

A Critical Reflection on Utilitarianism as the Basis for Psychiatric Ethics

Part I: Utilitarianism as an Ethical Theory

Michael Robertson MB BS FRANZCP

Director of Psychiatry, Royal Prince Alfred Hospital and Sydney South West Area Health Service (Eastern Sector), and Clinical Lecturer, Discipline of Psychological Medicine, University of Sydney, Sydney, Australia.

Garry Walter MB BS BMedSc PhD FRANZCP

Professor of Child and Adolescent Psychiatry, University of Sydney, and Area Clinical Director, Child and Adolescent Mental Health Services, Northern Sydney Central Coast Health, NSW, Australia.

ABSTRACT

Utilitarianism is one of the “grand Enlightenment” moral philosophies. It provides a means of evaluating the ethical implications of common and unusual situations faced by psychiatrists, and offers a logical and ostensibly scientific method of moral justification and action. In this first of our two papers, we trace the evolution of utilitarianism into a contemporary moral theory and review the main theoretical critiques. In the second paper we contextualize utilitarianism in psychiatry and consider its function within the realm of the professional ethics of psychiatrist as physician, before applying it to two dilemmas faced by psychiatrists as individuals and as members of a profession. We conclude that psychiatry must search beyond utilitarianism in grappling with everyday clinical scenarios.

tion *primum non-nocere* (“first, do no harm”) is one of the earliest utilitarian constructs.

These notions were formulated as a moral philosophy to provide an ethical framework for the political liberalism emerging in the post-Enlightenment West. Utilitarian ethics were to be the blueprints for social justice, and utilitarianism is habitually considered the starting point of contemporary moral philosophy (Kymlicka, 2002).

The original utilitarian ideas come from Jeremy Bentham (1748-1832), who constructed a hedonistic view of utilitarianism (Bentham, 1970/1823). To Bentham, man was at the mercy of ‘the pleasures’ and it was therefore preferable to be ‘a contented pig’ than ‘unhappy human’. Bentham did not valorise the ‘higher pleasures’, arguing that happiness arising from the mindless game of “pushpin” was as good as that from reading poetry. John Stuart Mill (1806-1873), by contrast, argued that cultural, intellectual, and spiritual pleasures are of greater value than the physical pleasures in the eyes of a competent judge (Mill, 1968). Mill viewed the maximization of some form of eudemonic happiness as the source of the good. In an assertion slightly undermining the secular humanism of his project, Mill sought to endorse his utilitarianism by proclaiming, “In the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility” (Mill, 1968 p. 16). Mill’s utilitarianism does not necessarily avoid the same difficulties as Bentham’s version, particularly the so-called ‘quantification problem’, i.e. how to measure overall pleasure. GE Moore (1873-1958) averred that no true conception of the good could be formulated, and that an intuitive view of maximizing “ideals”, like aestheticism, may be the ultimate goal of maximizing good (Moore, 1903/1988). Later, economist-driven formulations of the ultimate good of utilitarianism involved the satisfaction of preferences, allowing people to choose for themselves what has intrinsic value (Arrow, 1984).

Introduction:

Original Conceptions of Utilitarianism

The notion of maximizing pleasure, or avoiding pain, seems an intuitive *raison d’être*, and indeed forms the basis of a body of moral philosophies dating from antiquity. Ethical hedonism, first described by Epicurus (341 BC – 270 BC) (Epicurus, 1926), posits that the good life is one spent in pursuit of pleasure, defined simply as the avoidance of pain. A consequentialist philosophy holds that the rightness or wrongness of an action is determined solely by reference to the ‘goodness’ or ‘badness’ of the consequences of that action. Integrated, these two ideas give us the broad foundation of utilitarianism. Indeed in medicine, the injunc-

Recent Conceptualisations of Utilitarianism

Utilitarians writing since Mill have elaborated the original ideas and modified utilitarianism to make it more workable. RM Hare (1919-2002) distinguished between two levels of utilitarian thinking (Hare, 1981; Hare, 1997). Hare asserted the existence of more lofty 'critical' level of thinking, applying the so-called 'Golden-Rule Argument', as against an 'intuitive' level, utilising simple consequentialist principles and integrating emotional responses. The intuitive level applies at the ethical coalface, and its deliberations must be acceptable at the critical level, whereas critical levels of moral reasoning are the domain of the 'archangels'. This latter kind of elitist moral philosophy, assuming that the common man is incapable of any form of reflective moral agency, was first described by Henry Sidgwick (1838-1900) (Sidgwick, 1907) and has been termed "government house utilitarianism" (Williams, 1973)

The distinction between intuitive and critical levels has evolved into 'Act' and 'Rule' utilitarianism (Hare, 1963). Hare argued that his utilitarianism may have been more what Kant had in mind in his moral philosophy, and saw 'The Kingdom of Ends' of Kant's Categorical Imperative (Kant, 1964) as being utilitarian in nature (Hare, 2000).

Hare advanced his version of utilitarianism as a workable basis for psychiatric ethics (Hare, 1993), arguing that utilitarian accounts of psychiatric ethics are often abandoned because of the perceived duties of psychiatrists to their patients. Hare suggested that psychiatrists:

"need not think like utilitarians; they can cleave to principles expressed in terms of rights and duties and may, if they do this, achieve better the aims that an omniscient utilitarian would than if they themselves did any utilitarian calculation" (Hare 1993, p.30).

Rather than act automatically based on a simple calculation of maximized utility, the psychiatrist, as moral agent, acts on a utilitarian basis at the intuitive level, and reflects upon how rights and duties may be best served at a critical level.

Another formulation of utilitarianism is that of "negative utilitarianism", originally outlined by Karl Popper (1902-1994) in the aftermath of the political excesses of the 1930s and 40s (Popper, 1945). Negative utilitarianism argues that, as moral agents, we seek to prevent the greatest amount of harm or evil, as against maximizing preferences. An argument, *reductio ad absurdum*, against negative utilitarianism is the so-called 'pin-prick argument', which states it would be better to destroy humanity painlessly than allow one person to experience a pin-prick ("DP", 2006). Other, less straw-man arguments have also been made against negative utilitarianism (R. Smart, 1958).

The elaboration of utilitarianism by Peter Singer (1946-) follows on from Hare (Singer, 1993). Singer's principle of equality encompasses all beings with interests, and it requires equal consideration of those interests, whatever the species. This kind of universalization, Singer admits, is Kantian in spirit. Singer contends that suppressing individual need for that of the collective has a survival advantage, an argument for the naturalism of utilitarian ideas (Singer, 1981). All species may have an interest in avoiding pain

but few have an interest in cultivating their unique individual abilities and Singer considers this as justifying different treatments for different interests. This is manifest in his concept of 'diminishing marginal utility', a form of hyper-consequentialism whereby the distinction between interests is as much about the need as the desire for the preference. For example, a starving person's preference for food has greater utility in its allocation than someone who is only slightly hungry. In expanding this idea, Singer takes a 'journey' model of life, which measures the wrongness of taking a life by the degree this thwarts a life journey's goals. To Singer, only a personal interest in continuing to live brings the journey model into play. Singer's utilitarianism has led to heated debate, in particular over the manner in which his philosophy appears to validate euthanasia and abortion (Singer & Kuhse, 1985). The core of his argument here relates to the perceived value of life being linked to sentience and the capacity to reach a life's journey goals, two issues highly relevant to severe mental illness.

Advantages of Utilitarianism

The advantages of utilitarianism as an ethical theory lie in its intuitive appeal, particularly in the case of act utilitarianism, and its apparent scientific approach to ethical reasoning. Beauchamp and Childress (2001) have devised a set of criteria by which a moral theory can be assessed (Beauchamp & Childress, 2001). These include clarity, coherence, comprehensiveness, simplicity, explanatory power, justificatory power, output power and practicality. On criteria such as output power, practicality and clarity, utilitarianism fares well. However, on issues such as justificatory power and comprehensiveness, there are problems. The ethical decision making process in utilitarianism may be quite straightforward. However, the justification and practicality of many utilitarian based decisions are limited, and its comprehensiveness as a moral philosophy is also a source of criticism.

Beauchamp and Childress (2001) believe that the principle of utility approximates their principle of "beneficence" and that it has tremendous output power. They also regard one of utilitarianism's strengths is its fitting well with approaches to public policy.

Critiques of Utilitarianism

Over time, there have been a number of cogent criticisms of utilitarianism as a moral philosophy. The more practical critiques have focussed upon the simple issue of the measurement of outcome of a utilitarian choice. Whilst this problem is more difficult with the Benthamite version of utilitarianism, the matter of how robustly one can measure gratification of preferences is problematic. The issue of adaptive preferences, whereby people accept less because of low expectations (such as the 'contented slave'), is one such area (Elster, 1982). The issues of unexperienced preferences (i.e., ones we will never know existed) and granting harmful preferences are also challenges to preference utilitarianism (Kymlicka, 2002). Some have argued that this potential limitation can be overcome by only applying preference utilitarianism to goods which are universally desired or provide basic necessity (Goodin, 1995), or for some form of utilitarian elite, like that described by Sidgwick (1907), to

oversee and exclude irrational preferences (Rawls, 1980). Ronald Dworkin (1931) has distinguished between 'personal preferences', referring to self, and 'external preferences', referring to a person's choices about others (Dworkin, 1977). We have seen recent examples of this in the vexed area of gay couples being restricted in access to assisted fertility treatments or entitlements to social welfare on the basis of their relationship. The capacity of utilitarian calculations to manifest latent prejudices prompted Dworkin and other writers to call for external preferences to be proscribed (Harsanyi, 1976; Kymlicka, 2002).

There have been a number of other logically based challenges to utilitarianism. Among these are the so-called 'replaceability problem' (Foot, 1967), based upon a thought experiment involving the utilitarian justification of one healthy person being killed to provide transplant organs for a half a dozen others in need – a utilitarian calculation. This is as confronting as the metaphor offered by Le Guin (2000) in her short story, *The Ones Who Walk Away from the Omelas* (Le Guin, 2000), in which a thriving population's prosperity is contingent upon the torture and imprisonment of an individual.

The above consideration relates to the so-called 'doctrine of double effect', first outlined by Thomas Aquinas (c.1225-1274), which seeks to explain the permissibility of an action that causes a serious harm as a side effect of promoting some good end (Cavanagh, 1997). The distinction here is between the direct or wilful creation of harm as a means, rather than a regrettable consequence of seeking a good. In medical ethics, this issue has been discussed primarily in terms of the intentions of the moral agent, and the proportionality of the harm in relation to the good (Boyle, 1991). Those who see this as the morally vacuous side of utilitarianism have called for a degree of 'deontic constraint', to this principle, rather than tolerating completely impersonal considerations of the positive and negative effects of actions (Nagel, 1986). In other words, rather than be purely beholden to utility in a vacuum, the moral agent should also reflect upon duties to other persons.

Bernard Williams (1929-2003), one of utilitarianism's most comprehensive critics, challenged us with his thought experiment 'Pedro and Jim', as to whether we would execute one man to save ten (Williams, 1973). The utilitarian decides to shoot one man; however according to Williams, being compelled to act on the basis of utility alienates us from our moral agency. In answer to these, somewhat 'straw-man' arguments, some have asserted that utilitarian arguments only evolved to fit common or mundane situations, and therefore cannot credibly apply to the kind of extraordinary situations cited by critics of utilitarianism (Sprigge, 1965). Derek Parfit's "repugnant conclusion" argument (Parfit, 1984) also takes a logical knife to utilitarianism in that it is, according to the utilitarian calculation, better for the world to have 100 billion all living in marginal poverty than the current situation of wealth being concentrated in a comparatively small part of humanity. Bernard Williams argues that the utilitarian moral agent is both responsible for the consequences of the consequences of their actions, as well as failing to prevent the negative consequences of these. Williams charges that utilitarianism places the moral agent under the burden of unreasonable expectations (Williams, 1973), although others have taken the view that the responsibility for ongoing consequences of actions actually diminishes over time (J. Smart, 1973).

Williams has further charged that utilitarianism, like deontic ethics, alienates the moral agent from their moral agency (Williams, 1973). The idea of a 'U-Agent' (Brink, 1986), totally devoid of any personal morality and wedded to the utilitarian abacus, is clearly unrealistic, prompting some to soften the utilitarian stance to incorporate 'agent relative values' as against 'agent neutral values'. In the former, an act is considered morally wrong if its consequences have less overall value from the perspective of the agent; the latter where this is not a consideration (Sen, 1982). Indeed, any conceptualization of utilitarianism committed to our reneging on important personal commitments in order to promote the welfare of others, is unrealistic (Railton, 1984).

Kymlicka's criticism of utilitarianism (Kymlicka, 2000) is more historical than based on logic or thought experiments. Kymlicka argues that in Bentham's time utilitarianism was a progressive theory, but in modern liberal democracies it is a conservative one:

"In short, when the question is whether to defend an oppressed majority against a small privileged elite, utilitarianism gives us a clear, progressive answer. But when the question is whether to defend an oppressed minority against a large, privileged majority, utilitarianism gives us vague and confusing answers (p. 48)"

Before applying these considerations to three scenarios commonly faced by psychiatrists, we will summarize the strengths and criticisms of utilitarianism as a basis of ethics. In terms of the strength of utilitarianism it has the veneer of scientific and rational method; it fits well with decision making at a macro-policy level; and appears to parallel decision procedures in daily life.

The negative features of utilitarianism based moral choices are that they: involve assessments of preferences which may be biased or flawed; expect too much of the moral agent in responsibility for consequences of consequences and negative responsibility; may require abandonment of emotional or filial bonds; potentially involve alienation from moral agency; may involve the active disadvantage or harm of individuals; and, are based on a political and moral philosophy that is arguably anachronistic.

Conclusion

In this paper, we have described the evolution of utilitarianism as an ethical theory and considered advantages and disadvantages. There have been a variety of critiques of utilitarianism varying from practical concerns to well constructed, logically based arguments. For those who face complex decisions affecting many people, utilitarianism offers a valuable technique of ethical reasoning. For psychiatrists whose craft is steeped in the Hippocratic tradition, the principle of utility presents a challenge to their ethical obligations as physicians. This is particularly the case where conclusions based on utilitarian calculations may have negative consequences for the patient. We will consider these criticisms in the context of psychiatric practice in the second of our papers on this topic

This article has been peer reviewed.

Funding and Support: Nil

Competing Interests: None

Acknowledgements: None

References:

- "DP". (2006). The pin-prick argument. Retrieved 14th October, 2006, from <http://www.utilitarianism.com/pinprick-argument.html>
- Arrow, K. (1984). *Social Choice and Justice* (Vol. 1). Cambridge, MA: Bellknap.
- Beauchamp, T., & Childress, J. (2001). *Principles of Biomedical Ethics (5th Edition)*. New York: Oxford University Press.
- Bentham, J. (1970/1823). *An Introduction to the Principles of Morals and Legislation*. London: Althone Press.
- Boyle, J. (1991). Who is Entitled to the Double Effect? *Journal of Medicine and Philosophy*, 16, 475-494.
- Brink, D. (1986). Utilitarian morality and the personal point of view. *Journal of Philosophy*, 83, 417-438.
- Cavanagh, T. (1997). Aquinas's Account of the Double Effect. *Thomist*, 61, 107-121.
- Elster, J. (1982). Utilitarianism and the Genesis of Wants. In A. Sen & B. Williams (Eds.), *Utilitarianism and Beyond* (pp. 219-238). Cambridge: Cambridge University Press.
- Epicurus. (1926). Letter to Menoeceus (C. Bailey, Trans.). In *The Extant Remains*. Oxford: The Clarendon Press.
- Foot, P. (1967). Abortion and the doctrine of double effect. *Oxford Review*, 5, 28-41.
- Goodin, R. (1995). *Utilitarianism as a Public Philosophy*. Cambridge: Cambridge University Press.
- Hare, R. (1963). *Freedom and Reason*. London: Oxford University Press.
- Hare, R. (1993). *The philosophical basis of psychiatric ethics*. New York: Oxford University Press.
- Hare, R. (1997). A Philosophical Self Portrait. In *The Penguin Dictionary of Philosophy* (pp. 234-235). London: Penguin.
- Hare, R. (2000). Could Kant have been a utilitarian? In *Sorting Out Ethics* (pp. 147-191). Oxford: Oxford University Press.
- Harsanyi, J. (1976). *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.
- Kant, I. (1964). *Groundwork of the Metaphysic of Morals*. New York: Harper and Row.
- Kymlicka, W. (2002). *Contemporary Political Philosophy*. New York: Oxford University Press.
- Le Guin, U. (2000). The Ones who walk away from the Omelas. In *The Wind's Twelve Quarters*. London: Victor Gollancz.
- Mill, J. (1968). *Utilitarianism, Liberty, Representative Government*. London: JM Dent and Sons.
- Moore, G. (1903/1988). *Principia Ethica*. Amherst: Prometheus Books.
- Nagel, T. (1986). *The View from Nowhere*. New York: Oxford University Press.
- Parfit. (1984). *Reasons and Persons*. Oxford: Clarendon Press.
- Popper, K. (1945). *The Open Society and its enemies*. Princeton: Princeton University Press.
- Railton, P. (1984). Alienation, consequentialism, and the demands of morality. *Philosophy and Public Affairs*, 13, 134-171.
- Rawls, J. (1980). Kantian constructivism in moral theory. *Journal of Philosophy*, 77, 515-572.
- Sen, A. (1982). Rights and Agency. *Philosophy and Public Affairs*, 11, 3-39.
- Sidgwick, H. (1907). *The Methods of Ethics (7th Ed)*. London: Macmillan.
- Singer, P. (1981). *The Expanding Circle: Ethics and Sociobiology*. New York: Farrar, Straus and Giroux.
- Singer, P. (1993). *Practical Ethics (2nd Edition)*. Cambridge: Cambridge University Press.
- Singer, P., & Kuhse, H. (1985). *Should the Baby Live? The Problem of Handicapped Infants*. New York: Oxford University Press.
- Smart, J. (1973). *Utilitarianism and the Future*. In J. Smart & B. Williams (Eds.), *Utilitarianism: For and Against* (pp. 62-67). Cambridge: Cambridge University Press.
- Smart, R. (1958). *Negative Utilitarianism*. *Mind*, 67, 542-543.
- Sprigge, T. (1965). A Utilitarian reply to Dr McCloskey. *Inquiry*, 8, 264-291.
- Williams, B. (1973). A Critique of Utilitarianism. In J. Smart & B. Williams (Eds.), *Utilitarianism: For and Against* (pp. 75-150). Cambridge: Cambridge University Press.

Key Words: *Psychiatric Ethics, Utilitarianism, Critique, Professional Ethics*

Address for Correspondence:

e-mail: mrobertson@med.usyd.edu.au