Human Motivation: The Inadequacy of Economists' Models

Alfred F. MacKay

Follow this and additional works at: http://scholarlycommons.law.hofstra.edu/hlr
Part of the Law Commons

Recommended Citation
HUMAN MOTIVATION: THE INADEQUACY OF ECONOMISTS' MODELS

Alfred F. MacKay*

In The Book of Fallacies Jeremy Bentham cautions against what he calls the "Chinese Argument," which consists of appeals to the wisdom of ancestors:

From the facts of their times, much information may be derived—from the opinions, little or none. As to opinions, it is rather from those which were foolish, than from those which were well grounded, that any instruction can be derived. From foolish opinions comes foolish conduct; from the most foolish conduct, the severest disaster; and from the severest disaster, the most useful warning. It is from the folly, not from the wisdom of our ancestors, that we have so much to learn . . . .

In this article I argue that lawyers can learn something important from economists about human motivation, and, following Bentham, that it is from their folly rather than their wisdom "that we have so much to learn."

THE STANDARD PICTURE

A certain picture, or implicit background conception, underlies many economists' thinking about human motivation, which may be described rather simply. When one wants something, one has in mind a set of features, characteristics, or properties which together constitute the basis of the desire for the thing. In the case of wanting an automobile, for example, the set might include capacity, style, engine performance, efficiency, durability, resale potential, and price. Professor Kelvin Lancaster puts the point this way:

* Associate Professor of Philosophy, Oberlin College. B.A., 1960, Davidson College; Ph. D., 1967, University of North Carolina.


2. Id. at 401.


441
Goods, as such, are not the immediate objects of preference or utility or welfare, but have associated with them characteristics which are directly relevant to the consumer. . . . The consumer is assumed to have a preference ordering over the set of all possible characteristics vectors, and his aim is to attain his most desired bundle of characteristics subject to the constraints of the situation. The consumer's demand for goods arises from the fact that goods are required to obtain characteristics and is a derived demand.4

To desire something, then — or, more generally, to be motivated toward something — is to have desirously in mind this mental shopping list containing clusters of characteristics and, subject to situational constraints, seek whatever best or satisfactorily fills the bill. The most important feature of this picture is its generality — not so much its range of application (although the picture is intended to apply without exception and so be completely general in that sense) as its attribution of essential generality to the phenomenon of desire as such. Desire is always for whatever: whatever offers prospects of the best mix of listed characteristics; whatever happens to meet the requirements best. According to Lancaster, “[i]t is the characteristics in which consumers are interested. They possess preferences for collections of characteristics, and preferences for goods are indirect or derived in the sense that goods are required only in order to produce the characteristics.”5

Economists, of course, have specialized concerns, so it is worth noting that this picture offers them some definite advantages. For one thing, it explains aspects of the phenomenon of substitution,6 why margarine will typically be a substitute for butter but not for lumber. Margarine and butter, but not lumber, share many desire-relevant characteristics and, subject to price or income changes, one or the other will offer the best mix. If it is the cluster of table-spread characteristics rather than the particular goods toward which preference is aimed, and if desire is for whatever offers the best mix of said characteristics, then this substitution phenomenon is explained.

The theory also predicts how new products or services will fare in the marketplace. If desire for transportation is focused not on particular goods but on transport-relevant clusters of characteristics, we can estimate the demand for new, untried forms of transportation by

5. K. LANCASTER, supra note 3, at 7 (emphasis in original).
predicting how they will satisfy currently desired characteristics. Thus, we envision the reception of various currently unavailable, urban people-movers by projecting how they would rate in terms of convenience, reliability, comfort, speed, relative expense, and so forth. Notice that it is the generality of this picture—that desires are for whatever—that gives it the flexibility to accommodate things like product innovation, change, variation, and substitution. Desire is focused on characteristics, and characteristics are essentially general; they apply to and are exemplified by more than one thing.

This picture is not completely wrong; indeed it accurately describes a considerable range of motivational phenomena. But economists, as well as others influenced by this model, are mistaken in thinking that it accounts for all desire, all motivation. Common sense, of course, recognizes both general and specific desires. For general desires—for example, the desire not to remain unmarried for long; the desire to acquire some form of transportation—the cluster of desire-defining characteristics does not even appear to single out any particular, individual thing, and thus, attributing generality seems to follow. Specific desires, however, like the desire to marry Jerry Brown, or the desire to buy Secretariat, focus on unique, particular things, and the role of generality is less apparent. Still, in terms of the economic model under discussion here, some element of generality may apply even to specific desires, which are seen as being derived from general desires plus specifying beliefs. My purchase of this tub of Mrs. Filbert's margarine is seen as involving the general desire to attain the best mix of table-spread relevant characteristics, plus the specifying belief that this tub does the trick. From this, via what philosophers call a practical inference, we get the derived specific desire to buy this tub. This resulting desire, though, is thought of as implicitly trailing elements of its inferential ancestry. It somehow retains sufficient generality to account for my readiness to accept close substitutes if the price is right. Similarly, the desire to marry Jerry Brown is seen as deriving from some more general desire—perhaps, to associate with a famous, glamorous person—plus a specifying belief that Jerry Brown fills the bill.

Possibly some specific desires are thus inferentially derived and

---

7. See id. at 10.
8. See R. McKenzie & G. Tullock, The New World of Economics at x (rev. ed. 1978): "Economic method does not cover all phases of human life. Hence, when economics is applied to new areas . . . it does not give a complete picture."
possibly this inferred heritage of generality helps explain the readiness to accept substitutes, which is so pervasive in some contexts. The mistake I claim some economists make, and from which, a la Bentham, lawyers have much to learn, lies not in thinking that desire is sometimes this, but in thinking that all desire is like this — that this picture gives the very essence of human motivation.

Accordingly, I suggest direct desire, a competing picture of some human motivation. Direct desire focuses directly on a particular object, without intervening lists of characteristics or features in mind; it is fundamental and not derived; it operates without implicit prediction and assessment, without comparison and evaluation, and without potential substitutability. In short, direct desire is essentially particular, not general. Thus we have two pictures of human motivation to consider. One involves essential generality, implicit or explicit, and operates through clusters of characteristics as mediating conceptual mechanisms. The other is simple and direct, unmediated by conceptual structures with their attendant generality. Economists often tend to be monistic about motivation and think all desire is general. I hold, on the other hand, that human motivation is pluralistic, that at least both these pictures, neither reducible to the other, are necessary for an adequate model. Before sharpening this distinction, however, I suggest a useful comparison between theories of human motivation and theories of reference.

THEORIES OF REFERENCE

Philosophers have long been interested in the connection between language and reality. On the one hand, language is used mainly to talk about the world, although it can talk about language itself, as this remark does. On the other hand, language is itself in the world, a part of the world, a real thing. Still, it seems sensible to contrast language and world and contemplate the connection between them. How does language hook up with reality? A dictionary, especially an ideally complete dictionary that contained all words, would seem a gigantic, circular definition: words defined in terms of words, defined in terms of words. How does one break out of the circle of words and make contact with extra-linguistic reality? One way language seems to make contact with reality, maybe the only way, is through its mechanisms of reference: referring to things,
talking about them, speaking of them, mentioning them, calling them by name, designating or denoting them. Consequently, to inquire how language hooks up with reality is in large part to investigate reference.

Rationalists in these matters hold that all reference is achieved through meaning. An expression is about something, they hold, by virtue of what it means, and we learn what it is about by understanding what it means. The paradigm-referring expression on this view is the definite description, of which an example is “the winner of the 1948 Miss America competition.” This is a complex expression whose overall significance is a function of the significance of its structure and of its component parts. Changing “winner” to “winter” or “Miss America” to “International Rugby” changes the meaning of the resultant expression, and thus changes what the expression is about, what it refers to.

It is clear that definite descriptions operate in some such manner as this. What is controversial about rationalist accounts of reference — and hence about how language connects with reality — is the claim that all reference works this way. Proper names, for example, which do not appear to have content or meaning, are held by rationalists to “abbreviate” definite descriptions, or to be “disguised” definite descriptions, or anyhow to have sufficient “implicit” meaning to permit figuring out who or what they refer to. Thus, the name “Kenneth Arrow” is held really to mean (somehow) “the Nobel Prize-winning economist who proved an impossibility theorem of great significance for 20th-century social science.”

The intellectual father of this approach is Leibniz, whose idea of an individual concept is the archetype of all such views. Once he had noticed that the more content a concept contained the more determinate it became, and hence more narrowly applicable, it did not take an inventor of the calculus long to extrapolate that process to the limit and reach the idea of a concept so complete, so determinate, containing so many specifying constituent sub-concepts, that it applied to exactly one thing. Such a process could achieve particularity using only mechanisms of generality. Consider what we might

12. See id. at 172-73.
13. See id. at 174-78.
call "Leibnizian Twenty Questions."

"What am I thinking of?"
"Animal, mineral, or vegetable?"
"Animal."
"Human?"
"Yes."
"Living?"
"No."

The game continues down the line, adding more concepts, narrowing the focus—animal, human, not living, genius, inventor, diplomat, historian, linguist, mathematician, philosopher—with the aim of guessing the identity of the object in question. Leibniz' idea was to remove the guesswork altogether by adding enough concepts—each in itself containing content, characteristics, and hence generality, each in itself applying to many things, eventually reaching what he called an "individual concept," so narrowly focused, so super-determinate, that it applied to exactly one thing. Thus, from multiple, overlapping generality was supposed to come unique particularity.

This conception is called "rationalist" because rationalists share the conviction that the universe is completely intelligible, in principle containing nothing that cannot be understood. Given the ancient tradition equating intelligibility with generality, wherein the special function of the mind is to grasp universals, the motivation for this approach emerges quite clearly. The connection between language and world is so basic that it must be intelligible if anything else is to be. Hence all mechanisms of reference, even proper names, must operate through meaning and content, structures that can be understood. If a name, for example, worked without any content or meaning, there would be nothing intelligible about it, nothing for the mind to grasp, nothing to figure out. It would relate directly to its object, embodying a brute, arbitrary connection between word and world that would just have to be accepted as given. Such fundamental unintelligibility at the very heart of all thinking—the connection between language and reality—would be intolerable to the rationalist mind.

Empiricist theories of reference, on the other hand, are pluralis-

15. See id. at 41-43.
16. Id. at 19-20.
They grant that some referring expressions work the way rationalists claim all do, but they admit other kinds as well. In particular, they insist that some mechanisms of reference are brute, direct, given in experience, unmediated by conceptual content or meaning. They hold, in other words, that some linguistic expressions achieve reference to objects in ways that are importantly not understandable. This approach does not deprive us of seeing how reference works in such cases, but it does insist that at the core of such mechanisms lies at least one crucial link that is direct, given in experience. For this connection there is no reason, nothing to understand: It must just be accepted. Proper names, for example, are denied content or meaning; there is the name and the object named, and nothing is determined by hidden or implicit concepts.

Bertrand Russell was on both sides of this dispute. His celebrated theory of descriptions remains a monument of 20th-century rationalism; and his theory of logically proper names represents his empiricist heels digging in. Replying to Peter Strawson’s criticism of this latter doctrine, Russell says:

"What he objects to is the belief that there are words which are only significant because there is something that they mean, and if there were not this something, they would be empty noises, not words. For my part, I think that there must be such words if language is to have any relation to fact. . . . Unless fundamental words in the individual’s vocabulary had this kind of direct relation to fact, language in general would have no such relation."

More recently, Saul Kripke has urged considerations that appear to tip the balance in favor of empiricist pluralism. Kripke argues that only empiricist theories correctly capture our intuitions about the reference of proper names in counterfactual contexts, whereas
rationalist theories get those contexts all wrong.\textsuperscript{28} Proper names, according to rationalist theories, have associated clusters of descriptive content that determine the referent of the name.\textsuperscript{29} From this perspective, the name refers to whatever happens to satisfy uniquely the associated cluster of descriptions. But, Kripke argues, this account yields wrong predictions.\textsuperscript{30}

Let “Aristotle” mean (or have its reference determined by) the following cluster of descriptions: “the pupil of Plato and teacher of Alexander.” According to the rationalist view, the statement, “Suppose Aristotle had not gone into philosophy or pedagogy,” would not refer to Aristotle, and would probably be self-contradictory. Furthermore, if it should turn out that no student of Plato taught Alexander, then the name “Aristotle” does not denote anybody; and if somebody else, say Theophrastus, did both so study and so teach, then statements using the name “Aristotle” actually refer to Theophrastus. Kripke claims that all of these predictions are wrong.\textsuperscript{31} We intuit firmly that all of these counterfactual statements refer neither to someone else nor to nobody but to Aristotle.

Proper names, according to Kripke, are “rigid designators”\textsuperscript{32} that denote the same entity in all counterfactual situations as they do in actuality. A rigid designator is like a homing torpedo: Once locked on to its target it follows or tracks it no matter how it twists and turns through counterfactual circumstances. A non-rigid designator, on the other hand, — which is what definite descriptions typically are, according to Kripke\textsuperscript{33} — is more like a searchlight: It falls now on one thing, now on another, depending on circumstances. Thus, the winner of the 1980 United States presidential election was Ronald Reagan. But if things had gone differently with the Iranian hostages, the Cuban refugees, inflation, and productivity, the winner might have been Jimmy Carter. That is, the definite description, “the winner of the 1980 United States presidential election,” would have denoted Jimmy Carter, not Ronald Reagan. Thus, in different circumstances the same description will denote different objects, because, on the rationalist theory (which is correct for definite descriptions), the description denotes whatever happens to fit the require-

\textsuperscript{28} See id. at 91-97.
\textsuperscript{29} See id. at 96-97.
\textsuperscript{30} See id. at 28-34.
\textsuperscript{31} See id. at 61-63.
\textsuperscript{32} Id. at 48.
\textsuperscript{33} Id. at 59 n.22.
ments specified by its content. The implicit generality accounts for the shifting focus. Although in the actual world only one thing might fit the requirements, in different, counterfactual circumstances, various things might fill the bill.

By this point my comparison should be clear: The economists' conception of desire is like the rationalists' theory of reference. They both focus on particular objects only indirectly, when particular objects happen to satisfy certain content requirements. When they seek to attain particularity, they both use only mechanisms of generality—intermediary conceptual structures, clusters of characteristics, clusters of descriptions. They both attempt to give a monistic account of apparently pluralistic data — the one attributing generality to both general and specific desires, the other attributing meaning to both descriptions and names. They both have an easy time with one range of their data, and difficulty with the other.

The rationalist theory of reference is right and natural for definite descriptions but strained and awkward for proper names. The economists' theory of desire is right and natural for general desires (and some specific ones too). For what motivational phenomena is it strained and awkward? Economists in any case may want to see what they call "operational significance" attached to this discussion. What range of observable behavior resists their rationalist theory of motivation and calls for an empiricist account?

**Love and Loyalty**

A range of motivational phenomena has been discussed recently by philosophers, notably Michael Stocker\(^{34}\) and Andrew Oldenquist,\(^{35}\) that fairly cries out for an empiricist account. This area includes love (and self-love), loyalty (and various sorts of group identification, including patriotism), attachments of various kinds, caring, taking an interest in, being concerned with, and so forth. The area is not new to novelists and poets, nor to lawyers who deal with bequests, or divorce proceedings. Consider what Robert Nozick says about love:

> Love is an interesting instance of another relationship that is historical, in that (like justice) it depends upon what actually occurred. An adult may come to love another because of the other's characteristics; but it is the other person, and not the characteris-

---

tics, that is loved. The love is not transferable to someone else with the same characteristics, even to one who “scores” higher for these characteristics. And the love endures through changes of the characteristics that gave rise to it. One loves the particular person one actually encountered. Why love is historical, attaching to persons in this way and not to characteristics, is an interesting and puzzling question. 36

Andrew Oldenquist’s discussions of self-love and loyalties strike similar themes. 37 Egoism, according to Oldenquist, is wrongly construed in terms of impartial reasons, as if the egoist notices about himself a cluster of repeatable but uniquely specifying characteristics by virtue of which it is reasonable for him to promote his own welfare. 38 On the contrary, he argues, the motivational thrust of my self-love is focused directly on me, not mediated through any repeatable features or generalizable structures. I am interested in promoting my advancement, my happiness, my welfare, because it is mine. The motivation attaches to its object directly, not through general characteristics that the object is noticed to possess.

Egoism, Oldenquist thinks, involves one of a class of self-dependent judgments, 39 all of which appear to embody direct or empiricist motivation. Other examples are the judgments involved in parental affection (contrast “Save my child first!” with “Save the blond child first!”), group loyalties (my neighborhood, my profession), and patriotism. All these attach directly to their objects. I am not now loyal to my neighborhood because of such characteristics as location, amenities, and friendliness. If I were, I might be prepared to relocate upon discovering one better located or more cordial; but to move because I have found a neighborhood that offers a better mix of attractive characteristics would be precisely to be disloyal.

Michael Stocker’s discussion of love proceeds similarly:

[C]onsider those utilitarianisms which hold that an act is right, obligatory, or whatever if and only if it is optimific in regard to pleasure and pain (or weighted expectations of them). Such a view has it that the only good reason for acting is pleasure vs. pain, and thus should highly value love, friendship, affection, fellow feeling, and community. Suppose, now, you embody this utilitarian reason as your motive in your actions and thoughts toward someone.

37. See generally Oldenquist, supra note 35.
38. Id. at 26.
39. See id. at 32.
Whatever your relation to that person, it is necessarily not love (nor is it friendship, affection, fellow feeling, or community). The person you supposedly love engages your thought and action not for him/herself, but rather as a source of pleasure.40

What is lacking in these theories is simply — or not so simply — the person. For, love, friendship, affection, fellow feeling, and community all require that the other person be an essential part of what is valued. The person — not merely the person's general values nor even the person-qua-producer-or-possessor-of-general-values — must be valued.41

These philosophers make the common point that love and loyalty do not behave the way the rationalist theory predicts. Love and loyalty are sticky: Once attached to their object, they persist, resisting dislodgment. We speak of betrayal, defection, and treason in connection with these sorts of motivation, not merely of change, replacement, or substitution.

It may be that in the areas of economists' traditional concern the bulk of human motivation is adequately conceptualized by the rationalist picture.42 Economists, however, have been reluctant recently to stay within the confines of their traditional arena and have increasingly yielded to the temptation to stray off the reservation.43 Yet whatever the case with economists, lawyers who adopt their rationalist picture of human motivation — desire focused on clusters of characteristics; substitution, transferability, and tradeoffs the rule — are badly advised. The entire area of domestic relations, for example, because shot through with love, hate, loyalty, and betrayal, defies the rationalist picture. The peculiar intensities of motivation —

40. Stocker, supra note 34, at 458.
41. Id. at 459.
42. Even here, though, the notion of product loyalty and other sorts of loyalties make themselves felt. For instance, in occupational choice some people persist at their calling beyond what can be explained in terms of inertia and the pains of uncertainty and costs of retraining, even when faced with the perceived advantages of other lines of work. And savers notoriously leave their money in local, low-interest-bearing accounts, long beyond what rational behavior would predict. In these and other cases, of course, it is always possible to take refuge in the explanation of irrationality; but it seems preferable to recognize and make allowances for the impact of these various loyalties.
43. See, e.g., R. McKENZIE & G. TULLOCK, supra note 8, at 3: "In recent years, . . . economists have greatly expanded their field of concern, and, as a result, the boundaries of economics as a discipline are rapidly expanding outward, encroaching on areas of inquiry that have historically been the exclusive domain of other social sciences."
aims that persist long after their initiating circumstances are changed—these and other aspects of motivation in this area call for the direct, empiricist model.

In closing, I will try to shed some light on a difficulty noticed by Stocker.44 Following his point about the centrality to love of valuing the beloved, he continues:

Having said this, it must be acknowledged that there are many unclarities and difficulties in the notion of valuing a person, in the notion of a person-as-valuable. When we think about this—e.g., what and why we value—we seem driven either to omitting the person and ending up with a person-qua-producer-or-possession-of-general-values or with a person's general values, or to omitting them and ending up with a bare particular ego.45

I suspect the puzzlement here is generated by a false dichotomy, which may be resolved by appeal to our earlier comparison.46 In developing his empiricist account of reference, Kripke distinguishes between using descriptions to give the definition of a name, and using them to fix the reference.47 In the latter case, a name gets its reference initially established via descriptive mechanisms, but once the reference is thus fixed the name behaves as a rigid designator tracking the object so named, not shifting to whatever subsequently happens to fit the description. So, supposing “one yard” to be the name of a certain length, its reference was initially fixed as the distance when the arm of King Henry I of England was outstretched from the tip of his finger to his nose. According to Kripke, the name “one yard” is not synonymous with “the distance when the arm of King Henry I . . . .”48 That description was used merely to fix the reference of the name. Once that is done we can quite intelligibly say: “Had certain things happened to King Henry I in his youth, the distance between nosetip and fingertip would have been less (or more) than one yard.” The name, after getting its reference initially fixed, stays with the length, not with the description used to choose that length.

A similar maneuver seems available for Stocker's worry. Perhaps love, in its initial, formative stages, does attach to its object via noticed, attractive features. But once thus fixed, love tracks the be-

44. Stocker, supra note 34, at 460.
45. Id.
46. See text accompanying notes 28-33 supra.
47. See S. Kripke, supra note 27, at 43-48.
48. Id. at 54-55.
HUMAN MOTIVATION

loved, not whoever subsequently happens to satisfy the initiating cluster of characteristics. So we need not accept the implicit dichotomy Stocker seems to have in mind. He seems to be thinking along these lines: Either there is love which is reasonably, intelligibly initiated via attractive features, and thereafter consigned to shift its focus to whoever offers those features best; or love behaves as love should, remaining attached to its object through changes in characteristics — in sickness and in health, for richer and poorer — but at the price of being cut off from reasons, arbitrary of initiation and irrational of continuance.

If we follow Kripke, however, we can have it both ways: a rationalist beginning, and an empiricist (happy) ending. Of course, it might be objected that in neither case — Kripke’s reference-fixing via descriptive content, or my suggested love initiation via attractive characteristics — is the shift in mode from rationalist initiation to empiricist subsequent career really explained. It is described with whatever intuitive plausibility it has, but we are not told how or why it occurs. Still it does seem to occur, and that in itself is enough to show Stocker’s dichotomy false.

In conclusion, I should note that the economists’ picture of motivation has not gone entirely unchallenged. Among recent work questioning it, Professor A.K. Sen’s essay on “rational fools” is especially interesting. About 100 years ago Henry Sidgwick claimed that the Western tradition in ethics tended toward dualism — some thinkers operating with a goal-directed, maximization approach, others with a jural, law-based model. The former he traced back to Greek influence, the latter to Roman. This duality may appear in the distinction between teleological and deontological theories of ethics. What I have called the economists’ picture is obviously a descendant of the former, teleological approach. Most of Professor Sen’s criticisms seem to me to derive from the other, deontological approach.

This article suggests that a third model of human motivation is needed — an empiricist model that allows motivation to be connected to its object directly, not via content, clusters of characteris-

49. See text accompanying notes 31-33 supra.
50. See text accompanying notes 34-43 supra.
52. Sidgwick, Ethics, in 8 ENCYCLOPAEDIA BRITANNICA 574 (9th ed. 1878).
53. Id. at 574-75.
tics, or repeatable generality mechanisms. It also suggests that a combination of rationalist and empiricist models is possible in situations where motivation initially develops via clusters of characteristics but later attaches itself directly to the object, as is the case with love. Developing such a model, and exploring its implications extensively, remains a task for another occasion.