Journal of Social & Organizational Matters
Vol 3 No 2 (2024): 124-143

Utilitarianism, Social Justice, and the Trolley Problem: An Ethical Theory without Egalitarian Morality

Saad Malook*[1]

[1]*Lecturer, Department of Philosophy, University of the Punjab, Lahore, Pakistan.

Corresponding author: saad.phil@pu.edu.pk

ORCID No: https://orcid.org/0009-0000-3289-9346

Keywords: *Utilitarianism, Injustice, Egalitarianism, Bentham, Ethics, Politics*

DOI No: https://doi.org/10.56976/jsom.v3i2.70

This article examines the implications of utilitarianism for social justice, considering different cases of the trolley problems. Utilitarianism comprises a cluster of ethical theses, which have political and legal implications. In general, utilitarianism is assumed to augment the common good, such as pleasure, happiness, satisfaction, and utility, or to reduce pain, suffering, dissatisfaction, and disutility. The article investigates a key problem whether utilitarianism brings about social justice as a moral theory. In recent literature, many moral philosophers have developed several thought experiments, known as trolley problems, that help understand the utilitarian applications to social justice. Philippa Foot and others formulated several thought experiments to help explain the utilitarian difference between the maximum good and the minimum loss. Considering different cases of the trolley problem, I argue that utilitarianism does not do egalitarian justice to all people because it does not eliminate unjust inequalities in society. I hold that utilitarianism is an ethical theory without an egalitarian morality. John Rawls posited a landmark theory of social justice in contrast to the utilitarian notion of social justice. Utilitarianism does not provide social justice to the few, the minority, or the weak in the social world. So, utilitarianism is an inadequate moral theory for acquiring social justice. In qualitative research methodology, I adopt a method of an empirically informed philosophical analysis to examine documentary resources, including journal papers, academic books, and conferences and congresses.

1. Introduction

This article explores a key problem whether utilitarianism promotes egalitarian social justice. Utilitarianism is a consequentialist ethical theory that decides the rightness or wrongness of an action or a rule on the basis of its consequences. Jeremy Bentham, the founder of utilitarianism, developed two versions of his principle of utility, which provide the foundation for utilitarian consequentialism. The first version of Bentham's principle of utility asserts, "It is the greatest happiness of the greatest number that is the measure of right and wrong" (Bentham 2000, p. 3). I call this the *particular version* of the principle of utility. The second version of the principle of utility holds, "principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to increase or diminish the happiness of the party [individuals] whose interest is in question" (Bentham, 2000, p. 88). Let us call this the *general version* of the principle of utility. Utilitarianism, in its essence, is based primarily on the *particular version* of the principle of utility and explains the difference between maximum good for the maximum number of people. This version of utilitarianism is consistent with the trolley problem and is inconsistent with egalitarian social justice.

I argue that utilitarianism is an inadequate moral theory because it fails to eliminate the unjust inequalities in society. Utilitarianism does not do justice to egalitarian social justice because it justifies harming a minority when that harm is necessary for the good of the majority. For instance, utilitarianism could justify acts commonly thought to be immoral, such as waging unjust wars, killing innocent people, breaking promises, telling lies, and violating human rights in certain circumstances. Different thought experiments, explained as trolley problems, reveal that utilitarianism does not support social justice by eliminating unjust inequalities in society. So, utilitarianism does not support egalitarian morality, and by so doing, it does not condemn the killing of an innocent person, breaking promises, telling lies, humiliation, exploitation, and torture of innocent people.

2. Literature Review and Research Methodology

Utilitarianism has two and half centuries of history, and abundant literature has been produced to explain it. Francis Hutcheson, Jeremy Bentham, James Mill, John Stuart Mill, Henry Sidgwick, David Lyons, J. J. Smart, Peter Singer, Robert Goodin, and many other philosophers contributed to the development of utilitarianism. Although it is generally considered that Bentham founded the school of utilitarianism, the genesis of the concept, the

principle of utility, goes back to Francis Hutcheson (1694-1746) who explicitly stated in 1726 that "[a]n [a]ction is best, which procures the greatest [h]appiness for the greatest [n]umbers; and that, … worst, which, in like manner, occasions [m]isery" (Hutcheson 2004, p. 125). Hutcheson identified the best action with the greatest happiness while the worst action with the greatest misery. Bentham would have drawn his notion of the particular version of the principle of utility from Hutcheson.

'Utilitarianism' consists of several philosophical theses about the nature of morality (Scanlon, 1982, p. 108). Russell Hardin describes the social interactionist character of utilitarianism: "Of all moral theories perhaps none has originated more clearly from a consideration of the fullest range of human interactions than utilitarianism, the theory that judges the rightness and wrongness of all actions according to their likely effect on human and perhaps on animal welfare" (Hardin, 1988, p. 12). Hardin states that this social interactionist aspect of utilitarianism has been mostly adopted by the theorists of trolley problems. Bentham concentrated on pleasure and pain. Bentham, James Mill, John Stuart Mill, and Henry Sidgwick, who advanced the notion of pleasure as utility, were hedonistic utilitarians. Yet modern utilitarians explain utilitarianism beyond hedonism. Goodin states: "Modern utilitarians now go well beyond Bentham's simple calculus of pain and pleasures, so it is no longer true to say that utilitarianism is necessarily a crassly "hedonic" philosophy. But dress up they're maximined as they will, modern utilitarians – to deserve the name at all – must necessarily be involved in maximising satisfactions, somehow construed, of people, somehow specified" (Goodin, 1995, p. 10). Goodin correctly states that modern utilitarians go beyond hedonism: pleasure and pain.

What is utility? According to Bentham, "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" (Bentham, 1996, p. 12). The trolley problem is a modern approach to examining utility. Prominent scholars who contributed to the development of the idea of trolleyology, except Philippa Foot, are G. E. M. Anscombe, Judith Jarvis Thomson, Warren S. Quinn, and David Edmonds. Foot in *The Problem of Abortion and the Doctrine of Double Effect* (1967), *Anscombe in Action, Intention and Double Effect* (1982), Thomson in *Trolley Problem* (1985), Quinn in *Actions, Intentions, and Consequences: The Doctrine of Double Effect* (1989) and Edmonds in *Would You Kill the Fat Man*? have constructed different thought experiments. In

this article, I employ the method of empirically informed philosophical analysis to evaluate scholarly papers, books, and proceedings of the conferences and congresses to explain the difference between the maximum good and the minimum loss.

3. The Standard Account of Utilitarianism

Utilitarianism incorporates a cluster of consequentialist theories, which comprise a myriad of doctrines relevant primarily to psychology, ethics, politics, and law. The standard account of utilitarianism maintains that consequences determine whether an action or a rule is right or wrong. Bentham founded a philosophical sect called utilitarianism. The standard account of utilitarianism has three core theses: maximise utility or minimise disutility, be impartial, and only consequences determine the rightness of actions. Considering these theses together in a moral and political context, classical utilitarianism asserts that an action or a rule is morally right if and only if the action or the rule will maximise utility (or minimise disutility) for the maximum number of people or for everyone if possible. The central doctrines of utilitarianism are explained to provide a comprehensive understanding of utilitarianism and its moral and political implications.

3.1 Principle of Maximising Utility

The concept of 'utility' is the bedrock of utilitarianism because it provides the core principle to theorize moral and political decision-making. Yet, the meaning of term, 'utility' differs in classical and modern utilitarianism. Most classical utilitarians support hedonistic conceptions of utility or disutility while many modern utilitarians support non-hedonistic conceptions of utility. Bentham the early advocate of classical utilitarianism, and his disciples, including James Mill, John Stuart Mill, and Henry Sidgwick, developed hedonistic utilitarianism. This conception of utility considers pleasure and pain as two "sovereign masters" that regulate human behaviour (Bentham, 2000, p. 87). In contrast, many modern utilitarians, such as John Harsanyi, Richard Brandt, and Peter Singer have developed accounts of preference utilitarianism, while Robert E. Goodin and David Parfit have developed accounts of welfare utilitarianism.

As it is mentioned earlier that Bentham developed two versions of his principle of utility, which I call the particular and the general: The particular version of the principle of utility asserts, "It is the greatest happiness of the greatest number that is the measure of right and wrong" (Bentham 2000, p. 3). This version of the principle of utility provides the foundation for the principle of maximizing utility. Yet, the general version of the principle of utility states:

"principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to increase or diminish the happiness of the party [individuals] whose interest is in question" (Bentham, 2000, p. 88). Alasdair MacIntyre states:

> To have understood the polymorphous character of pleasure and happiness is of course to have rendered those concepts useless for utilitarian purposes; if the prospect of his or her own future pleasure or happiness cannot for reasons which I have suggested provide criteria for solving the problems of action in the case of each individual, it follows that the notion of the greatest happiness of the greatest number is a notion without any clear content at all. It is indeed a pseudo-concept available for a variety of ideological uses, but no more than that (MacIntyre, 2007, p. 64).

MacIntyre criticises the particular version of the principle of utility, which he calls a pseudo-concept. MacIntyre is not wrong because this version of utility has some malign implications. This version of utility divides people into the majority and the minority, and the minority has to sacrifice their utility for the majority.

For J. S. Mill, 'principle of utility' is "a creed, a doctrine, a philosophy; in one among the best senses of the word, a religion" (J. S. Mill, 1981, p. 69). Mill's statement is correct because the principle has a wide scope. Interestingly, Francis Hutcheson developed an idea before Bentham that has a close affinity with his theory of utility. Hutcheson states "that action is best, which procures the greatest Happiness for the greatest numbers; and that, worst, which, in like manner, occasions misery" (Hutcheson, 2004, p. 125). Accordingly, Hutcheson identified the best action as maximising happiness and the worst action as causing misery. Bentham's particular version of the principle of utility and Hutcheson's principle of utility have striking resemblance.

Bentham's principle of utility has both psychological and ethical elements. The underlying assumption is that pleasure and pain are the basic drives of human motivation and, therefore, drive human action (Bentham, 2000, p. 88). From the perspective of psychology, Bentham's pleasure principle is a kind of psychological hedonism. It is the theory that human action is always motivated by a desire to experience pleasure and avoid pain. Bentham states:

> Nature has placed mankind under the governance of two sovereign masters, pain and pleasure. It is for them alone to point out what we ought to do, as well as to

> determine what we shall do. On the one hand the standard of right and wrong, on the other the chain of causes and effects, are fastened to their throne. They govern us in all we do, in all we say, in all we think: every effort we can make to throw off our subjection, will serve but to demonstrate and confirm it (Bentham, 1996, p. 3).

According to Bentham, pleasure and pain provide a criterion of right or wrong actions, in this way Bentham extends psychological hedonism to ethical hedonism.

3.2 Principle of Impartiality

Impartiality is a significant notion in ethics and politics, and it has numerous conceptions. Prominent accounts of impartiality, include Adam Smith's 'Impartial Spectator', Henry Sidgwick's 'The Point of View of the Universe', Immanuel Kant's 'Formula of Humanity', Kurt Baier's 'The Moral Point of View, John Rawls's Veil of Ignorance, and Amartya Sen's 'Open and Closed Impartiality'. There are two significant accounts of impartiality in utilitarian and non-utilitarian moral and political theories. The utilitarian concept of impartiality does not treat a person from an individual point of view but from the point of view of the universe (Sidgwick, 1962, p. 382, 421). Methodologically, the utilitarian concept of impartiality is not concerned with particular individuals but the entire humanity. Utilitarians determine whether an action is impartial by looking at whether the happiness of everyone affected is counted as equally important.

For instance, a judge can decide cases in two ways. First, one can decide the cases by looking at the consequences of the decision. Second, one can decide the cases by not looking at the consequences of the decision. In *Enquiry Concerning Political Justice* (1793), William Godwin presents one of the best illustrations of the principle of utilitarian impartiality. Godwin states, "At the social level, no regulations may be legitimately enforced by public authorities which do not promote the utility of the people. Moral duty and justice can be jointly defined as the impartial treatment of every [hu]man in matters that relate to his happiness" (Godwin, 1793, p. 69-70). Godwin asserts that no ethical and political decision-making can be legitimised until it treats human persons impartially in order to ensure the decision has the best effect on overall happiness. John Stuart Mill in *Utilitarianism* (1861) developed a similar position to Godwin's. Mill writes, "The happiness which forms the utilitarian standard of what is right in conduct, is not the agent's own happiness, but that of all concerned. As between his own happiness and that of others, utilitarianism requires him to be as strictly impartial as a disinterested and

benevolent spectator (J. S. Mill, 1969, p. 218). Like Bentham, Mill identifies the rightness of an action with its effect on the happiness of all concerned members. He states that actors must be impartial when determining which action is right, acting as if they were disinterested and benevolent spectators. Mill's idea of the disinterested and benevolent spectator has striking resemblance with Adam Smith's idea of impartial spectator. Smith used the metaphor of the impartial spectator for an actor to be "a judge or arbiter" (Smith, 2009, p. 290) who decides the cases under a "candid and impartial light" to bring about justice (Smith, 2009, p. 28).

Godwin's thought experiment helps explain the notion of utilitarian impartiality (Godwin, 1793, p. 82-3). Imagine that archbishop Fenelon of Cambray and his chambermaid are caught in a house fire. What should one do if only one of the two can be saved? This question is concerned with the criterion of preference. Should one save the chambermaid or Fenelon? Godwin asserts that these people are not alone; they belong to a society, nation or in some sense the whole family of humankind. Godwin holds that the criterion for determining who should be saved should be based on consequences of their lives to society. Just as a human is more significant to society than a beast, an archbishop is more significant to society than a common girl. Even if the chambermaid is the mother of the actor, Godwin supports saving Fenelon because of the utilitarian point of view (Godwin, 1793, p. 82-3). Thinking from another way, suppose that the chambermaid is a genius and with her intelligent mind, she can be more productive than Fenelon. What should a saviour do? If she can contribute better to society than Fenelon, the chambermaid ought to be saved. However, again, the reason for saving her is not because the actor ought to act with partiality towards the chambermaid because of her role as mother of the actor. So, utilitarian impartiality means that an act is impartial if and only if it treats the happiness of each affected person as equally important when determining what action will maximise overall happiness.

3.3 Principle of Consequences

Utilitarianism is a consequentialist moral theory. Being a consequentialist moral theory, utilitarianism asserts that the morality of actions or rules depends only upon their respective consequences. The contemporary standard account of utilitarianism maintains that a right action maximises the overall good in the world (Barry, 1995, p. 23). The role utilitarianism gives to maximising good consequences means that the moral rightness of particular forms of action, such as killing, lying or gossiping, varies depending on the consequences of acting that way in a particular situation. For example, if a breaking promise brings about the best overall

outcome, then breaking the promise is the morally right thing to do. As a result, utilitarians claim that in some situations people ought to act in ways that are inconsistent with common social morality.

I consider two examples to discuss whether utilitarianism is consistent with commonly accepted social moral values, promise-keeping and social justice. In *Moral Life* (2000), Louis P. Pojman creates a thought experiment. He writes that there is a dying rich person on an island. This rich man has been a great supporter of the New York Yankees, a baseball team, for five decades. At the eleventh hour of his life, he decides to donate his entire property of six million dollars to the Yankees for the development of the baseball team. He reveals his will to you and you promise him that you will transfer the money to George Steinbrenner, the owner of the New York Yankees. On reaching New York, you notice a newspaper advertisement, by the World Hunger Relief Organisation, requesting a donation of six million dollars for saving and sustaining the lives of about one hundred thousand starved people in East Africa. What will you do? You have two options, first, respect the will of the deceased person by fulfilling your promise to deliver the money to Steinbrenner, second, ignore the will of the deceased person, break the promise, and present the money to the World Hunger Relief Organisation. If you are committed to the morality of promise-keeping, you will complete your promise and deliver the money to Steinbrenner. If you are a utilitarian, you will give the money to the World Hunger Relief Organisation because saving one hundred thousand people brings about more utility than empowering a sports team (Pojman, 2000, p. 219). Certainly, a consistent utilitarians ought to break promises, tell lies, and kill innocent people when these actions will bring about the best possible consequences.

In a similar vein, Philippa Foot (2002) posits an example illustrating the consequences of utilitarianism for social justice. Suppose someone committed a crime in a community and the civil society protested against the crime. A large, violent mob of people demands the judge executes a criminal for the crime. The problem is that if someone is not executed in time, the violent mob might vandalise public property. The judge understands that the culprit is unidentified and there is no strong evidence against anybody. What ought the judge do in this state of affairs? One option is that the judge waits for the investigation to identify the culprit so that he may execute the criminal. This may take some time and the enraged mob may start destroying public property. Another option would be that the judge declares an innocent person a criminal and executes him so that the enraged public may come to its senses with the thought

that the real culprit has been executed (Foot, 2002, p. 23-4). From the point of view of utilitarianism, killing an innocent person can be justified if no other action can bring about more happiness. In contrast, from the non-utilitarian perspective, killing an innocent person is considered to be morally wrong and unjust.

These two cases show that if the morality of actions or rules is judged on the grounds of consequences, commonly accepted social moral values are not consistent. Consequentialist theorists assert that the moral rightness or wrongness of actions, rules, and policies is decided by their particular consequences. Before considering the main problem of whether utilitarianism is compatible with egalitarian social justice, it is significant to explain the difference between the two kinds of utilitarianism.

4. Two Kinds of Utilitarianism

Utilitarianism has two major kinds: Act-Utilitarianism and Rule-Utilitarianism.

4.1 Act-Utilitarianism

Act-utilitarianism is considered as a mainstream form of utilitarianism. Yet, the term, 'act-utilitarianism' was not used by classical utilitarians. Richard Brandt introduced the terms, act-utilitarianism and rule-utilitarianism (Brandt, 1979, p. 71, 74). Brandt describes 'act-utilitarianism' as the holding that "an act is morally right if and only if the total welfare-expectation for everyone affected by it is at least as great as from any alternative action open to the agent" (Brandt, 1979, p. 271). There are numerous moral situations which demand utilitarian decision-making at the individual level. This kind of utilitarian decision-making is what R. M. Hare calls, the rational level of moral reasoning. Russell Hardin holds that the dominant spirit of twentieth century utilitarianism is bases on "personal problems of choice in small number of interactions such a personal promising, and face-to-face charity" (Hardin 1988, 12-13). Trolley problems are a numerous set of thought experiments to analyse different moral conflicts at interpersonal level. Many moral and political philosophers use trolley problems in their writings, particularly in the explanation of act-utilitarianism.

Bentham's act-utilitarianism is predominantly interpreted as a theory that directly deals with individual actions and their consequences, yet many contemporary utilitarian scholars have developed theories that consider the consequences of rules or public policies. In the contemporary epoch, trolley problems are the best illustrations to explain personal moral problems. In contrast, rule-utilitarianism primarily deals with public policies. When utilitarians

develop public policies they rely primarily on rule utilitarianism. Act-utilitarianism deals with personal moral problems that direct individuals to act for greater utility in the context of interpersonal moral and political interactions. Bentham claims that individuals are real entities while a "community is a fictitious body, composed of the individual persons who … [constitute] its members. The interest of the community… is the sum of the interests of the several members who compose it" (Bentham, 2000, p. 3). Bentham argues that the interests of the individuals determine the interests of the community. The interests of individuals mean "the sum total of … [their] pleasures: or, what comes to the same thing, [the diminishment of] … the sum total of … [their] pains (Bentham, 2000, p. 88). Bentham's words show that it is the enhancement of individuals' pleasures and removal of their pains that matters for his form of classical utilitarianism.

In *Habits of the Heart: Individualism and Commitment in American Life* (1985), Robert N. Bellah, Richard Madsen, William M. Sullivan, Ann Swidler, and Steven M. Tipton develop an argument that supports utilitarian individualism. According to Bellah and others, 'utilitarian individualism' is the thesis that: " in a society where each [individual] vigorously pursued his own interest, the social good would automatically emerge" (Bellah, Madsen, Sullivan, et al, 1985, 47). The essential features of these individuals include autonomy and self-interest. A utilitarian individual is an "autonomous individual, presumed able to choose the roles he will play and the commitments he will make, not on the basis of higher truths but according to the criterion of self-effectiveness as the individual judges it" (Bellah, Madsen, Sullivan, et al, 1985, 47). I endorse the argument, developed by Bellah and others, hold that individuals judge what it means to be right or wrong on the basis of the expected consequences of his / her actions.

Bellah et al. maintain, "That is also the voice of a utilitarian self-seeking its separate identity in the exercise of its own growing powers, ever freer of a restraint by others and ever farther out in of front of them" (Bellah, Madsen, Sullivan, et al, 1985, 68-9). So, in a utilitarian society, people with a "utilitarian self-established by "becoming one's own man" and then "settling down" to progress in a career" what one judges to be self-effective (Bellah, Madsen, Sullivan, et al, 1985, 69).

Michael Oakeshott holds that classical utilitarian theorists could neither understand the problem of the relationship between the individual and the society nor realise its consequences (Oakeshott, 1993, p. 52-3). He holds that classical utilitarian theorists could not explain the social relations between individuals. These relations, according to Oakeshott, do not only

depend upon the economic and physical necessities of life but also on the "unity of mind" (Oakeshott, 1993, p. 57). This sense of the 'unity of mind' in Oakeshott's terms, refers to a shared mind, which is essential for a social life. Classical utilitarian theorists focus on interpersonal relationships, which do not develop a unity of mind among people but material interests. Utilitarianism is a consequential theory. This consequential character of utilitarianism implies individualistic decision-making. To explain this consequential character and its individualistic implications, I explain numerous trolley problems in section 3.1.

4.2 Rule-Utilitarianism

In this section, I explain rule-utilitarianism to understand its implications for egalitarian social justice. In response to the criticism of act-utilitarianism, the idea of rule-utilitarianism appeared. Rule-utilitarianism asserts that an action is ethical if it would be required by a rule that, if it were followed by everyone in society, would bring about the most utility (Smart and Williams, 1973, p. 9). Certainly, rule-utilitarianism has advantages to act-utilitarianism because it supports public policy and public values. The significant feature of rule-utilitarianism which makes it distinctive is the role it gives rules. If everyone follows a rule, and everyone following that rule reliably results in the best outcomes, that rule ought to be followed, if not, it ought not to be followed.

David Lyons highlights the significant role of rule in utilitarian decision-making. In *Forms and Limits of Utilitarianism* (1965), Lyons makes a distinction between two strands of utilitarianism with an example. He mentions a dialogue between a husband and his wife travelling across an apple orchard. The wife asks the husband to pick some apples from the orchard. He refuses, arguing that it is wrong to do that because if everyone passing through the orchard picking the apples, there would be a grave loss to the grower. She responds to her husband that it may not be the case that everyone would pick the apples, and if just he picks the apples, the grower would not miss them (Lyons, 1965, p. 2). Lyon uses this dialogue between wife and husband to make a distinction between two kinds of utilitarianism. He calls these 'simple' and 'general' kinds of utilitarianism instead of terming them act-utilitarianism or rule-utilitarianism. According to Lyons, simple utilitarianism is concerned with the value-criteria that determines the rightness or wrongness of an act on the basis of the value of the effect of that single act. It is simply the value of utility of the action in comparison with the value of the utility of other possible actions (Lyons, 1965, p. 3). While the general kind of utilitarianism is concerned with judgements or rules. This means that an action is determined

to be wrong or right or obligatory depending on the value of the action's generalised utility (Lyon, 1965, p. 3-4).

The main problem in rule-utilitarianism is that sometimes following a rule does not bring about more utility than other possible actions. If breaking the rule in a particular situation will bring about more utility, should the rule still be followed? There is a diversion on this problem. Some theorists argue that rule-utilitarianism is not inconsistent with act-utilitarianism because making rules up to the requirement of the time or situation is promising. So, amending, evolving or changing a rule is not objectionable when it makes a theory more successful. Utilitarians hold that rules should be fine-tuned at the time of requirement. So, rules will change if we have a rule "drive on the left in New Zealand", it will change to "drive on the right in New Zealand" if the road rules change. In contrast, some theorists argue that rule-utilitarianism collapses into act-utilitarianism if utilitarians change the rules in this way. However, amending rules in accordance with the requirement of the problem is promising.

The issue with rule-utilitarianism is that rules are transitory, that they need to be changed over time to maximise utility. If rules are changed to those rules which would create more utility, the question arises that what would be the scope of the old rule? Rules would only need to be changed if the utility of following the rule changes. J. J. C. Smart argues that it would be irrational for a person to follow a rule in a situation where not following the rule will bring about more happiness (Smart and Williams, 1973, p. 10). In a utilitarian way of thinking, the agent should not follow the rule in any situation where following the rule does not maximise overall utility. If one wants to act rationally, but remain a rule-utilitarian, when following the old rule no longer maximises utility, there is a need to create a new rule. If actors create new rules time and again, it means that there is, perhaps, not much difference between act-utilitarianism and rule utilitarianism. So, rule-utilitarianism collapses into act-utilitarianism.

R. M. Hare attempted to synthesise act-utilitarianism with rule-utilitarianism. Hare argues in support of basing actions on two "levels" of utilitarian reasoning, a level of intuitive moral reasoning and a level of rational moral reasoning (Hare, 1981, p. 43-4). Hare's two-level utilitarianism holds that people should resolve their moral conflicts by means of a set of intuitive moral rules, except particular moral conflicts where critical moral reasoning might be required (Hare, 1981, p. 43). The nature of moral conflicts determines the main distinction between these levels of utilitarianism. At an intuitive level, moral conflicts appear to be

irresolvable. Whereas at a rational level, moral conflicts appear to be resolvable. For instance, assume that there is a moral conflict between actions A and B. On one hand, 'one ought to do A', and on the other hand, 'one ought to do B', and 'one cannot do both A and B'. This might be permissible at the intuitive level but not permissible at the rational level (Hare, 1981, p. 26). This means that ignoring the distinctions between the critical and intuitive levels of moral thinking creates the conflict between act-utilitarianism and rule-utilitarianism. According to Hare, utilitarianism is characterised by both kinds of levels.

Hare's intuitive level of moral reasoning is underpinned by prima facie principles that constitute rule-utilitarianism, so it is consistent with rule-utilitarianism. In contrast, the rational level of moral reasoning is consistent with act-utilitarianism, which focuses on specific cases (Hare, 1981, p. 43). Rule-utilitarianism provides rules for unlimited specificity. In short, two kinds of utilitarianism can coexist at their respective levels. A critical rational actor "considers cases in an act-utilitarian or specific rule-utilitarian way,…general prima facie principles for use, in a general rule-utilitarian way, at the intuitive level" (Hare, 1981, p. 43). If it is at the intuitive level, does it follow a common rule? However, rules are not eternal.

One of the central criticisms of utilitarianism is that it requires a utilitarian decision-maker to calculate the utilities of different possible actions to determine the optimal choice. Hare argues that the practice of calculating interpersonal utilities is not only impossible but also dangerous (Hare, 1981, p. 121). Hare's concern is correct, not just because of his reasons for holding this position, but also because if every individual is required to calculate the utilities of all their potential actions, the emerging outcome may be inconsistent with the traditional common morality. Instead, Hare holds that general principles usually provide a reliable ethical guide for what people should do. Interpersonal comparison is not a problem at an intuitive level of moral thinking. For instance, prima facie principles, such as the principle of beneficence, would guide people to bring about the good. Hence, Hare's two level's of utilitarian reasoning helps understand utilitarian actors can decide what to do and what do not to do while taking decision-making.

Utilitarianism is inconsistent with egalitarian social justice. For instance, it justifies telling lies, killing innocent people and breaking promises. In certain circumstances, act-utilitarianism justifies harming a minority to benefit the majority. To refute this aspect of utilitarianism, John Rawls states, "I do not believe that utilitarianism can provide a satisfactory account of the basic rights and liberties of citizens as free and equal persons, a requirement of

absolutely first importance for an account of democratic institutions" (Rawls, 1999a, p. xii). I agree with Rawls's argument that act-utilitarianism fails to develop a promising account which can guarantee the basic rights and liberties of people as free and equal citizens in democratic countries. Thus, act-utilitarianism does not support freedom or equal opportunity unless these things maximise utility.

5. The Trolley Problem, Doctrine of Double Effect, and Egalitarian Social Justice

In his classic novel, *Crime and Punishment*, Fyodor Dostoevsky states:

> "In my opinion, if, as the result of certain combinations, Kepler's or Newton's discoveries could become known to people in no other way than by sacrificing the lives of one, or ten, or a hundred or more people who were hindering the discovery, or standing as an obstacle in its path, then Newton would have the right, and it would even be his duty... to remove those ten or a hundred people, in order to make his discoveries known to mankind. It by no means follows from this, incidentally, that Newton should have the right to kill anyone he pleases…" (Dostoevsky, 1917, p. 263) .

Dostoevsky remarkably makes a case for the trolley problem: to acquire the good from the bad. This case is also related to the doctrine of the double effect (the doctrine is explained later). The term 'trolleyology' or 'trolley problem' has appeared to a large extent in the contemporary literature of moral theory. Philippa Foot introduced different thought experiments to understand whether a person's killing is permissible under certain circumstances (Foot, 2002, p. 21). Prominent scholars who contributed to the development of the idea of trolleyology include Philippa Foot G. E. M. Anscombe, Judith Jarvis Thomson, Warren S. Quinn, and David Edmonds.

Trolley problem is consistent with the principle of the doctrine of double effect. The doctrine of the double effect evolved from the works of Thomas Aquinas, Saint Augustine, Philippa Foot and G. E. M. Anscombe (Rovie, 2006, p. 1). The doctrine appeared as the result of the dispute between Aquinas and Augustine on the problem of whether killing a person is permissible. Augustine argues killing a person in self-defence is not legal. Augustine's argument holds that killing a person cannot be moral in any state of affairs (Augustine, 2010, p.10). In *Summa Theologica*, Aquinas refuted Augustinian's argument by declaring that killing in self-defence is legal and permissible. Aquinas articulates that an act can have two effects:

one of them is intended while the other is unintended. Acts of self-defence have two effects: the primary intended effect is saving one's life, the secondary unintended effect is killing the offender. Thus, the intentional act of self-defence is permissible because the secondary effect was not intended. However, an intentional act may be unlawful if its good outcome is not proportional to its bad outcome (Aquinas, 1988, p. 226-7).

The doctrine of the double effect consists of a set of four moral theses: First, the act should not be immoral. Second, one must intend a good effect rather than a bad effect. Third, a good effect should not causally arise from a bad effect. Fourth, a good effect should be at least as good as a bad effect is bad. The fourth thesis judges the action of an actor on the ground of the proportionate value between good consequences and bad consequences. Most trolley problems show that the agents make moral decisions relying on their evaluative skills. Agents decide on the action based on their evaluation of the consequences of the action.

There is collective but "the individual agent[s] to be deciding, quite often, what is the right thing to do" (Sen and Williams, 1990, p. 2). If individual agents are required to decide what it is right to do and are required to employ their own rationality when doing so, they depend upon their own individual minds. Individual agents can use their own rationality while considering information from others in their community, from libraries - both fiction and non-fiction books, from being raised within a good family that helps them learn what will and will not hurt others. Thus, act-utilitarianism deals with the individual aspects while rule-utilitarianism deals with the social aspects.

Act-utilitarianism is consistent with such thought experiments which make explain utility and disutility in certain circumstances. Consider a fat man who is stuck in the narrow mouth of a cave while five of his companions are enclosed inside the cave. The trapped people might wait until the fat man becomes thin enough to pass through the mouth of the cave and make the way open for the remaining five people. But unfortunately, flood water is rising, and the trapped people do not have enough time to open the mouth of the cave. The trapped people hold a rod of dynamite that can be used to blow up the fat man to open the way. They have the choice to blast the fat man and save their lives or drown (Foot, 2002, p. 21). From a utilitarian point of view, saving five people is better than saving one. However, five people could not be saved without taking the life of one person. An act-utilitarian response to the situation would justify the killing of an innocent person.

Foot also makes a case that has become the basis of the trolley problem. A driver of a tram has a choice to drive his tram to a track where five are working or to another track where one person is working. Anyone on the track the tram drives down will be killed. What should the driver do? (Foot, 2002, p. 23). Similar to the previous case, the utilitarian answer would appeal that the driver should behave in such a way that the five people be saved while one might be killed. Could one person's killing be justified? There are two divergent arguments for answering the question. Killing a person cannot be justified. Instead, utilitarian theorists would argue that one person's killing can be justified if that death saves five persons and no alternative action would bring about as much utility.

In certain circumstances, the humiliation of a person may bring about more utility than any alternative action. Similarly, the action that maximises utility in a particular context could also involve exploitation, discrimination or coercion. Utilitarians argue that wars should be fought when fighting wars maximises utility - that is, in situations where there is no other action that could be carried out that will bring about as much good in the world. Pettit makes a juxtaposition between deontologists and consequentialists on the problem of peace:

> "Every pacifist thinks that peace is a very important value, in particular an important value for states to recognise. But a pacifist may be deontological or consequentialist about the value in question; these are the salient possibilities, if not strictly the only ones. The pacifist who thinks that peace is a value that should always be honoured by a state, even if this means that there is less peace thereby realised, is a deontologist. He will say that an individual state should never go to war, even if the war would do better than appeasement in promoting peace. On the other hand, the consequentialist pacifist will say that the state should go to war, if this is the best way in the long run of promoting peace: if this is truly, for example, the war to end all wars" (Pettit, 1996, p. 303).

The utilitarian standpoint which provides the logic of war for ending war does not seem plausible because good cannot be acquired from evil. War is evil in any sense because it opens a new way to start an indefinite series of wars in the future.

Consider the following thought experiment to investigate whether exploitation is immoral. Suppose a surgeon has five patients in a hospital. Each patient needs a particular organ. One patient requires a heart, two require a kidney each and similarly two need lungs each. They all will die if these organs are not transplanted today. A young person just came to

get regular clinical check-ups. Coincidentally his blood group is similar to the five patients. This young man could be a donor of five required organs. If the surgeon takes the life of the young man and saves five people, this will bring about better in the world than any other action the surgeon could carry out. What act is permissible for the surgeon to do? (Thomson, 1985, p. 1395). I hold that taking organs from one person and using them for another without his consent is exploitation and immoral. Thus, utilitarianism is inconsistent with egalitarian social justice because it does not respect the equal rights of people.

6. Conclusion

This article investigated the problem whether utilitarianism is consistent with egalitarian social justice. Benthamite utilitarianism is based on two versions of the principle of utility: the particular version and the general version. The particular version of the principle of utility asserts "It is the greatest happiness of the greatest number that is the measure of right and wrong" (Bentham 2000, 3). The general version of the principle of utility holds, "principle which approves or disapproves of every action whatsoever, according to the tendency which it appears to have to increase or diminish the happiness of the party [individuals] whose interest is in question" (Bentham, 2000, p. 88). These versions of the principle of utility give rise to two kinds of utilitarianism: act-utilitarianism and rule-utilitarianism. Act-utilitarianism embarks on the people's acts to measure whether they are moral and immoral acts while rule-utilitarianism focuses on the rules to measure whether they are moral and immoral. If rules are required to change, rule-utilitarianism turns into act-utilitarianism.

The article defends the claim that utilitarianism is inconsistent with egalitarian social justice because it justifies killing an innocent person, breaking promises, telling lies, humiliation, exploitation, and torture. Different thought experiments, explained as trolley problems, reveal that utilitarianism fails to eradicate unjust inequalities because it does not respect the equal rights of people.

Acknowledgements: *The author would like to thank Professor Mark Francis and Dr Carolyn Mason, faculty members of the Department of Political Science and Philosophy, University of Canterbury, Christchurch, New Zealand, for reading the original manuscript and making valuable comments and suggestions.*

7. References

Aquinas, T. (1988). *Summa Theologica* (II-II, Q. 64, art. 7), "Of Killing", in *On Law, Morality, and Politics*, eds. William P. Baumgarth and Richard J. Regan, S. J. Indianapolis / Cambridge: Hackett Publishing Co.

Augustine, S. (2010). *On the Free Choice of the Will, On Grace and Free Choice, and Other Writings*, ed. Peter King. New York and Melbourne: Cambridge University Press.

Barry, B. M. (1995). *A Treatise on Social Justice: Justice as Impartiality*. Oxford and New York: Oxford University Press.

Bellah, R. N., Richard Madsen, William M. Sullivan, Ann Swidler and Steven M. Tipton. (1985). *Habits of the Heart: Individualism and Commitment in American Life*. California: University of California Press.

Bentham, J. (1990). "Bentham's Dream", in James E. Crimmins, *Secular Utilitarianism: Social Science and the Critique of Religion in the Thought of Jeremy Bentham* (Oxford and New York: Oxford University Press.

Bentham, J. (1996). *An Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart (Oxford and New York: Oxford University Press.

Bentham, J. (2000). "A Fragment on Government". In *Jeremy Bentham: On Utilitarianism and Government*. Ross Harrison (ed.). UK: Wordsworth.

Brandt, R. B. (1979). *A Theory of the Good and the Right*. Oxford: Clarendon Press.

Dostoevsky, F. (1917). *Crime and Punishment*. Constance Garnett (Trans.). New York: PF Collier and Son.

Foot, P. (2002). *Virtues and Vices and Other Essays in Moral Philosophy*. Oxford: Oxford University Press.

Godwin, W. (1793). *An Enquiry Concerning Political Justice and its Influence on General Virtue and Happiness*. Vol. 1. London: G. G. J. and J. Robinson.

Goodin, R. E. (1995). *Utilitarianism as a Public Philosophy*. Cambridge: Cambridge University Press.

Hardin, R. (1988). *Morality with the Limits of Reason*. Chicago and London: The University of Chicago Press.

Hare, R. M. (1981). *Moral Thinking: Its Levels, Method and Point*. Oxford and New York: Oxford University Press.

Hare, R. M. (1999). *Objective Prescriptions and Other Essays.* Oxford and New York: Oxford University Press.

Hutcheson, F. (2004). *An Inquiry into the Original of our Ideas of Beauty and Virtue in two Treatises* (ed.). Wolfgang Leidhold. Indianapolis: Liberty Fund.

Lyons, D. (1965). *Forms and Limits of Utilitarianism*. Oxford: Clarendon Press.

MacIntyre, A. (2007). *After Virtue: A Study in Moral Theory*. (3rd ed.). Notre Dame, Indian: University of Notre Dame Press.

Mill, J. S. (1969). *Essays on Ethics, Religion and Society*. (ed.). John M. Robson. Canada: University of Toronto Press.

Mill, J. S. (1981). *Autobiography and Literary Essays*, eds. John M. Robson and Jack Stillinger (University of Toronto Press.

Oakeshott, M. (1993). *Religion, Politics and the Moral Life*, ed. Timothy Fuller. Yale University Press.

Pettit, P. (1996). *The Common Mind*: An Essay on Psychology, Society and Politics. 2nd edition. Oxford University Press.

Pojman, L. P. (2000). *The Moral Life* (Oxford and New York: Oxford University Press.

Rawls, J. (1999). *A Theory of Justice*. (revised ed.). Cambridge, Massachusetts: Harvard University Press.

Rovie, E. (2006). "Reevaluating the Historical Evolution of Double Effect: Anscombe, Aquinas, and the Principle of Side-Effects", *Studies in the History of Ethics* 2, no. 1.

Sen, A. and Bernard Williams. (1990). *Utilitarianism and Beyond*, Reprint. London and New York: Cambridge University Press.

Sidgwick, H. (1962). *The Methods of Ethics*. (7th ed.). London: Palgrave Macmillan.

Smart, J. J. C. (1973) and Bernard Williams. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.

Smith, A. (2009). *The Theory of Moral Sentiments*. Penguin Books.

Thomson, J. J. (1985). "The Trolley Problem", *The Yale Law Journal* 94, no. 6.