Paidéia, 2004, 14(27), 9-16

UTILITARISMO: UMA PERSPECTIVA PSICOFÍSICA

Lawrence Marks*
John Pierce Laboratory, Yale University - USA

Resumo: As doutrinas psicológicas do Empirismo, Associacionismo e Hedonismo serviram de fontes intelectuais para o desenvolvimento do Utilitarismo no século XVIII e da psicofísica no século XIX. O Utilitarismo, articulado primeiramente por Bentham em 1781, apresenta quatro pressupostos psicofísicos implícitos, embora importantes: 1) que utilidade, que reflete "benefício, vantagem, prazer, bem, ou felicidade", são conceitos quintessencialmente psicológicos; 2) que utilidades são quantitativas; 3) que utilidades são comensuráveis através de diferentes objetos; e 4) que utilidades são comensuráveis entre indivíduos. Embora as utilidades algumas vezes refletem a satisfação de necessidades biológicas, elas comumente representam valências ou valores psicológicos, cujas forças subjetivas podem elas mesmas derivar, dinamicamente, de processos de tomada de decisão.

Palavras-chave: Utilitarismo, psicofísica, filosofia, mensuração

UTILITARIANISM: A PSYCHOPHYSICAL PERSPECTIVE

Abstract: The psychological doctrines of empiricism, associationism, and hedonism served as intellectual sources for the development of utilitarianism in the 18th century and psychophysics in the 19th. Utilitarianism, first articulated by Bentham in 1781, makes four implicit but nevertheless important psychophysical assumptions: (1) that utilities, which reflect "benefit, advantage, pleasure, good or happiness," are quintessentially psychological concepts; (2) that utilities are quantitative; (3) that utilities are commensurable across different objects; and (4) that utilities are commensurable across individuals. Although utilities sometimes reflect the satisfaction of biological needs, they commonly represent psychological valences or values, whose subjective strengths may themselves derive, dynamically, from processes of decision-making.

Key-words: Utilitarianism, psychophysics, philosophy, measurement

Over the centuries, in weaning itself from its philosophical parents and emerging as an independent discipline, psychology established, developed, and elaborated a small number of core concepts and principles, three of which are notable. The principle of empiricism takes experience with the world as a primary means for shaping our minds and behaviors. The principle of essociation claims that spatial and temporal contiguities can account for the links between mental events, and between behavioral responses and stimuli. And the principle of hedonism regards pleasures and pains as central psychological forces that spur people to act, through biologically and culturally shaped systems of rewards and punishments.

Artigo recebido para publicação em 18/09/2003; aceito em 02/12/2003. *Endereço para correspondência: Lawrence E. Marks, John B. Pierce Laboratory 290 Congress Avenue, New Haven, CT 06519, USA E-mail: marks@jbpierce.org

The principles of empiricism, associationism, and hedonism merged in the 18th century, notably in the writings of psychologically minded philosophers, such as David Hume. Hume played an especially noteworthy role in the history of psychology, and, implicitly in the history of psychophysics. Building on the work of Hobbes, Locke, and Berkeley, Hume sought psychological answers to questions that had theretofore largely been treated from a "rational", or philosophical, perspective. To try to answer epistemological questions - How may we know whether a statement is true? How can knowledge be justified - Hume first reformulated the questions themselves, asking: How is knowledge acquired and constituted? Hume then answered by saying (to summarize crudely): through experience (empiricism), and by the workings of the mind (association, hedonism). The very title of Hume's major work, <u>A Treatise of Human Nature</u> (1739), indicates the thrust of his approach. To Hume, perceptions constitute the raw ingredients of mental life, and within less than a century, scientists would begin investigating the mechanisms, psychophysical and neural, that underlie perceptual experience.

Bentham's utilitarianism

Already by the late 18th century, the psychological principles of empiricism, associationism, and hedonism would find a comfortable haven, by way of Hume, in the writings of Jeremy Bentham, where these principles would form the backbone to his theory of utilitarianism. Utilitarianism, broadly construed, sought a scientific framework for political and moral decision-making (Bentham, 1781/1948). Utilitarianism aimed to do this by applying computational formulas to subjective (psychological, mental) quantities. From the perspective of utilitarianism, even the Aristotelian virtues, as I shall argue, may be translated into a set of psychological quantities – utilities, in Bentham's terms - that operate within appropriately formulated mathematical equations.

Bentham's (1781/1948) proposal is well known: "By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good or happiness ... or to prevent the happening of mischief, pain, evil or unhappiness to the party whose interest is considered" (p. 2). Utility, designated here as <u>U</u>, represents the psychological value or subjective worthiness of a stimulus object. To broaden the concept, <u>U</u> may be negative as well as positive, where a negative utility would refer to a property in an object that, to paraphrase Bentham, produces rather than ameliorates mischief, pain, evil or unhappiness. As a psychological entity, utility implicitly entails a psychophysics.

At the very least, Bentham's formulation requires that, for a given individual at any moment in time, we can rank order the values of \underline{U} associated with objects or properties that give "pleasure, good or happiness", such as wealth, gains or losses in income, and so forth. By ranking the relevant set of utilities (and assessing their probabilities, or subjective

probabilities), a decision-maker might then take whatever course of action leads to the greatest value of <u>U</u>. Of course, the strengths and weaknesses of various forms of utilitarianism have been debated for more than three centuries now, and it is not possible here even to outline the main ethical, logical, and psychological issues (Smart & Williams, 1973) – although I should note that several variants of utilitarianism are considerably more subtle than Bentham's. The present goal is not to enter the debate over the adequacy of utilitarianism as a philosophical, economic, or psychological theory, but instead to identify and consider the psychophysical implications of Bentham's theory.

Bentham's utilitarianism went well beyond the ranking of utilities. For Bentham wanted utilitarianism to contribute to social, economic, and political decision-making as well as individual judgment, and in order to accomplish this he had to assume that utilities are measurable quantities. Further, Bentham proposed to compute utilities not only for a given individual but also across individuals, for a society at large. Thus, utility to society as a whole could be computed from "the sum of the interests of the several members who compose it" (Bentham, 1781/1948, p. 2), providing a measure of the overall or net utility. Net utility, U*, may be written as

$$U^* = \Sigma U^+_{i,k} + \Sigma U^-_{i,k}, \qquad (1)$$

where \underline{U}^+ and \underline{U}^- are the positive and negative utilities associated, respectively, with objects \underline{i} and \underline{j} for individual \underline{k} . For the purposes of political decision-making, and in a spirit consistent with contemporary communitarianism, Bentham offered a rule for maximizing or optimizing. Simply stated, one should choose whichever option produces the greatest value of \underline{U}^* . "The greatest good for the greatest number", defined here as the maximal value of \underline{U}^* , became the credo of utilitarianism.

Psychophysical Consequences

Bentham's utilitarianism has several consequences that are relevant to psychophysics. To start, let me reiterate that utilities are psychological

$$U = F(W) \tag{2}$$

and

$$\delta^2 U/(\delta W)^2 < 0 \tag{3}$$

Given a rule of diminishing marginal gains, it may be possible to satisfy Bentham's principle of maximizing total utility U* by distributing total W uniformly over individuals. Note that this conclusion requires several auxiliary assumptions, for example, that any steps taken to spread W equally would not substantially decrease the total value of W, and that the psychophysical function F governing utility (and disutility) is uniform over individuals. Neither of these auxiliary assumptions is, of course, indubitable. First, it is likely that the very process of distributing wealth, or certainly of redistributing it, would modify people's behaviors in a way that might well affect the total quantity of \underline{W} to be distributed. And second, it is likely that the psychophysical functions do vary over individuals, an issue discussed later (Sen. 1973).

There is a related matter that cuts even more deeply to the heart of Bentham's utilitarianism. Equation 1 itself makes several tacit assumptions,

and an important one is that the measures of \underline{U}^+ and \underline{U}^- are commensurable over individuals (Harsanyi, 1955). This assumption does not require that people be psychophysically equivalent – that everyone's utility function obey the same form of Equation 2, for example – but simply that different individual utility functions, $\underline{F}_1, \underline{F}_2, \dots, \underline{F}_n$, characterize values of \underline{U} that have a common unit. What the principle of commensurability does say is that we must be able to perform arithmetic operations by combining Jill's pleasure with Jack's pain, that in the Benthamian Equation 1 we may freely sum utilities over the individuals.

If a utilitarian system is to be adequate, it have to address doubts regarding commensurability over individuals. Commensurability is a matter to which psychologists have given relatively short shrift – we seem most comfortable comparing behavioral measures, measures of performance, such as the heights of pole vault or standardized scores on Wechsler Adult Intelligence Scale, where the units of measurement come from physics or from counting correct and incorrect responses. And psychologists tend, quite properly, to be circumspect in what they say about inferred mental entities for which there are no "natural units" or whose constancy over individuals seems dubious. Despite the deep-seated reluctance to compare mental entities per se across different individuals, it is important to keep in mind that commensurability, like utility itself, has important practical as well as theoretical implications.

Leaving ontological considerations aside, a coherent system for comparing utilities across people is at least plausible. It is plausible because, in one domain of psychophysics at least, namely the perception of fatigue during dynamic physical work, laboratory and clinical research by (Borg, 1982) has provided a practical and widely applied methodology for interindividual comparison. Borg's system for measuring and comparing perceptions of fatigue is based on several explicit postulates. The most important of these states that, given some constraining conditions, different individuals working at their maximal physical capacity all experience approximately the same level of perceived exertion. It follows that, by measuring each person's

perception relative to his or her own maximum, the resulting measures become commensurable over individuals. Borg's scheme has been remarkably successful in predicting, for example, physiological responses to submaximal levels of physical exercise.

If it were generally the case that everyone experiences the same perceptual magnitude at the sensory maximum in a given domain, the problem of interindividual comparison would be readily resolved. Unfortunately, the domain of perceived exertion seems an exception rather than the rule. In my own laboratory, for instance, when presenting subjects with strong stimuli taken from two sensory domains (taste and hearing) we have often observed clear individual differences in relative response some subjects indicating that the taste sensations were markedly stronger than the auditory, other that the sounds were markedly stronger than the tastes. The implication is that the magnitude of the experiences at (extrapolated) maximum differs across individuals in at least one of the modalities, and perhaps in both. Measuring perceptions relative to maximum does not always ensure commensurability. Even if the domain of exertion is unique in this respect, however, it might be possible to start with measures of exertion and use these as leverage, or common currency, to measure commensurable quantities in other perceptual and cognitive domains, such as utility.

Even if the matter of commensurability across individuals can be resolved, there is still the matter of commensurability within the domain of utility itself. That is, it is possible that utility is a disjunctive concept - constituting a manifold of noncommensurable quantities. Recall Bentham's (1781/1948) definition – that utility comprises whatever confers "benefit, advantage, pleasure, good or happiness", a list that may not even be inclusive. It is not prima facie evident that "benefit", for example, is identical to "good", or that "pleasure" is identical to "happiness". To be sure, the term utility itself suggests a property of being useful, or beneficial (and hence to many connotes what is crass, making it in some ways a most unhappy choice of terms) – yet Bentham's principle is often called a maxim of greatest happiness or greatest good, as if there is sufficient equivalence in Bentham's list for utility to rest on some kind of common currency.

Elsewhere (Marks, 1992b), I've called this assumption of equivalence Plato's principle – for Plato was among the first to claim a possible unitary basis for deciding when actions are just, and Plato was among the first to claim, in the Protagoras (though elsewhere he would deny it) that "pleasure" mediates the "good". Furthermore, in that same dialogue, Plato argued that pleasure and pain are quantities, and thus are directly comparable. This argument has had a long history, pervading quarters of psychology as disparate as psychoanalysis and psychophysics (Marks, 1992a).

The matter of commensurability over objects, events, situations, is orthogonal to the matter, mentioned earlier, of commensurability over people. Even so, there are two ways to assert that objects are commensurable – a strong way and a weak way. The strong claim is Plato's, namely, that utility has a common currency – that pleasure, good, happiness, benefit, and whatever else share a common metric. A weaker claim is that various kinds of pleasure or good or happiness are, though qualitatively different, nevertheless commensurable and thus interchangeable.

It would be fruitless, if not impossible, to try to uphold the stronger claim without also maintaining the weaker one: If we cannot compare different kinds of happiness or pleasure or benefit, how are we to locate the Platonic "one" in the "many"? But viewed optimistically, happiness or utility is amenable to natural law, and in the utilitarian metric of Equation 1, every <u>i</u>, every <u>j</u>, every <u>k</u>, is a Newtonian particle, buffeted by the prevailing forces of socio-physical nature, if not quite "created equal" under the <u>ethos</u> or laws of the land then at least commensurable under the laws of natural philosophy.

The problems posed by commensurability become exaggerated the more sophisticated the utilitarian theory becomes. John Stuart Mill's (1863/1964) version, for example, while seeking to promulgate the virtue of virtue (so to speak), in order to clear the grounds for "higher pleasures" and thereby avoid the seemingly crude hedonism implied by Bentham, only exacerbates the problem of comparability. But commensurability may be unavoidable. For even when (or though) the demands,

desires, and obligations of the moment appear independent of one another, and thus qualitatively dissimilar, the process of choosing itself forces comparison. And by forcing comparison, choosing may thereby itself impose commensurability. In this regard, commensurability is intrinsic also to approaches to decision-making and ethical theories that are non-utilitarian, such as the neo-Kantian position of Rawls (1971) e Marks (1992b).

Utility, Value, and Valence

Psychology has shared part of its history with other disciplines, including economics, philosophy, and biology, and one consequence has been the development of cross-disciplinary connections, albeit fuzzy ones, among utilitarian theories in the domains of ethics, economic decision-making, and even neo-Darwinism. Some years ago, Cooper (1987) argued that theories of rational choice and utility could be derived from principles of evolutionary biology. In Cooper's view, utility is closely allied to the notion of fitness.

A decade ago, Cabanac (1992) recast Plato's claim in a contemporary biological context. Cabanac's goal is to zero in on the physiological systems or devices, behavioral and physiological, that mediate the ways that organisms choose their responses. By focusing on homeostatic mechanisms, Cabanac infers the "paramount importance of pleasure in the determination of behavior" (p. 173): an experimental confirmation, he asserts, of two millennia of philosophic speculation. Like Bentham and Mill, Cabanac argues that utility rests on pleasure, but he goes even further by claiming that pleasure can therefore serve as a theoretical bridge from biology to psychology, and by implication to related disciplines such as economics.

Although the data obtained from Cabanac's clever experimental designs - which produced behavioral/psychophysical/economic indifference curves – suggest that people behave as though they optimize something, and although his subjects report on their "pleasure" in ways that suggest that the internal mental/biological states stand in one-to-one correspondence with what they optimize, still, a skeptic may remain unconvinced of Cabanac's theoretical claims. It is a far cry from the conditions tested in his experiments – where Cabanac's subjects trade off physical work for changes in ambient temperature, or trade off greater sweetness for less sourness in a beverage - to the conclusion that different forms of pleasure are sufficiently commensurable to serve as the basis for a utilitarian theory of behavior. Consequently, a skeptic is likely to remain unconvinced that these data provide adequate support for what is ultimately a nonempirical claim, to wit, that "[human liberty] is the freedom to choose one's own way to maximize pleasure" (Cabanac, 1992, p. 197). When a person chooses among the alternatives that are available within a well-defined and strongly circumscribed set of conditions, that person's decisions, however systematic and well articulated in their structure, may nonetheless reflect contingencies that are unique to their contextual setting.

Cabanac (1992) like others implies that wealth, or a bundle of commodities, is a convenient substitute for something else – a stand-in for what is commonly called utility, which Cabanac characterizes in terms of pleasure, but which I would prefer to call valence, appropriating for this purpose a term used by Lewin (1936). Although Cabanac's stance serves a useful purpose in reminding us that biology undoubtedly plays an ineluctable role in economic and other kinds of behavior, nevertheless, I believe that his argument needs to be reconceptualized and reformulated. Most importantly, it needs to be rescued from an overly reductive physiological framework, and to do this it is critical to recognize that pleasures may, and commonly do, become valences or values, but that this happens only when biology becomes cognition, and thus it is critical to recognize that not all pleasures-that-become-values are easily derived and some do not derive at all - from biological needs.

Perhaps most importantly, in my view it may be necessary to conceptualize human behavior so that it is seen not as a hierarchically-organized system comprising subsets of well-defined and universal mental and behavioral processes, but instead as a looser set of capacities linked only weakly one with another, perhaps through heuristic strategies, which in turn are defined by and instantiated in the particular tasks in which they arise.

On the Contextuality of Valence

That all decisions are contextual is, no doubt, a truism in psychology. But my argument goes further: Contextual constraints may define the utilities themselves, and thus delimit the valences afforded by those outcomes that matter to us. Valences and values are construals and constructions. This position is related to, and sympathetic with, views that have been voiced by other psychologists, notably by Kahneman and Tversky (1979). Birnbaum, Mellers, and their colleagues (Birnbaum, Coffey, Mellers, & Weiss, 1992; Mellers, Ordónez, & Birnbaum, 1992) have also focused on contextual processes in decision-making. Their interpretation may seem to differ from mine, but I suspect that our stances are ultimately compatible – a matter of where in the mathematical equations to place weighting coefficients.

Biologically oriented thinkers, such as Cabanac (1992), commonly fail to make a distinction that I believe is crucial: Where needs are biological, values are social - and often cultural. Indeed, the very language in which we formulate and represent the "human condition", the terms through which we assay alternatives, by which we scale consequences, is a social construction and convention. To be sure, the capacity for language is biological, an evolved characteristic. But discourse has what philosophers call intentionality or meaning, and meaning, like utility, valence, and value, is psychological. We humans can assess situations requiring choice before they happen, and we can weigh the values of the objects, events, and conditions that are relevant to our choices. We are able to do this because we have evolved cognitive systems and strategies for representing knowledge. These systems enable us to predict outcomes, to imagine consequences, to hope for pleasures of the flesh and mind, and to fear pains that may afflict the body and the soul. Our mental and behavioral organs are highly flexible and adaptive. We can run on automatic pilot, as when we make implicit visual-motor decisions while, say, riding a bicycle or driving a car; and we can deliberate, as when we decide whether to change jobs or buy a home.

Because we often must choose among incompatible alternatives and competing obligations, we defer to – and perhaps in doing so define – the valence that applies to each. However biological their origins, valences are typically modulated or channeled through social practices or cultural experiences, as when we choose between red wine and white. How much more so, then, when Antigone had to choose between obedience to her uncle Creon and the social duty to bury her brother, or when Thomas More had to choose between obedience to his king and duty to his church! To be sure, many day-to-day decisions rest on pleasures based in biological needs; but even our biological needs, our drives for nourishment and for sex, filter through a cultural colander, unfolding through socially constrained or formulated valences or values. By this token, the common, Platonic view – that what is good or pleasurable is unitary – is erroneous, based on the misconception that utilities, valences, or values necessarily preexist or inhere in the mind's calculus prior to the act of deciding. Instead, it is often those very acts that delineate the valences or values, and thus confer commensurability. Perhaps there would be virtue to eliminating the essentialism that pervades the common view.

These considerations speak to one of the unhappy consequences of Bentham's brand of utilitarianism, as noted by Sen (1973, 1990), among others. As mentioned earlier, utilitarianism has been taken by some to endorse or underpin an equal distribution of resources, but it need not do so. Indeed, to the extent that people differ in their utilities, optimization may dictate a distribution of resources that exacerbates inequality rather than ameliorating it. Utilitarian schemes may become pernicious "when the mental characteristics of pleasure or desire adjust to situations of persistent inequality. In circumstances of long-standing disparity and inequity, the underdogs may come to regard their fate as fairly inescapable ... [and] ... learn to adjust their desires and pleasures accordingly..." (Sen, 1990, p. 51). This state of affairs has led Sen to affirm the priority of human freedom over utilities, and to focus on human capabilities.

These concerns may be ameliorated,

however, by taking utilitarianism as a framework for decision-making based more broadly upon valences and concomitant social values, for valences and values can depend on wide range of cultural, social, and behavioral norms and practices. Consequently, the utilitarian calculus could give heavy weight, to, say, opportunities to maximize individual potential, or more generally to Aristotelian virtues - this perhaps being accomplished more easily when utilitarianism is based in rules (rule utilitarianism) rather than acts (act utilitarianism). But any quantitative scale, or valuation, of the virtues would, of course, itself reflect their cultural and social priorities (MacIntyre, 1984; Nussbaum, 1990). Although the claim is often made that such an approach trivializes utilitarianism, the situation is quite the opposite: The approach suggested here highlights the importance of relating the psychophysical properties of utility theory to the psychological processes that give rise to the underlying valences that serve as psychophysical quantities. As suggested earlier, these valences may sometimes derive from the process of decisionmaking itself.

A deeper understanding of the mechanisms that underlie valences may be critical to political, social, and economic decision-making. Decisions are made all of the time, at both the "micro" and the "macro" level, by individuals, by corporations, and by governments. Many values and valences are learnable, and many surely are learned. To give one example, Eisenberger (1992) has investigated the "work ethic", mounting considerable evidence to show that "industriousness" is, or at least can be, conditioned through well-known principles of associative learning. More generally, valences and values arise within a framework of social and cultural practices, themselves transmitted through experience and learning. Not only can valences and values be learned, but surely they can also be taught. This message was not lost on John Dewey (1916), who recognized the singular role of educational reform as a means to achieve social progress.

References

Bentham, J. (1948). An introduction to the principles of morals and legislation. New York: Hafner

- (originally published, 1781).
- Birnbaum, M.H., Coffey, G., Mellers, B.A., & Weiss, R. (1992). Utility measurement: Configuralweight theory and the judge's point of view. Journal of Experimental Psychology: Human Perception and Performance, 18, 331-346.
- Borg, G. (1982). A category scale with ratio properties for intermodal and interindividual comparisons. In H.-G. Geissler & P. Petzold (Eds.), Psychophysical judgment and the process of perception (pp. 25-34). Berlin: VEB Deutscher Verlag der Wissenschaften.
- Cabanac, M. (1992). Pleasure: the common currency. Journal of Theoretical Biology, 155, 173-200.
- Cooper, W.S. (1987). Decision theory as a branch of evolutionary theory: A biological derivation of the Savage axioms. Psychological Review, 94, 395-411.
- Dewey, J. (1916). Education and life: An introduction to the philosophy of education. New York: Macmillan.
- Eisenberger, R. (1992). Learned industriousness. Psychological Review, 99, 248-267.
- Harsanyi, J.C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. Journal of Political Economy, 63, 309-321.
- Hume, D. (1739). A treatise of human nature: being an attempt to introduce the experimental method of reasoning into moral subjects. London: John Noon.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.
- Lewin, K. (1936). Principles of topological psychology. New York: McGraw-Hill.
- MacIntyre, A. (1984). After virtue (2nd edition). Notre Dame, Indiana: University of Notre Dame
- Marks, L.E. (1992a). Freud and Fechner, desire and energy, hermeneutics and psychophysics. In H.-

- G. Geissler, S.W. Link, & J.T. Townsend (Eds.), *Cognition, information processing, and psychophysics* (pp. 23-42). Hillsdale, NJ: Erlbaum.
- Marks, L.E. (1992b). Quantifying mental qualities: From utility to *ethos* and *polis*. In G. Borg & G. Neely (Eds.), *Fechner Day 92. Proceedings of the Eighth Annual Meeting of the International Society for Psychophysics* (pp. 129-134). Stockholm: ISP.
- Mellers, B.A., Ordónez, L., & Birnbaum, M.H. (1992). A change-of-process theory for contextual effects and preference reversals in risky decision making. *Organizational Behavior* and Human Decision Processes, 52, 331-369.
- Mill, J.S. (1964). Utilitarianism. In E.A. Burtt (Ed.), *The English philosophers from Bacon to Mill* (pp. 855-948). New York: Random House. (originally published, 1863).
- Nussbaum, M.C. (1990). *Love's knowledge*. New York: Oxford University Press.
- Plato. (1899). The dialogues. New York: Scribner's.
- Rawls, J. (1971). *A theory of justice*. Cambridge: Harvard University Press.
- Sen, A. (1973). *On economic inequality*. Oxford: Clarendon Press.
- Sen, A. (1990). Individual freedom as a social commitment. *New York Review of Books*, *37*,(June 14), 49-54.
- Smart, J.J.C., & Williams, B. (1973). *Utilitarianism* for and against. Cambridge: Cambridge University Press.
- Stevens, S.S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects.* New York: Wiley.