

Binder, Martin

Working Paper

Some considerations regarding the problem of multidimensional utility

Jena Economic Research Papers, No. 2009,099

Provided in Cooperation with:
Max Planck Institute of Economics

Suggested Citation: Binder, Martin (2009) : Some considerations regarding the problem of multidimensional utility, Jena Economic Research Papers, No. 2009,099, Friedrich Schiller University Jena and Max Planck Institute of Economics, Jena

This Version is available at:
<http://hdl.handle.net/10419/32569>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



JENA ECONOMIC RESEARCH PAPERS



2009 – 099

Some Considerations Regarding the Problem of Multidimensional Utility

by

M. Binder

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

© by the author.

Some Considerations Regarding the Problem of Multidimensional Utility[☆]

M. Binder^a

^a*Max Planck Institute of Economics, Evolutionary Economics Group, Kahlaische Str.10, 07745 Jena, Germany*

Abstract

The concept of ‘utility’ is often used in ambiguous ways in economics, from having substantive psychological connotations to being a formal placeholder representing a person’s preferences. In the accounts of the early utilitarians, it was a multidimensional measure that has been condensed during the marginalist revolution into the unidimensional measure we know today. But can we compare different pleasures? This paper assesses the evidence from psychology and neurosciences on how to best conceive of utility. It turns out that empirical evidence does not favor a view of multidimensional utility. This does not eliminate the possibility to make a normative argument supporting a multidimensional notion of utility.

Key words: utility, pleasures, neuroeconomics, multidimensionality of utility

JEL-classification: D87, B41, B12

1. Introduction

The concept of ‘utility’ is not an easy one. Not only is it a fuzzy notion which is being used with considerable ambiguity in economics (Broome, 1991b,a; Sen, 1991). It also has undergone major transformations in the way it is commonly used within the discipline (Stigler, 1950a,b; Bruni and Sugden, 2007): In ordinary English, utility means ‘usefulness’ but this usage is not what economists have in mind when using the term. In the economic context, originally, the notion of utility has been defined by Bentham who borrowed it from Hume (cf. Hume, 1994, 1740, Book III). For the classic utilitarians such as Bentham and Mill, the notion of utility had a substantive meaning that was clearly related to pleasures and pains (Sen, 1991, p. 279). Moreover, Bentham and Mill have understood utility to be an irreducibly multidimensional notion, of which the dimensions were the different types of pleasure and pain (Warke, 2000a,b,d).

During the marginalist revolution and in positivist economics afterwards, this conceptualization changed dramatically, when utility was purged of any psychological content and streamlined in order to allow for the formal analysis of marginal utility and constrained maximization (see Lewin, 1996; Bruni and Sugden, 2007). Today, the concept of utility is very far removed from the meaning it had in its beginnings. What Samuelson (1947, p. 91) describes

[☆]The author wants to express his gratitude for helpful comments and suggestions to the participants of the ‘8th Buchenbach Workshop for Evolutionary Economics’. Remaining errors are mine.

Email address: binder@econ.mpg.de (M. Binder)

as the “hollowness of utility” is the fact that utility has become an abstract magnitude that is only used as a placeholder for “*that which represents a person’s preferences.*” (Broome, 1991b, p. 3, italics in original) The concept that is fundamental and which (axiomatic) utility theory starts with are individual preferences. From these, one arrives at a utility function if preferences do obey some well-known standard textbook assumptions (Mas-Colell et al., 1995, chs.1&3).

It seems indisputable that this turn toward a formally specified but materially empty utility notion has considerably increased the “mathematical fitness” of the concept (Warke, 2000b). However, its inherent ambiguity has come at a certain cost. First, utility is used in economics in a dual role that is often conflated, namely as a motivator of action and as measuring rod for welfare (Sen, 1973, p. 253): An action is undertaken because of a forward-looking preference which is directed to the future (‘wanting’). The corresponding (expected) utility is motivating the action. But having acted, the individual derives a certain amount of utility from the action, the reward (‘liking’). Both notions are subsumed under the heading utility. While Berridge (1999) has shown that both wanting and liking often go hand in hand, they can be dissociated; for instance can we desire things and ultimately not like them, an interesting case for welfare analysis (see also Kahneman et al., 1997).

Second, and for the present paper more importantly, utility is nowadays usually understood to be a onedimensional concept. But what about the classical utilitarians’ idea of different pleasures? Is the utility we derive from reading one of Shakespeare’s sonnets the same utility which we derive from eating an opulent meal at a fine restaurant? Or to take this question one step further and put it in Mill’s words: Is it better to be a dissatisfied Socrates than a satisfied pig because some utility is of a higher quality than the other? We can see that a multidimensional utility concept would, on the one hand, create difficulties specifying determinate choice because when pleasures are qualitatively different, on what basis can one decide how to choose between them? On the other hand, the allusion to Mill’s dictum shows that the multidimensionality of utility also creates problems for the evaluation of individual well-being and social welfare since it is not clear how to differentiate between more and less (morally) valuable pleasures.

In this paper, I will argue that the question of such a multidimensionality of utility can be answered by recourse to recent findings from the behavioral and neurosciences, or what one could call a “naturalistic approach” to economics (Witt, 1987). I will start from the methodological premiss that the positivist purging of all substantive psychological content from economics has impoverished economic theorizing that now often relies on folk psychological intuition instead of well corroborated findings from the behavioral sciences. In this paper, the concept of utility shall be formulated therefore so as to be compatible with empirical findings from other behavioral sciences, as these can help in settling substantive economic disputes or aid in clarifying core economic concepts such as the utility notion (cf. Camerer et al., 2005; Braeutigam, 2005; Kenning and Plassmann, 2005). With this argument, the paper extends a discussion by Landreth and Bickle (2008), who have centered on a similar question regarding neural substrates of a forward-looking utility notion relevant for decision-making (‘wanting’). Based on the available evidence, I here argue that utility as measuring rod for welfare, i.e. ‘liking’, seems to be processed as unitary in terms of a common currency (McClure et al., 2004).

The paper is structured as follows: I will first state the problem of uni- versus multidi-

mensional utility and how it is usually dealt with in economics in section 2. In the main part of the paper I then discuss relevant evidence from neurosciences and psychology to evaluate whether utility is best understood as a multidimensional magnitude or not (section 3). Utility here is linked to how the brain processes reward. While utility might be plural at its sources, evidence points to the fact the utility is unidimensional as regards its neural substrate, i.e. there exists a common neural currency. While I maintain that the empirical evidence does not support a conceptualization of utility as multidimensional, this does not preclude that one could give normative arguments in favor of conceiving of utility as a multidimensional magnitude. I will discuss some difficulties regarding such an argument in section 4 before finishing with a conclusion (section 5).

2. Multidimensional Utility in Economics

In the economics of today, utility is conceptualized as a onedimensional variable derived from the consumption of different goods. That means, drinking water, enjoying one's fame or reading a stimulating novel all result in a comparable utility that can be evaluated on the same scale. But is this really the case? Contrary to the use in economics today, the classical utilitarians did seem to hold, as Warke (2000a,b,c) has persuasively shown, a multidimensional view of utility. And doubts regarding the comparability of utility resulting from different activities have also been succinctly summed up by Amartya Sen:

“How do we sum up, on the basis of some objective measures of intensities, the respective desires for an ice-cream, freedom from a headache, writing the most beautiful sonnet ever written, going to bed with one's favorite film star, and being morally impeccable?” (Sen, 1981, p. 200)

Prima facie intuition of the above sort against the idea that all pleasure should be similar and easily comparable is usually taken as one main counter-argument against hedonist accounts (Sumner, 1996, p. 90). Often, even proponents of hedonism do reject such ‘internalist’ versions of hedonism in favor of ‘externalist’ versions, where pleasures are allowed to differ while the common defining feature is a pro-attitude one has towards these qualitatively different pleasures.¹

But a potential multidimensionality of utility does pose difficulties on multiple levels. For example in the case of aggregating these dimensions into one measure of social welfare (when comparing two or more different economic states with respect to the utility profiles of the individuals involved), one encounters the index number problem, which describes the impossibility to arrive at a ranking of some states that differ in respect to more than one independent quality: Given n persons and m qualitatively different pleasures, one needs a matrix of $m * n$ intra- and interpersonal aggregation weights in order to arrive at a numerical representation of social welfare. Without these weights, it would be impossible to use such calculus except in trivial situations (such as those where differences in pleasures and pains

¹This type of refutation comes also by the name of “Brentano's cigar”, who asks us “how ridiculous it would be if someone said that the amount of pleasure he has in smoking a good cigar is such that, if it were multiplied by 127, or say by 1.077, it would be precisely equal to the amount of pleasure he has in listening to a symphony of Beethoven or in viewing one of Raphael's madonnas.” (Brentano, 1969; Feldman, 2004)

only relate to one type of pleasure) because there is no common natural denominator, no generic pleasure according to which acts can be ranked.

But let us put normative considerations aside and think about the positive side of an irreducible multidimensional utility notion. When contemplating which action to choose, there arises a severe problem for the specification of determinate choice. To exemplify, consider a person and her motivation to act in a multidimensional utility framework. When facing two possible actions A and B and the question whether to choose A or B , that person would be guided by the utility associated with action A and B , viz. $u(A)$ or $u(B)$. Let us assume equal duration, and exclude questions of uncertainty. The intensity of the pleasure associated with actions A or B still seems to be subject to considerable subjectivity. And in order to come to a conclusion whether to choose A or B as an action, an individual has to be able to trade-off different pleasures with each other and trade-off different pleasures with different pains. Suppose for example, the simplest case of only one type of pleasure and one type of pain related to both actions. Assuming cardinality, if A results in a pleasure of 5 units and a pain of -3 units and B in pleasure of 7 and pain of -6 , what action to choose? The individual would have to be able to specify how many units of pain are traded-off against one unit of pleasure (assume the ratio is one to one, then one could calculate $u(A) = 2$ and $u(B) = 1$ and the individual would prefer A over B . If such a trade-off cannot be specified, it is unclear how to order $u(A) = \{5, -3\}$ and $u(B) = \{7, -6\}$). If there exists the possibility that there are even more than two pleasure dimensions, one would have to have more weights to internally compare the components and thus calculate an utility index. Such an index number problem then is straightly transferred into the theory of consumer behavior, where commodities are bundles of qualitatively different pleasures and pains (Warke, 2000a, p. 375).

The question of whether different pleasures are comparable with respect to their intensity—even if only restricted to the intrapersonal context—thus creates a pronounced problem: If intensities of different pleasures were comparable, one would have a common scale to calculate trade-offs, otherwise different intrapersonal weights would be needed to make these qualitatively different pleasures comparable. Sen has already pointed out that depending on the degree of comparability one can reasonably assume, only partial orderings might be possibly arrived at (Sen, 1981, pp. 204-5): The minimum partial ranking would be the intersection of the rankings reflecting different pleasures, i.e. no weights are specified and the respective $u(\bullet)$ can only be ranked under restricted conditions, viz. $u(A)$ can only be ranked as higher than $u(B)$ if all elements of $u(A)$ are ranked higher than those of $u(B)$.² In the other extreme case, if it were possible to rank pleasures on a common scale, the different weights are given and a complete ordering (the weighted sum) is possible. Of course, intermediate cases are conceivable if one might specify ranges for the different weights. Finally, this problem of intrapersonal comparisons of pleasures does transfer straight to the interpersonal case (of which is abstracted here).

In a more systematic fashion, there are three different ways for the economist of dealing with the index number problem in order to be able to cope with the problem of determinate choice (Warke, 2000a). The first strategy is to conceive of pleasure and pain as sufficiently

²Setting aside complications if components are negative (the ranking would have to be reversed) or if pleasures are not monotonic (when more of a pleasure at a certain point would be evaluated negatively).

homogeneous to be lumped together. It could be argued that with each different type of pleasure and pain comes a feeling of enjoyment or satisfaction that is homogeneous across the types of pleasures and pains. If such a feeling of enjoyment were assumed and it could be argued to be sufficiently homogeneous, the index number problem would not arise and utility as an index number could be calculated in summing up the enjoyment of different activities. Then determinate choice would favor the action that maximizes this feeling, a complete ordering would be possible (note this pertains to the general possibility, practical measurement might complicate this issue).

A different strategy (2) would be the assumption of a state of ‘indifference equilibrium’ in which the individual is when contemplating over different actions. While the first strategy solves the problem of determinate consumer choice from bottom up in asserting the homogeneity of pleasures and pains, the assumption of an indifference equilibrium does so top down (Warke, 2000a, p. 382). The argument basically goes that the consumer chooses between bundles of goods in such a way that the pleasures resulting from a marginal increase of expenditure on any of the goods are equal, even if the pleasures are qualitatively different. To put it with the words of Alfred Marshall, who used this method:

“For instance the pleasures which two persons derive from smoking cannot be directly compared: nor can even those which the same person derives from it at different times. But if we find a man in doubt whether to spend a few pence on a cigar, or a cup of tea, or on riding home instead of walking home, then we may follow ordinary usage, and say that he expects from them equal pleasures.” (Marshall, 1925, p. 15)

With this kind of equilibrium as a starting point and the assumptions of reflexivity and continuity, again, a complete and transitive ordering over the set of alternatives can be arrived at. However, it is a strong assumption to put the consumer into such a state of indifference equilibrium all the time. Such an assumption rules out the idea of out-of-equilibrium consumer behavior and, as Warke succinctly puts it, it is at least questionable whether the chosen pepperoni/ham ratio on our pizza reflects our long-run indifference equilibrium regarding the pepperoni/ham ratio on a pizza (Warke, 2000a, p. 383).

The third strategy would be to impose *a priori* restrictions on an individual’s preference set. These restrictions are rationality assumptions that are chosen so as to imply that determinate choice will be possible. This is done for example in axiomatic (expected) utility theory, which starts with an individual’s preferences. From these, one arrives at a utility function if the preferences do obey some standard textbook assumptions such as completeness and transitivity (Mas-Colell et al., 1995, chs.1&3): If preferences satisfy the conditions of completeness and transitivity, they are called “rational”. This notion of rationality implies that an individual has a complete ranking of all commodities in commodity space X which means he has a complete specification of his preferences for even the most exotic bundles of commodities and this ranking must additionally be consistent in the sense of the transitivity requirement.³

³Sometimes the third condition of reflexivity is also added. Since it is implied by the completeness assumption, it is redundant however, cf. Mas-Colell et al. (1995), p. 6

If the preference relation is also assumed to be continuous on X , there exists a continuous utility function $u(x)$. This function assigns a numerical value to each element in X and ranks these according to the preferences. We say $u : X \rightarrow \mathbb{R}$ is a utility function that represents these preferences if $\forall x, y \in X : x \succeq y \leftrightarrow u(x) \geq u(y)$ holds (where x and y are commodity bundles from X and \succeq denotes a preference relation).

The drawback of the third strategy is that in such a framework, all connotations with sensory experiences or psychological qualities are lost in the technical meaning of the term utility. If there is some psychic experience that is actually measured and what it would be like is unclear. And of course, the rationality assumptions imposed on human preferences are far removed from empirical reality as the many well-known findings from behavioral economics have shown (Loomes and Taylor, 1992; Spash and Hanley, 1995; Knetsch, 1992). This lacuna also extends into the sphere of normative economics and social choice: Instead of the psychologically oriented calculus, today only subjective preference orderings are what utilitarian social philosophy is considering. What motivates people, and why their —only formally specified— preference ordering is the way it is, is left unexplained in this kind of economics.

Is there a possibility to decide which of the strategies to choose? Is utility best conceived of as uni- or multidimensional? While Warke argues in favor of a multidimensional concept (see Warke, 2000a, p. 373) on the grounds that none of the other strategies “commands empirical support”, he owes the reader the evidence. From a naturalistic point of view, such a question should not be decided in an *ad hoc* or armchair fashion based on introspection but based on sound empirical scientific evidence. The question of such a multidimensionality of utility can be answered by recourse to recent findings from the behavioral and neurosciences (what one could call a “naturalistic approach” to economics, cf. Witt, 1987). In order to decide the case, I will review some evidence from psychology and the neurosciences in the following section.

3. A Neuroeconomic Perspective on Utility

Neuroeconomics has recently been excessively promoted as being able to help economists in finally understanding human behavior with so far unknown scientific precision. Although many of such sweeping claims in neuroeconomics can rightly be seen critically (see McCabe, 2008; Harrison, 2008, for a discussion of related methodological issues and the imprecision still involved in this young sub-discipline), a plausible case can be made that neuroeconomic methods do provide additional data for economic theorizing in general and can help shed light on the underlying neural correlates of utility in our particular context (cf. also Jamison, 2008). Data gained in neuroeconomic experiments (or neuro-scientific data more general) thus provides the economist with evidence to evaluate competing economic theories, if not enabling the development of altogether new theories. Whether it finally will be possible to conduct measurements of individual well-being directly in the brain is not a matter that needs to be settled here. Nevertheless, linking the economic utility concept to its biological and neural correlates offers the behavioral economist the possibility to turn back to a more substantive utility notion, one that is not solely linked to choice behavior but to subjective well-being, understood as the hedonic evaluation of an individual’s actions.

In this section, I will therefore review a notion of utility that is in accordance with

empirical evidence from psychology and the neurosciences. Studies on reward processing have opened up the black box of the human brain and allow us to understand when an individual encounters a rewarding object (e.g. McClure et al., 2004). While I will not go into the details of brain imaging (e.g. PET, fMRI etc.), I will link the findings of brain sciences with those established results from psychology.

By way of *prolegomenon*, recall the vital distinction in terminology made in the introduction: When using the notion of utility so far, there has been involved a certain ambiguity as to its dual role as motivator of choice and as measuring rod for welfare. This double function is often confounded (Sen, 1973, p. 253). It is therefore worthwhile to spill some ink on this distinction. When facing a decision, choice is made on the basis of certain preferences and a forward looking calculation (or estimation) of expected utility is made. That is, the individual arguably decides on a course of action on the basis of his desires that are directed to the future. This has also been called ‘wanting’ (Berridge, 1996, 1999). In biology, this notion of wanting relates to appetite or incentive motivation. Having made the decision and taken an action, the individual then faces the outcomes of his decision, for example the consumption of a food item. Such an act of consumption or more broadly speaking the satisfaction of a desire then leads to a feeling of reward, a pleasurable experience (this feeling of reward has a clear biological role in homeostasis, see below). This pleasure aspect of utility, or in other words, the ‘liking’ aspect (the evaluative side, not the motivational side) is what I am interested in the following discussion. It is this notion of utility that plays a role for assessing individual well-being. Note that this pleasure aspect is an *ex post* evaluation or can also be the instantaneous evaluation of the moment. Contrary to wanting (or expected utility), it denotes the hedonic qualities of reward (of course, the forward looking motivational utility can be based on previous experiences of ‘liking’). Berridge has clearly shown that both wanting and liking often go hand in hand but can be dissociated. And as Knutson and Peterson (2005) have made clear, both processes have different neural correlates, a distinction that has not been made by Landreth and Bickle (2008): Opposed to their predominantly methodological investigation that implicitly revolves around the decision-based notion of utility, what I am looking into here is the biological and neurological representation of utility as pleasures, i.e. the liking aspect of utility.

Seeking rewards is one essential feature for the survival and reproduction of all animals. From a biological perspective, not the pursuit of pleasure and pain is the ultimate goal but reproductive success. Nevertheless, nature ingeniously has endowed animal and man alike with an elegant solution that favors these actions that are in general beneficial for reproductive success. This is achieved by making these actions rewarding for the organism. Rewards can be understood to be all those stimuli that positively reinforce behavior; rewards thus increase the probability of behavior (McClure et al., 2004, p. 260). There are many reinforcers that are genetically hardwired to be reinforcing. These are called primary reinforcers and include food, water, air, sexual stimuli etc. I will take the different innate reinforcers provided by Millenson (1967) and as discussed in Witt (2001) for granted. They are a well established *fact* in psychology (Damasio, 2003, pp. 131-2). Via associative learning, previously neutral stimuli can be learned to be reinforcing as well. For example, one can conjecture that the most prominent secondary reinforcer money is conditioned on other primary rewards. In principle, all learned reinforcers are conditioned on primary reinforcers (even if the chain of associations may sometimes be quite large and idiosyncratic; cf. Cabanac, 1979; McClure

et al., 2004).

Being biological products, humans have (perhaps: unfortunately) not evolved to be happy but to survive and reproduce (Camerer et al., 2005, p. 27). This is practically achieved by a process called homeostasis, a process that seeks to regulate bodily functions in reaction to changes in external and internal conditions. As such, they compensate for changes in environmental conditions. One can think of the regulation of body temperature, the regular intake of air, water and food as examples of homeostatic processes. Homeostatic processes consist of two subprocesses. First, the organism detects a deviation from ‘set points’ as regards the parameters that are necessary to allow continued functioning of the organism. For instance, the hot summer has increased our body temperature or the hunting for food has dehydrated the organism. Then, mechanisms set in to restore the balance condition in these monitored parameters (so to speak, bring back the organism into equilibrium; set points comprise mostly of parameter ranges, not sharp levels). The organism begins to sweat to reduce body temperature or becomes thirsty to motivate the replacement of lost fluids. These homeostatic processes mostly work without any deliberate intervention. The negative hedonic feeling that is associated with the deviation from set points, i.e. the deprivation of reinforcers discussed above, is nature’s way of motivating action of the organism to restore homeostatic balance. In that sense, sensory pleasure is a sign for the presence of a useful stimulus (Cabanac, 1979). Closely related, and complicating the account a little more, is the concept of “alliesthesia” (greek for: ‘changed sensation’, see Cabanac, 1979), meaning that a stimulus can be perceived to be pleasant or unpleasant depending on the inner state of the organism. While deprivation is the negative motivator for action, alliesthesia is the positive one, making stimuli more rewarding when an organism is deprived of them. In that sense, pleasure and pain are “sovereign masters” (Bentham) motivating human action (see also Cabanac, 1992).

What I want to review in this section, is however, how these rewards are mapped onto brain functions. There have been established certain brain regions in which rewards are processed (these regions differ from the regions where ‘wanting’ or ‘expected utility’ is processed). The idea is now to get a deeper understanding of how reward is processed in the brain to understand the neural substrates of (experienced) utility. As Camerer et al. (2005, p. 10, their italics) show, “the study of the brain and nervous system *is* beginning to allow direct measurement of thoughts and feelings”. Especially the neurosciences have furthered our understanding of what happens when an individual encounters a natural or conditioned reinforcer. In this section, I review evidence drawing mainly from the literature of Kahneman et al. (1997), Kahneman et al. (1999), Berridge (1996, 1999), Shizgal and Conover (1996), Shizgal (1999) and others regarding where and how reward is processed in the brain. Although there is diversity at the level of different reinforcers, the rewarding experience (i.e. the ‘pleasure’) that is associated with the removal of deprivation is largely a unitary one as these authors show.

Before expanding on the hypothesis that reward is always processed “along a single final common pathway” (McClure et al., 2004, p. 261) and translated into one single “common currency” (McFarland and Sibly, 1975), let us shortly get a picture on the brain areas involved in reward processing. While different tasks activate different brain regions, it can be taken as a quite well established fact that rewarding stimuli are consistently increasing activity in the regions of the orbitofrontal cortex (OFC), the amygdala and the ventral

striatum/nucleus accumbens (NAcc).⁴ Note that ‘activated’ brain regions consume more oxygen. From increased blood flow in these regions, one thus infers increased neural activity. This is measured non-invasively through functional magnetic resonance imaging/tomography (fMRI). The measures taken thus do not report absolute levels but relative increases in blood flow (Kenning and Plassmann, 2005, p. 345).

To establish the role of these regions in brain processing, a wide array of brain imaging experiments have been conducted with a host of different stimuli that were presented, such as food, juice, water (e.g. Berns et al., 2001; O’Doherty et al., 2002), smells (e.g. Gottfried et al., 2002), sexual stimuli (e.g. Arnou et al., 2002), sexual behavior (e.g. Komisaruk et al., 2004), conditioned rewards such as money and positive feedbacks (e.g. Delgado et al., 2000; O’Doherty et al., 2001), but also abstract conditioned rewards such as light flashes (e.g. Pagnoni et al., 2002) and social rewards (e.g. Aharon et al., 2001; Rilling et al., 2002; Rolls et al., 2003). Although the functional mapping onto the OFC-Amygdala-NAcc complex is crude in terms exact location, it seems fairly well established that reward processing (evaluation) in general takes place in this region.

While the OFC has been found to receive direct input from primary taste and olfactory cortices as well as from higher-order visual and somatosensory areas (Rolls, 2000) and is thus an ideal place for the storage of reward values of sensory stimuli (McClure et al., 2004, pp. 260-1), the amygdala seems involved in processing the intensity of positive and negative stimuli (Anderson et al., 2003). The ventral striatum also plays an important role regarding rewards because of its connection with the mesencephalic (midbrain) dopamine system: The release of dopamine from the ventral striatum on the NAcc is experienced as highly rewarding (Johnston, 1999, p. 116).

“Output from the limbic system, the pleasure pathway [i.e. the MFB, the medial forebrain bundle, the pathway from striatum to NAcc], appears to be the common neurophysiological correlate of hedonic tone, and this output establishes the important link between feelings and learning.” (Johnston, 1999, p. 116)

As such, this pathway plays an important role in assessing reward and pleasure:

“Incredibly, the release of dopamine onto the nucleus accumbens has now been established to underlie almost every form of pleasure that animals can experience. Blocking the action of dopamine at the nucleus accumbens will [...] block the pleasurable effects of a host of natural rewards, such as food and water and sex. In addition, it has now been established that all addictive drugs [...] either directly or indirectly activate the pleasure pathway and eventually release dopamine onto the nucleus accumbens. The release of dopamine onto the nucleus accumbens, then, appears to underly all of our reward feelings.” (Johnston, 1999, p. 116)

Shizgal (1999) discusses evidence in research on brain stimulation reward (BSR). BSR means basically that rats receive direct electrical stimulation in the brain reward centers. This

⁴Cf. Rolls (2007) for a more detailed discussion of these brain areas. Descriptions of the location and functions of the OFC can be found in Rolls (2000) and of the amygdala and the ventral striatum circuits in Everitt et al. (1999) and Delgado (2007). The OFC is approximately in Brodmann areas 45, 47, while the amygdala can be found in areas 28, 34 and the NAcc/ventral striatum is located in area 34.

stimulation differs from a natural situation in that no deprivation of physiological resources is required (*ibid.*, p. 501) but it generally simulates natural reward very well (see also Conover and Shizgal, 1994a,b). It does so in driving up what Kahneman et al. (1997) have called “instant utility”. Instant utility is the continuous evaluation of our sensory experience over time. It has two aspects, namely a hedonic one (positive or negative valence; affective evaluation) and an action component (insofar it biases or impels action; there is a pro-approach bias for positive evaluation and an avoidance bias for negative evaluation). Now, direct stimulation of brain areas in rats⁵ is subject to problems of precision. Can one give evidence how reward is processed as regards natural reinforcers? The answer is yes.

While the neural activities are still partly mysterious to science, reinforcer-related behavior is increasingly well-researched. I will stick to the model by Shizgal (1999) of what happens when a so called ‘goal object’ (a reinforcer) is encountered. Three different brain processes work here in parallel: First there is perceptual processing to discern the identity, location and physical properties of the object. Then there is evaluative processing which determines goal worth. And third, a stop-watch timer process computes when the goal object will be available and how often this will be the case. The latter two processes are then the basis to calculate the payoff of the object (i.e. the reward or utility).

Note that regarding the two processes, there is a certain multidimensionality of the ‘payoff record’ of a reinforcer because reward depends on intensity, amount, time of occurrence, kind of reinforcer etc. But this multidimensionality differs from one that would assert that the pleasures resulting from different reinforcers differ regarding their qualities. Recall that reward processing itself has been found to take place in certain regions of the brain:

“A large body of evidence implicates specific regions in reward processing, including midbrain dopaminergic nuclei and target areas such as the striatum, as well as orbitofrontal cortex and amygdala (see Everitt et al., 1999; Rolls, 2000; Schultz et al., 2000). Two distinct components to reward processing are an anticipatory component, often signaled by the presentation of a cue which reliably signals the subsequent delivery of reward, and a consummatory component relating to reward receipt (Berridge, 1996).” (O’Doherty et al., 2002, p. 815)

These areas are frequently cited for the computation of reward. But if stimuli differ regarding time, reward intensity etc., there is the question how the organism computes the reward and adjusts behavior. The information contained in the multidimensional payoff record is condensed into a common internal unitary valuation by means of a “currency function” that expresses the value of the goal object on a common scale (Shizgal, 1999, p. 509). The same is argued by (Montague and Berns, 2002, p. 279): “Together, these results suggest that the OFS [orbitofrontal cortex and striatum] circuits act to generate a common internal currency (scale) for the valuation of payoffs, losses, and their proxies (predictors of payoffs and losses; ...).” Note that

“[w]ithout internal currencies in the nervous system, a creature would be unable to assess the relative value of different events like drinking water, smelling food,

⁵Or for that matter, direct stimulation of brain areas in humans as in the experiments of Heath (1964), whose subjects reported pleasure when stimulated in that way (cited from Berns, 2005).

scanning for predators, sitting quietly in the sun, and so forth. To decide on an appropriate behavior, the nervous system must estimate the value of each of these potential actions, convert it to a common scale, and use this scale to determine a course of action. This idea of a common scale can also be used to value both predictors and rewards.” (ibid., p. 276)

One can argue that this internal condensation of a multidimensional profile of the payoff record results in the continuous calculation of “instant utility” (Kahneman et al., 1997; Shizgal, 1999).

Summing up, a case can be made for the following view: Reward is processed in some fixed brain regions. The brain converts multidimensional payoff records into one common (neural) currency which is in turn the basis of our decisions but also of our evaluation of goal worth. Were this not the case, we would be the proverbial Buridan’s ass, not being able to decide between sonnet and food. The internal currency can be conjectured to be the feeling of reward that is associated with goal objects. There is no evidence that this reward (which we would term ‘utility’ in economics) is multidimensional as regards its neural correlates. This lends support to the hypothesis that reward is also best understood to be unidimensional. Consequentially, this comes to arguing that while of course there is a difference in the feeling when one eats something to remove hunger or one drinks something to remove thirst, the evaluation in terms of hedonic tone is of the same quality.

4. Some Speculations on the Normative Relevance of the Findings

From an empirical point of view, I think it is fair to conclude that the evidence for a truly multidimensional utility concept is lacking. At least internally, the brain is very able to compute a single currency of pleasure (through an internal value function) and to trade-off different goal objects regarding their ‘worth’. From a naturalistic perspective, it turns out that Warke’s strategy number one commands the most empirical support. Overall, I think that empirical findings do not lend much credibility to a true “vector view” (Sen, 1981) of the nature of utility. However, while conceiving of utility as ultimately onedimensional concept, this does not preclude disaggregating utility as regards the different types of reinforcers in more concrete examinations.

Note that although there are many studies that relate to how subjective reward is modified by many components of a decision task (e.g. grade of deprivation, expectancy of future reward, ...), I find no evidence in favor of a real multidimensional construct of reward regarding the pleasure component; one could say, as do Shizgal and Conover (1996) and Shizgal (1999), that utility is multidimensional in the sense that $u(x) = f(\textit{pleasure}, \textit{grade of deprivation}, \textit{time of next expected reinforcement}, \dots)$. Certainly, such a sort of multidimensionality is not the one of qualitatively different pleasures and pains implied by Sen’s quote in section 2. But do individuals systematically err when they attribute qualitatively different kinds of pleasure to such distinct events as they were exemplified above? It could be conjectured that since different neural processes (beside the reward process) are associated with the different enjoyable situations, it might well be the case that humans are ‘fooled’ into believing that the qualities of pleasure involved in these different acts are different. But maybe the difference in perception (sensory inputs) is different while the associated hedonic value nevertheless ranges on the same scale.

Moreover, conceptually speaking, the account depicted above lends credibility to an internalist account of hedonism that has been recently suggested by Crisp (2006a). According to his internalist pluralist account of hedonism, there is a plurality of different feeling tones of pleasure but the unifying element is that

‘... there *is* a way that enjoyable experiences feel: they feel enjoyable. ... Enjoyment, then, is best understood using the determinable-determinate distinction, and the mistake in the heterogeneity argument is that it considers only determinates. Enjoyable experiences do differ from one another, and are often gratifying, welcomed by their subject, favored, and indeed desired. But there is a certain common quality —feeling good— which any externalist account must ignore.” (Crisp, 2006a, p. 628)

Such an interpretation is plausible to the extent that the common factor in the diverse enjoyable experiences cannot be understood as a sensation or feeling that is in a certain body part. While ‘feeling good’ is not a particular kind of determinate feeling, it is a determinable. Such an account of enjoyment is not far away from our common everyday use of the concept. One can ask an individual whether he enjoyed a particular experience, be that strolling through the park, reading a novel or consuming a piece of cake. Moreover, the individual can give an ordering after he enjoyed any of these experiences, if he had them. The interesting point about such an internalist interpretation of hedonism is that it receives support from the argument I have given on the neural representation of utility in section 3: A unitary brain representation of reward can be argued to be the correlate of the determinable quality of ‘feeling good’. By way of this argument, it is possible to preserve the appeal of the internalist account of hedonism (namely its link to pleasures ‘feeling good’, i.e. being rewarding) without being vulnerable to objections from heterogeneous pleasures.

Of course, the whole discussion so far is notwithstanding the fact that one could make a normative argument why one should conceive of utility (or well-being) as multidimensional. One argument in favor of a multidimensional utility construct from a normative point of view can be found exactly in the Senian quote provided in the beginning. If we introspectively experience utility of such diverse actions to be unrelated and qualitatively different, should this intuition not be taken into account? A view that argues that utility is unidimensional here runs risk on imposing paternalistically a utility notion that runs counter to common (normative) intuitions. This is the standard objection that is advanced against typical internalist accounts of pleasure (Sumner, 1996, pp. 92-98).

This argument has been refined in terms of a thought experiment (see also Crisp, 2006b, pp. 112-7, for an extensive treatment). If pleasure is homogeneous and ranked only according to intensity and duration, the hedonist is committed to swallowing the following bitter pill: Consider that you could choose between living the life of the composer Joseph Haydn (having a duration of 77 years, with all the ups and downs of a creative artist’s life) or living the life of an oyster that is subjected continuously to a very small pleasure. Imagine that the duration of the oyster life is successively extended. Then at some point, the sum of pleasures of the oyster would outweigh the sum of pleasures of Haydn’s life. Would not the hedonist be committed to the absurd conclusion of choosing the life of the oyster? On the other hand, if we acknowledge for higher pleasures, then one could claim that there is some discontinuity between the

pleasures contained in Haydn's life and the ones in the life of the oyster. According to such a view, no matter how much oyster pleasure one could get, it would be always less worth than the higher pleasures involved in the life of Haydn. How can a unitary account of utility possibly cope with this counter-argument?

Such a type of refutation can be qualified. First of all, the thought experiment is one which is not exactly easy to fantasize about. How exactly do we envision, say, 2000 thousand years of oyster life being continuously pleasant in a low sort of pleasure. Would that amount to the mild pleasure of being tickled all these 2000 years? Can we really imagine how it is to experience a life of 2000 years of that pleasure (keeping in mind that we do not adapt to the pleasure). Part of the experiment is the idea that the pleasure does not lose its appeal. It is as strong and pleasant in the beginning as in the end. If we enjoyed the first second, we will similarly enjoy all the remaining 63 billion seconds of the experience following that second. Such an imagination taxes our faculty of correctly assessing the comparison. Schreiber and Kahneman (2000) have shown that when forming these kinds of aggregative judgements, we tend to use heuristics like the so called "peak-end-average" (the hedonic evaluation of this episode would be the average of peak and end pleasure) resulting in duration neglect. These findings are usually not mentioned in the oyster example. Of course, when assessing the oyster life according to such a heuristic, the resulting peak-end-average will be pitifully low, therefore leading to an unwarranted bias in the intuitive judgement. I think this is a prime example of how a naturalistic approach can also enrich normative theorizing: The thought experiment draws its force from a faulty aggregation humans employ in assessing periods of pleasure and pain. I submit that this explains why our intuition tells us that the oyster life is unattractive.

Ryberg (2002) has given arguments that cast further doubt on the thought experiment. While our preference for Haydn's life may be clear, a restatement of the problem can challenge this idea. Consider a 'higher pleasure' such as listening to Bach's organ works on the one hand, and consider a 'lower pleasure', maybe eating an apple on the other hand. Now imagine how many apples it would take to make the pleasure equivalent. Many would say no (finite) amount of apples would suffice. But what about some other higher pleasures, which are ranked not so high in enjoyment but are longer in duration. It would certainly be easier to trade-off half an hour of Bach's organ works against, maybe, one hour of the violin concertos of Schubert. And these maybe against seeing two hours of Shakespeare at a theater. And that maybe against a longer (three hours) fine meal. And so on. Giving some more examples, we would probably arrive at the least high of the higher pleasures and the highest of the lower pleasures. Would one not be willing to sacrifice then the higher pleasure for a much longer lower pleasure? Such a chain of trade-offs would then show (against our *prima facie* intuition) that there is a continuum of pleasure between the highest and the lowest pleasure and that there is no discontinuity involved.

What Ryberg here shows is that the test of whether we would prefer Haydn to the oyster may be an example where there seems to be no trade-off possible, but that there are many other cases where a trade-off indeed is possible. Therefore the objection is weakened. It is just not so clear cut anymore. Ryberg has refined the argument further with introducing probabilities in the calculus, for example getting the high pleasure only with a very small probability (say listening to Bach with $p = 0.00000001$ but the small pleasure with very high probability (having the apple with $p = 0.99999999$). Here our preference might be reversed

and many of us would probably stick with the lower pleasure. One could even add pain to the example. Imagine a life filled with many lower pains and a small chance of achieving the higher pleasure. On the other hand, imagine a life filled with many lower pleasures. According to the view that higher pleasures are discontinuously more valuable, one should choose the first alternative. But would our intuition still suggest that to be the case in the example?! Based on these counter-examples, it is fair to conclude that the Haydn-oyster objection is not as impeccable as it *prima facie* seemed.

Lastly, not only from a prudential, but also from a moral point of view, the idea that some pleasures should be ranked higher (and thus be morally more relevant) than other pleasures can constitute a powerful argument in favor of a multidimensional view of utility as well. This argument is exemplified in Mill's dictum to better be Socrates dissatisfied instead of a pig satisfied (Mill, 1863, ch.2). A "vector view of utility" then seems to be able to incorporate different classes of moralities as Sen has argued and it would enable the welfare economist to differentially weigh different pleasures according to moral worth (Sen, 1981, p. 193). However, the question that results is of course, on what grounds to distinguish between different utility components. One would have to come up with a *normative* weight scale to trade-off different utility components. Whether there can be made an impeccable argument how this scale would have to look like (equal weighting or discounting of some pleasures and not of others, etc.) is not clear and can be doubted (see also Sen, 1981).

The weighing problem thus translates to any utilitarian prescriptive theory of the right, in which the rightness of actions is determined through a felific calculus that forms a net total of pleasures over pains. While the standard utilitarian calculus suffers from the high rationality and perfect foresight requirement on the individual or social planner who would have to calculate the future expected consequences of an action (in terms of utility), this problem were even more pronounced when utility were assumed to be irreducibly multidimensional.

Moreover, if one were to base a welfare measure on a multidimensional concept of utility (i.e. assuming multidimensional welfarism such as the position discussed by Blackorby et al., 2005), it also becomes crucial what informational requirements the underlying utility notion can meet. This has been formalized in a social choice theoretic setting: Similar to the impossibility result obtained by Arrow (1963) for a unidimensional utility concept and a welfare aggregation based thereon, List (2004) has shown that in a multidimensional framework, one would have to meet certain informational requirements regarding "interdimensional comparability" (List, 2004, p. 120) in order to avoid an Arrovian impossibility result for that case: Especially interdimensional non-comparability leads to an Arrovian impossibility result in multidimensional utility space. Thus, from a normative point of view, the stipulation of various dimensions of utility that are non-comparable cannot be judged to be attractive. Not being able to meet some more demanding requirements on interdimensional comparability would thus constitute another normatively relevant objection (for the welfarist) against a multidimensional utility framework.

Summing up this discussion, I am not sure whether it's attractive to make a claim that utility on the conceptual level should be conceived as multidimensional for normative reasons. In accordance with the argument and evidence given in this paper, a welfare measure could be built based on the notion of "experienced utility" (Kahneman et al., 1997), i.e. on the idea that pleasure is a plausible normative maximand, not preference (Binder, 2009). It could be used to solve the problems associated with the preference satisfaction view such as the fact

that many preferences being satisfied do actually not increase well-being (e.g. because they were uninformed or biased). Going one step back to a unitary utility notion that retains its reference to pleasure and pain such as in the framework above has thus interesting normative ramifications and might have the potential to put the focus of welfare economics on that level where welfare should really lie, viz. at the level of pleasure. Indeed, basing a notion of welfare on these constructs seems to even allow for the measurement of pleasures on a common scale, as Kahneman et al. (1997) have shown. I agree with Feldman (2004) that remaining practical measurement problems cannot be a challenge to the idea of a hedonist account of utility or well-being on the conceptual level.

5. Conclusion

The concept of ‘utility’ was never an easy one and it has undergone major changes in its history. In Bentham’s and Mill’s classical utilitarian accounts, utility was a multidimensional measure clearly related to many qualitatively different pleasures and pains. During the marginalist revolution and in the heyday of positivism, this notion has been condensed into the onedimensional measure known to economists today in order to increase the “mathematical fitness” (Warke, 2000b) of the theory and to remove any psychological content which would not belong to the domain of economics proper (Bruni and Sugden, 2007).

This paper discussed the problems that a multidimensional utility measure imposes for a theory of determinate choice and for the assessment of welfare. In order to decide not on *ad hoc* methodological grounds whether the utility notion should be one- or multidimensional, I have assessed the evidence from psychology and neurosciences on how to best conceive of utility. It turned out that empirical evidence does not favor a view of multidimensional utility. While this does not eliminate the possibility to make a normative argument supporting a multidimensional notion of utility, I have highlighted some problems that such an argument would face. Implementing the results from other behavioral sciences (the naturalistic method) thus yields fruitful impetus for economic research, and in a case such as this one, it can even confirm conventional economic wisdom: Utility can best remain onedimensional. Economists rejoice!

Date: June 8, 2009; 8036 words

References

- Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., and Breiter, H. C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, 32:537–551.
- Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D. E., and Sobel, N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6(2):196–202.
- Arnow, B. A., Desmond, J. E., Banner, L. L., Glover, G. H., Solomon, A., Polan, M. L., Lue, T. F., and Atlas, S. W. (2002). Brain activation and sexual arousal in healthy, heterosexual males. *Brain*, 125:1014–1023.
- Arrow, K. J. (1963). *Social Choice and Individual Values*. Yale University Press, New Haven and London, 2nd. edition.
- Berns, G. (2005). *Satisfaction - The Science of Finding True Fulfillment*. Henry Holt and Company, New York.
- Berns, G. S., McClure, S. M., Pagnoni, G., and Montague, P. R. (2001). Predictability modulates human brain response to reward. *The Journal of Neuroscience*, 21(8):2793–2798.
- Berridge, K. C. (1996). Food reward: Brain substrates of wanting and liking. *Neuroscience and Biobehavioral Reviews*, 20(1):1–25.
- Berridge, K. C. (1999). Pleasure, pain, desire, and dread: Hidden core processes of emotion. In Kahneman et al. (1999), pages 525–557.
- Binder, M. (2009). Elements of an Evolutionary Theory of Welfare. Dissertation, accepted for publication by Routledge.
- Blackorby, C., Bossert, W., and Donaldson, D. (2005). Multi-profile welfarism: A generalization. *Social Choice and Welfare*, 24:253–267.
- Braeutigam, S. (2005). Neuroeconomics - from neural systems to economic behaviour. *Brain Research Bulletin* 67, 67:355–360.
- Brentano, F. (1969). *The Origin of Our Knowledge of Right and Wrong*. Routledge & Kegan Paul, London.
- Broome, J. (1991a). A reply to sen. *Economics and Philosophy*, 7:285–287.
- Broome, J. (1991b). 'Utility'. *Economics and Philosophy*, 7:1–12.
- Bruni, L. and Sugden, R. (2007). The road not taken: How psychology was removed from economics, and how it might be brought back. *The Economic Journal*, 117:146–173.
- Cabanac, M. (1979). Sensory pleasure. *The Quarterly Review of Biology*, 54(1):1–29.

- Cabanac, M. (1992). Pleasure: The common currency. *Journal of Theoretical Biology*, 155:173–200.
- Camerer, C., Loewenstein, G., and Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature*, 43:9–64.
- Conover, K. L. and Shizgal, P. (1994a). Competition and summation between rewarding effects of sucrose and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, 108(3):537–548.
- Conover, K. L. and Shizgal, P. (1994b). Differential effects of postingestive feedback on the reward value of sucrose and lateral hypothalamic stimulation in the rat. *Behavioral Neuroscience*, 108(3):559–572.
- Crisp, R. (2006a). Hedonism reconsidered. *Philosophy and Phenomenological Research*, 73:619–645.
- Crisp, R. (2006b). *Reasons And The Good*. Oxford University Press, Oxford.
- Damasio, A. R. (2003). *Looking for Spinoza*. Heinemann, London.
- Delgado, M. R. (2007). Reward-related responses in the human striatum. *Annals of the New York Academy of Sciences*, 1104:70–88.
- Delgado, M. R., Nystrom, L. E., Fissell, C., Noll, D., and Fiez, J. (2000). Tracking the hemodynamic responses to reward and punishment in the striatum. *Journal of Neurophysiology*, 84:3072–3077.
- Everitt, B. J., Parkinson, J. A., Olmstead, M. C., Arroyo, M., Robledo, P., and Robbins, T. W. (1999). Associative processes in addiction and reward - the role of amygdala-ventral striatal subsystems. *Annals of the New York Academy of Sciences*, 877:412–438.
- Feldman, F. (2004). *Pleasure and the Good Life - Concerning the Nature, Varieties, and Plausibility of Hedonism*. Clarendon Press, Oxford.
- Gottfried, J. A., O'Doherty, J., and Dolan, R. J. (2002). Appetitive and aversive olfactory learning in humans studied using event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 22(24):10829–10837.
- Harrison, G. W. (2008). Neuroeconomics: A critical reconsideration. *Economics and Philosophy*, 24(3):303–344.
- Heath, R. G. (1996[1964]). *Exploring the Mind-Brain Relationship*. Moran Printing, Baton Rouge.
- Hume, D. (1978[1740]). *A Treatise of Human Nature*. Oxford University Press, Oxford, 2nd. edition.
- Hume, D. (1994). *Enquiries concerning Human Understanding and Concerning the Principles of Morals*. Oxford University Press, Oxford, 3rd. edition.

- Jamison, J. C. (2008). Well-being and neuroeconomics. *Economics and Philosophy*, 24(3):407–418.
- Johnston, V. S. (1999). *Why We Feel - The Science of Human Emotions*. Perseus Publishing, Cambridge/Mass.
- Kahneman, D., Diener, E., and Schwarz, N., editors (1999). *Well-Being: The Foundations of Hedonic Psychology*. Russell Sage Foundation, New York.
- Kahneman, D., Wakker, P. P., and Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, 112(2):375–405.
- Kenning, P. and Plassmann, H. (2005). Neuroeconomics: An overview from an economic perspective. *Brain Research Bulletin* 67, 67:343–354.
- Knetsch, J. L. (1992). Preferences and nonreversibility of indifference curves. *Journal of Economic Behavior & Organization*, 17:131–139.
- Knutson, B. and Peterson, R. (2005). Neurally reconstructing expected utility. *Games and Economic Behavior*, 52:305–315.
- Komisaruk, B. R., Whipple, B., Crawford, A., Grimes, S., Liu, W.-C., Kalnin, A., and Mosier, K. (2004). Brain activation during vaginocervical self-stimulation and orgasm in women with complete spinal cord injury: fMRI evidence of mediation by the vagus nerves. *Brain Research*, 1024:77–88.
- Landreth, A. and Bickle, J. (2008). Neuroeconomics, neurophysiology and the common currency hypothesis. *Economics and Philosophy*, 24(3):419–429.
- Lewin, S. B. (1996). Economics and psychology: Lessons for our own day from the early twentieth century. *Journal of Economic Literature*, 34(3):1293–1323.
- List, C. (2004). Multidimensional welfare aggregation. *Public Choice*, 119:119–142.
- Loomes, G. and Taylor, C. (1992). Non-transitive preferences over gains and losses. *The Economic Journal*, 102(411):357–365.
- Marshall, A. (1925). *Principles of Economics*. Macmillan, London, 8th edition.
- Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, New York/Oxford.
- McCabe, K. A. (2008). Neuroeconomics and the economic sciences. *Economics and Philosophy*, 24(Special Issue 03):345–368.
- McClure, S. M., York, M. K., and Montague, P. R. (2004). The neural substrates of reward processing in humans: The modern role of fMRI. *The Neuroscientist*, 10(3):260–268.
- McFarland, D. J. and Sibly, R. M. (1975). The behavioural final common path. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 270(907):265–293.

- Mill, J. S. (1998[1863]). *Utilitarianism*. Oxford University Press, Oxford/New York.
- Millenson, J. (1967). *Principles of Behavioral Analysis*. MacMillan, New York.
- Montague, P. R. and Berns, G. S. (2002). Neural economics and the biological substrates of valuation. *Neuron*, 36:265–284.
- O’Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, 4:95–102.
- O’Doherty, J. P., Deichmann, R., Critchley, H. D., and Dolan, R. J. (2002). Neural responses during anticipation of a primary taste reward. *Neuron*, 33:815–826.
- Pagnoni, G., Zink, C. F., Montague, P. R., and Berns, G. S. (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nature neuroscience*, 5:97–8.
- Rilling, J. K., Gutman, D. A., Zeh, T. R., Pagnoni, G., Berns, G. S., and Kilts, C. D. (2002). A neural basis for social cooperation. *Neuron*, 35:395–405.
- Rolls, E. T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10(3):284–294.
- Rolls, E. T. (2007). *Emotion Explained*. Oxford University Press, Oxford.
- Rolls, E. T., O’Doherty, J., Kringelbach, M. L., Francis, S., Bowtell, R., and McGlone, F. (2003). Representations of pleasant and painful touch in the human orbitofrontal and cingulate cortices. *Cerebral Cortex*, 13:308–317.
- Ryberg, J. (2002). Higher and lower pleasures - doubts on justification. *Ethical Theory and Moral Practice*, 5:415–429.
- Samuelson, P. (1961[1947]). *Foundations of Economic Analysis*. Harvard University Press, Cambridge/Mass.
- Schreiber, C. A. and Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *Journal of Experimental Psychology: General*, 129(1):27–42.
- Schultz, W., Tremblay, L., and Hollerman, J. R. (2000). Reward processing in primate orbitofrontal cortex and basal ganglia. *Cerebral Cortex*, 10(3):272–283.
- Sen, A. K. (1973). Behaviour and the concept of preference. *Economica*, 40(159):241–259.
- Sen, A. K. (1981). Plural utility. *Proceedings of the Aristotelian Society*, 81:193–215.
- Sen, A. K. (1991). Utility - ideas and terminology. *Economics and Philosophy*, 7:277–283.
- Shizgal, P. (1999). On the neural computation of utility: Implications from studies of brain stimulation reward. In Kahneman et al. (1999), pages 500–524.
- Shizgal, P. and Conover, K. (1996). On the neural computation of utility. *Current Directions in Psychological Science*, 5(2):37–43.

- Spash, C. and Hanley, N. (1995). Preferences, information and biodiversity preservation. *Ecological Economics*, 12:191–208.
- Stigler, G. J. (1950a). The development of utility theory. I. *The Journal of Political Economy*, 58(4):307–327.
- Stigler, G. J. (1950b). The development of utility theory. II. *The Journal of Political Economy*, 58(5):373–396.
- Sumner, L. W. (1996). *Welfare, Happiness, and Ethics*. Oxford University Press, Oxford.
- Warke, T. (2000a). Classical utilitarianism and the methodology of determinate choice, in economics and in ethics. *Journal of Economic Methodology*, 7(3):373–394.
- Warke, T. (2000b). Mathematical fitness in the evolution of the utility concept from Bentham to Jevons to Marshall. *Journal of the History of Economic Thought*, 22(1):5–27.
- Warke, T. (2000c). Multi-dimensional utility and the index number problem: Jeremy Bentham, J. S. Mill, and qualitative hedonism. *Utilitas*, 12(2):176–203.
- Warke, T. (2000d). A reconstruction of classical utilitarianism. *Journal of Bentham Studies*, 3.
- Witt, U. (1987). *Individualistische Grundlagen der evolutorischen Oekonomik*. J.C.B. Mohr (Paul Siebeck), Tuebingen.
- Witt, U. (2001). Learning to consume - a theory of wants and the growth of demand. *Journal of Evolutionary Economics*, 11:23–36.