1$ people and ask if he is X. If he guesses right, he and 1 get a payoff of 1. Otherwise, they both get 0. 2 cannot observe the exact heights of the people at the airport, but does observe relative heights. That is, he sees the height ranking of each person.

What is the optimal language when $\#M = 2$? Surprisingly, the answer is independent of n and depends on F only through its median. It is for 1 to say m_1 when X's height is

same preferences over A, not over lotteries over A. In other words, risk attitudes could differ between h and h' even if $h \sim^* h'$. As we will see, such a difference will be irrelevant.

between 0 and μ and m_2 otherwise. 2's strategy is to try the shortest person in the first case and the tallest in the second. In fact, this is the unique equilibrium in which 2 uses a pure strategy.

To see this, let 2's strategy be to choose the $(k+1)^{\text{th}}$ tallest person in response to one word and the $(k+d+1)^{\text{th}}$ tallest in response to the other, where $d > 0$. It is not difficult to show that given k and d, there is a unique c with the property that 1's payoff to sending the first word exceeds his payoff to sending the second iff $h \in [0, c]$. Hence given any pure strategy by 2, 1's best reply will always partition $[0, 1]$ into two intervals. Therefore, we know that the language will have 1 send one message if $h \in [0, c]$ and the other if $h \in (c, 1]$ for some c.

Consider 2's best reply if he receives the message corresponding to $(c, 1]$. His payoff to choosing the $(k+1)^{\text{th}}$ tallest person is proportional to

$$\Pr[h \in (c,1] \text{ and } k \text{ people taller}] = \binom{n}{k} \int_c^1 [1 - F(h)]^k [F(h)]^{n-k} f(h)\, dh.$$

Define this expression to be $\varphi(k)$. For $k \geq 1$, we can integrate by parts to obtain

$$\begin{aligned} \varphi(k) = \tbinom{n}{k} \quad & \tfrac{1}{n-k+1} [1 - F(h)]^k [F(h)]^{n-k+1} |_c^1 \\ & + \tbinom{n}{k} \ \int_c^1 \tfrac{k}{n-k+1} [1 - F(h)]^{k-1} [F(h)]^{n-k+1} f(h)\, dh. \end{aligned}$$

But $F(1) = 1$ and

$$\binom{n}{k} \frac{k}{n-k+1} = \binom{n}{k-1},$$

so

$$\varphi(k) = \varphi(k-1) - \binom{n}{k} \frac{1}{n-k+1} [1 - F(c)]^k [F(c)]^{n-k+1}.$$

Because the term being subtracted on the right–hand side must be positive, we see that $\varphi(k)$ is decreasing in k. Hence the optimal choice for 2 is $k = 0$ — that is, to choose the tallest person. An analogous argument shows that when 2 conditions on $[0, c]$, the optimal choice is the shortest person.

Given these options, it is easy to see that 1 prefers 2 to pick the tallest person iff $[F(h)]^n > [1 - F(h)]^n$ or $F(h) \geq 1/2$. Thus we must have $c = \mu$. Because this is the only pure strategy equilibrium, Theorem 1 implies that it gives the optimal language.

This game typically has mixed equilibria as well. For example, if the distribution is uniform and $n = 2$, there is an equilibrium where 1 sends m_1 if $h \in (.5 - (\sqrt{3}/6), .5 + (\sqrt{3}/6))$ and m_2 otherwise. In this equilibrium, 2 chooses the middle person when receiving m_1 and randomizes between the tallest and shortest when receiving m_2. The expected payoff in this equilibrium is about .43, while it is .583 in the pure equilibrium. ∎

This simple model illustrates my claim that many seemingly obvious advantages of vague terms cannot be made part of anything like a standard model. Perhaps the most immediate argument for the usefulness of vague terms is that one cannot observe the height of an individual precisely enough to be sure in all cases of how to classify someone in such a precise language. But this objection runs into an immediate reply as follows. Suppose I cannot observe height exactly. Then I should form subjective beliefs regarding which category an individual falls into. The theorem above then says that the optimal language will be precise about such probability distributions. More formally, reinterpret the h above not as height but as a signal of height. The above result then says that it is optimal to partition the set of signals in a precise way. This is equivalent to partitioning the set of induced beliefs and communicating them precisely.

An obvious reply to this is that real people do not form precise subjective beliefs. I believe this is true, but it cannot fit into anything like a standard model of bounded rationality. For example, it is not enough to replace probability distributions with non-additive probabilities. If agents have nonadditive probabilities, then it is surely optimal to partition the set of such beliefs precisely.

Another seemingly obvious advantage to vague language is that it makes context–sensitivity easier. For example, if "tall" is not given a precise definition, I can use it to describe a newborn whose height is 2 feet or a professional basketball player whose height is 7 feet. Again, this objection cannot fit: we could make the precise definitions context–specific. That is, "tall" could be defined to be greater than or equal to 18 inches for a newborn and greater than or equal to 6 foot 10 for a professional basketball player. In terms of the model above, it would be simple to add a random variable which is observed by both players and interpreted as the context. The result above would imply that in each context, it is optimal to have a precise language, though the language might vary with the context in general.

A natural objection to this point is that it is cognitively difficult to remember all the relevant cutoff points. Of course, the key question is not whether it is difficult to remember the cutoff points corresponding to words in a precise language but whether it is more difficult than remembering a vague meaning. Again, this requires a model of bounded rationality different from anything in the existing literature.

The existing models often consider the complexity of rules such as the one giving the appropriate cutoffs for a given context or the complexity of the language in other senses. However, this is unlikely to deliver the conclusion that vagueness is optimal. Such a model could well deliver the conclusion that fewer words are used than would be used in the absence of complexity costs since this would be a way of simplifying. However, this effectively gives only a model of how many messages there are. As noted earlier, the size of the message set has no bearing on the conclusion that vagueness is suboptimal. The real problem is that these models do not consider simplifying the language by introducing some form of vagueness.

Another argument might be that the model above ignores the fact that people must learn to use a language. In particular, one explanation for the prevalence of vague terms which has been suggested in the philosophy literature is that vague terms might be easier to learn. Sainsbury [1990, page 262] argues that "we acquire the concept from the inside, working outwards from central cases ... rather than from the outside, identifying boundaries and moving inwards." As I understand the argument, the idea is rather natural. Suppose that I enter a population of people who have agreed on the cutoff height between "short" and "tall." The only way I can learn about this cutoff, however, is to observe how they classify various people. Naturally, unless I get very lucky and observe the classification of two individuals very close to but on opposite sides of the boundary, I will never learn the cutoff being used. Instead, I will end up learning that people below, say, 5 foot 5, are short, while people above, say, 6 foot 6, are tall and I would have no clear idea about people in the middle. In this sense, what I have learned is a vague term. That is, even if we begin with a precise term, as new agents come into the population and old ones die out, the term will become vague.

While this idea has some appeal for terms based on variables that are difficult to quantify (such as "nice" or perhaps Sainsbury's example of "red"), it is hard to accept when applied to a notion like height. Surely in the fictitious world described above, I should be able to open a dictionary and get out a tape measure to see exactly how tall is "tall."[7]

A related but different critique is that words are vague because we don't know the strategies of other people. If 2 does not know 1's pure strategy, then 2 does not know the intended meaning of the words 1 uses. This idea can also be incorporated into the model. Suppose we add a payoff irrelevant signal to what player 1 observes and assume this is not observed by 2. In such a model, there might be equilibria in which player 1's usage of words to describe the h he has seen depends on this signal. Since 2 does not observe the signal, 2 does not know how 1 is using words. This would give an equilibrium quite similar to what one would obtain using mixed strategies. In particular, the criticism of mixed strategy equilibria above would apply equally well to these equilibria. In other words, just as mixed strategy equilibria are (at least weakly) Pareto inefficient, an equilibrium in which people aren't sure of the pure strategies chosen by others is Pareto inefficient.[8]

[7]Another problem is that it is hard to see why the language won't ultimately collapse. After all, each generation learns a vague version of the previous generation's usage. For example, one can give processes for which the language converges to a distribution of cutoffs which is uniform on the set of heights, as uninformative a language as one could imagine.

[8]One could argue that the players might not have a common prior over these various languages 1 might use. However, this introduces a form of "cheating." As noted earlier, if players have different preferences, it is not surprising that one may try to hide information from another. Different priors is

Theorem 2 implies this claim directly.

Another version of this argument would be to suppose that player 1 observes *two* payoff relevant signals. This would make the set H above a product space. In such a world, even a precise language will not give a precise statement on each dimension. To see the point, return to the airport example and suppose 1 observes both height and weight of Mr. X. A partition of the set of height–weight pairs will not necessarily induce a clear partition on the set of heights. For example, it might be optimal to lump height and weight together into some general comparison of "large" versus "small."

I think this objection also misses the mark for two reasons. First, the optimal language here is precise, even if it is not precise in either dimension. Real language is not precise in this sense. Second, "tall" seems to be a clearly one dimensional notion and this argument cannot explain vagueness in one dimension. Many other vague terms ("thin" and "bald" for example) seem similarly one dimensional.

Some readers may perceive a different flaw in the argument. The sorites problem, for example, does not seem to depend on communication, only the inability of a single agent in isolation to made a categorization. I believe such a criticism is misplaced. There is no meaning to the word "tall" aside from what people interpret it to be. If a person says he cannot say for sure whether a particular person is "tall," surely this means he is not sure how most people would categorize this person or how he would best describe this person to others, not that he is unable to make a choice. Put differently, we use words to communicate: there is only an answer to the question of the minimum height of a tall person if we decide to use the word in such a way. Inherently, then, it is indeed a game theory problem.

3 Vagueness and Bounded Rationality

As noted numerous times in the preceding section, it may be possible to understand the prevalence of vague terms in a model which incorporates a different kind of bounded rationality than has been studied in the extant literature. In this section, I first try to say why existing models of bounded rationality will not help. I then sketch three possible directions in increasing order of ambitiousness (and vagueness!).

As I see it, vagueness requires a different way of thinking about information and what it is. In most models of bounded or unbounded rationality, information takes the form of an event in some state space. That is, when an agent receives information, what he learns is always treated as being the fact that the state of the world lies in some set. The

simply a way to introduce differences in preferences.

learning may be incomplete and even systematically flawed, but the ultimate conclusion takes this form. In this sense, what is learned is precise. I don't know how we can mathematically represent vague knowledge, but I believe that this is what is called for.

Turning to some possible approaches, first, I think that a natural intuition is that vagueness is easier than precision, for the speaker, listener, or both. Intuitively, for the speaker, deciding on which precise term to use may be harder than using a vague term. For the listener, information which is too specific may require more effort to analyze. With vague language, perhaps one can communicate the "big picture" more easily. Several of the points mentioned in the previous section can be thought of as versions of this approach. This requires a different model of information processing than any that I know of in the literature.

An approach which seems more difficult still would be to focus on the relationship between vagueness and unforeseen contingencies. That is, when I use a word, I do not know all the possible situations where you would use my information and hence I might want to "hedge" my bets and be vaguer.[9] A very concrete example of this comes from contract theory. Instead of attempting to specify in exact terms what a party to the contract is supposed to do, contracts often use vague terms such as "taking appropriate care" or "with all due speed." If agents fear that circumstances may arise that they currently cannot imagine, then they may wish to avoid being too precise in order to avoid being trapped later. Instead, they require the other party to respond to unexpected circumstances "appropriately," relying on the hope that the meaning of this word will be sufficiently clear *ex post*.[10] While there has been more work in recent years on this topic, I think it fair to say that there is still no reasonably plausible model of unforeseen contingencies which could be used to analyze such contracts. For a survey, see Dekel, Lipman, and Rustichini [1998].

To understand the third suggestion, let us return to one comment I made in the previous section. I noted that if 1 did not know the height of the person he is describing to 2, he should form a subjective probability distribution over it and communicate this distribution precisely to 2. A natural reply is that people don't form precise subjective beliefs and hence cannot communicate them. If one takes the Savage view of subjective beliefs, one must interpret this reply as saying that agents do not have precise preferences or, perhaps, that agents do not truly "know" their own preferences. If we think of preferences over, say, flavors of ice cream, this sounds ridiculous. If we think of preferences over

[9]Pierce [1902] seems to *define* vagueness in terms of unforeseen contingencies, saying "a proposition is vague when there are possible states of things concerning which it is *intrinsically uncertain* whether, had they been contemplated by the speaker, he would have regarded them as excluded or allowed by the proposition ... because these questions never did ... present themselves ..." " [italics in original].

[10]This idea bears more than a passing resemblance to the Grossman–Hart–Moore approach to incomplete contracts. See Hart [1995] on this approach and Dekel, Lipman, and Rustichini [1998] for a discussion of the connection between it and formal models of unforeseen contingencies.

state–contingent sequences of commodity bundles over one's lifetime, it seems obviously correct. In 1967, Savage described the problem as "the utter impracticality of knowing our own minds in the sense implied by the theory." He went on to comment

> You cannot be confident of having composed the ten word telegram that suits you best, though the list of possibilities is finite and vast numbers of possibilities can be eliminated immediately; after your best efforts, someone may suggest a clear improvement that had not occurred to you.

Put differently, the vastness of even very simple sets of options suggests it is ludicrous to think a real person would have well defined ideas, much less well behaved preferences, regarding the set.

In short, it is not that people have a precise view of the world but communicate it vaguely; instead, they have a vague view of the world. I know of no model which formalizes this. I think this is the real challenge posed by the question of my title.

References

[1] Crawford, V., and J. Sobel, "Strategic Information Transmission," *Econometrica*, 50, November 1982, 1431–1451.

[2] Dekel, E., B. Lipman, and A. Rustichini, "Recent Developments in Modeling Unforeseen Contingencies," *European Economic Review*, 42, 1998, 523–542.

[3] Dow, J., "Search Decisions with Limited Memory," *Review of Economic Studies*, 58, January 1991, 1–14.

[4] Hart, O., *Firms, Contracts, and Financial Structure*, Oxford: Clarendon Press, 1995.

[5] Keefe, R., and P. Smith, *Vagueness: A Reader*, Cambridge, MA: MIT Press, 1996.

[6] Monderer, D., and L. Shapley, "Potential Games," *Games and Economic Behavior*, 14, May 1996, 124–143.

[7] Savage, L. J., "Difficulties in the Theory of Personal Probability," *Philosophy of Science*, 34, 1967, 305–310.