Telling Moral Tales: Exploring Ways of Enhancing the Realism and Explanatory Power of Ethical Thought Experiments

Terry Hyland

Free University of Ireland, Dublin, Ireland
Email: hylandterry@ymail.com

Abstract

James Wilson has criticised thought experiments in philosophical ethics for failing to reflect the reality of everyday life. As against this, many such experiments owe their potential value and efficacy to an abstract simplicity which gives them widespread explanatory power and application. Through the examination of two particular arguments—Peter Singer’s “Shallow Pond” and Philippa Foot’s “Trolley Problem”—a number of suggestions are made for strengthening the realism of ethical thought experiments by reference to sources drawn from literature and social science. The principal aim is to enhance the real-world relevance and cogency of the thought experiments by offering suggestions for material which can reinforce the main points of the arguments. Such supplementary support—including references to the current Coronavirus pandemic—for the ethical thought experiments is intended to enhance teaching and learning in moral education and philosophical ethics at all levels.

Keywords

Moral Thought Experiments, Literature, Social Science and Morals, Ethics and Education, Moral Philosophy, COVID-19 Pandemic

1. Introduction

James Wilson (2016) has explored the notions of internal and external validity in ethical thought experiments and offered trenchant critiques of a number of famous arguments in this sphere. Thought experiments in this sphere are described as “toy ethical cases that are designed to simplify an ethical problem along a number of dimensions, thus making the problem more philosophically
tractable” (p.4). Making use of the paradigm of the randomised control trial (RCT) common in medical science, an ethical thought experiment would be internally valid “to the extent that it has been carefully and systematically designed with sufficient care to give a high degree of confidence that the results reported accurately measure the nature and the effect size of the intervention tested.” However, even if such a trial was satisfied such criteria “it may tell us little about whether the same intervention will work in other circumstances.” (p.7).

In a similar way, external validity—which in relation to the RCT would mean that its results are “applicable to a wide variety of contexts” (p.8) other than those specified—is said to be crucial to the question of whether thought experiments are genuinely useful and relevant to ethical decision-making in the real world. Wilson claims that certain famous ethical thought experiments—Peter Singer’s “Shallow Pond” and Philippa Foot’s “Trolley Problem” are cited in particular—fail to meet the internal/external validity criteria and, moreover, “unavoidably depict worlds that are under-described” (p.13). Suggestions for enhancing the quality and validity of such experiments include the move from a top-down linear model of reasoning to a “translational model” drawing on social science which concentrates on the practical applications of ethical theory, and the idea of seeing such experiments as “artworks” which may provide ethical insights in the way that works of fiction can do (pp.21-22). In the following sections, the Singer and Foot thought experiments—highlighted by Wilson for their widespread popularity and prominence in philosophy and mainstream academic literature—will be examined in detail with a view to enhancing their potency and status by means of a range of supplementary material from the spheres of social science and literature. The principal aim of this project is to provide additional material for analysis and discussion in the context of learning and teaching in ethics and moral education at all levels from school to university.

2. Peter Singer’s Shallow Pond

Singer’s famous “Shallow Pond” thought experiment was originally constructed in the context of an ethical argument for encouraging charity donations in the light of the 1971 refugee crisis in East Bengal and outlined in very brief form in Famine, Affluence and Morality (1972/2016). It is worth reporting in full:

If I am walking past a shallow pond and I see a child drowning in it, I ought to wade in and pull the child out. This will mean getting my clothes muddy, but this is insignificant whereas the death of the child would presumably be a very bad thing (p.7).

Singer goes on to claim that it “makes no moral difference whether the person I can help is a neighbour’s child ten yards from me or a Bengali whose name I shall never know, ten thousand miles away” (ibid.).

Over the years Singer has elaborated this thought experiment and defended its implications against a range of criticisms. A key challenge turns on the fact that the child in the pond is near to you whereas the starving children may be far
away in foreign countries. The response to this challenge is that—if we use our impartial reasoning capacities—it will be acknowledged that “there is no justification for having strong condemnation in one situation (i.e. in not saving the drowning child), and no such response in the other one” (ibid., p.xxvii) concerned with relieving poverty abroad.

The argument seems to satisfy the criteria for internal validity in that all the elements are clear, coherent and amenable to elaboration/adaptation without loss of potency and meaning. However, the fact that a significant number of people fail to make the inferences that Singer would wish to endorse indicates that there is a need of further support to strengthen the external validity of the principal elements. Moreover, there seems to be a strong case that the plausibility of moral argument may be “enhanced when one’s empirical claims are underpinned by the findings of empirical research” (De Ruyter, 2019: p. 644).

2.1. Social Science Perspectives

In the fifty years since Singer outlined the original thought experiment, his advocacy of charitable giving has advanced by leaps and bounds. Moreover, the movement—now labelled “effective altruism” (https://www.effectivealtruism.org/)—has been boosted by the organisation The Life You Can Save (2019) established by Singer to provide just that body of empirical social research which critics have demanded. Common misconceptions about charitable donating—such as philanthropy causing dependency, corruption or harming economic growth—are systematically dismantled by Singer who demonstrates forcefully the unfairness of the global economic system and explains in fine detail how systematic research by organisations such as Give Directly and Give Well has fostered public “confidence that people in extreme poverty will benefit from your donation, and benefit in a manner that is highly cost-effective” (Kindle edn, loc.1013).

In a similar way, the Oxford philosopher, William MacAskill, offers powerful reasons in Doing Good Better (2015) why all in the developed world should engage in effective altruism (EA). In reply to the common objection that anything we might do to try to relieve world poverty would amount to a futile drop in the ocean, he reminds us that:

we have the opportunity to provide a benefit for others that is one hundred times greater than the benefit we could provide for ourselves. That we can’t solve all the problems in the world doesn’t alter the fact that, if we choose, we can transform the lives of thousands of people (p.30).

MacAskill identifies five key questions which need to be addressed in debates about charitable giving:

1) How many people benefit, and by how much?
2) Is this the most effective thing you can do?
3) Is this area neglected?
4) What would have happened otherwise?
5) What are the chances of success, and how good would success be?

In taking these questions seriously and answering them fully as MacAskill does, we avoid the common pitfalls in addressing the topic of charitable donations and provide a comprehensive answer to the question of how we can do the most good which is at “the core of effective altruism’s approach to making a difference” (ibid., pp.17-18).

To be sure, some of the answers to the standard questions appear to be, at first glance, somewhat counter-intuitive. For example, MacAskill argues convincingly that those who wish to work in charitable organisations would do well to first find employment in effective and successful private companies rather than non-profit projects. Similarly, qualified physicians can do more for third world charities—not by volunteering to work in underdeveloped countries (since this work is generally covered by the “Doctors Without Borders” and similar organisations)—but by earning as much as they can in their home countries then donating as much as they are able to EA vetted charities. We are provided with a list of the potentially best EA causes by MacAskill in addition to some very practical suggestions for furthering the appeal and expansion of such work (ibid., pp.222ff.).

2.2. Evolution and Human Values

Such arguments echo Singer’s own justifications in The Life You Can Save (2019) which he admits that—though logically cogent and consistent—they have to face and overcome the alleged human tendency to favour our own interests over those of others. This question brings us face to face with the evolutionary basis of human behaviour concerned with competition vs co-operation in relation to ourselves, our families, tribes, nations as against all others—an issue which has been discussed at length by philosophers and social scientists.

Having shown in his earlier study The Better Angels of Our Nature (Pinker, 2011) that—in spite of our current pessimistic perspectives and preoccupations—“violence has declined over long stretches of time, and today we may be living in the most peaceable era in our species’ existence” (p.xix), in Enlightenment Now (2018) Steven Pinker goes on to outline positive progress in all spheres of human development. In a concluding section on the future of progress, he provides an impressive list of mighty human achievements over the centuries since the late 18th century Enlightenment. As Pinker comments:

When the Enlightenment began, a third of children born in the richest parts of the world dies before their fifth birthday; today, that rate befalls 6% of the children in the poorest parts … The world is about a hundred times wealthier than it was two centuries ago, and the prosperity is becoming more evenly distributed across the world’s countries and peoples … War between countries is obsolescent, and war within countries is absent from five-sixth of the world’s surface … Life has been getting safer in every way … Two centuries ago a handful of countries, embracing one per cent
of the world’s people, were democratic; today two-thirds of the world’s countries, embracing two-thirds of its people, are ... People are putting their longer, healthier, safer, freer, richer and wiser lives to good use (pp.322-323).

None of this good news would, of course, carry any weight with the poor people of Sub-Saharan Africa or in war-torn countries such as present day Yemen or Syria but, as Pinker shows over and over again, our news is dominated by negative images of violence, crisis and disaster whereas positive good news hardly ever gets reported. There may also be objections which point to any putative advances or examples of human progress as merely temporary blips in the long history of human folly and wickedness. Critical work by Taleb (2018) and Gray (2018) take such a pessimistic line, and point to the picture of the natural world as, in Tennyson’s words, “red in tooth and claw” with human evolution as a blind and purposeless struggle for existence.

However, Richard Dawkins’ (2017) interpretation of evolutionary history and development seems to lend more support to the Pinker thesis. Although it is now undisputed that we are “Darwinian creatures, our forms and our brains sculpted by natural selection, that indifferent, cruelly blind watchmaker” (p.34), this does not mean that our future development must be strictly determined by the blind watchmaker. Darwin had allowed for the development of moral instincts in humans which rise above the “selfishness” of our evolutionary endowment to build communities defined by trust and benevolence. Our big brains were ideally suited to the precarious and desperate struggle for existence by our ancestors as hunter-gatherers living on the Savannah plains 200,000 years ago but, as Pinker points out, once we had such brains they could then be used for purposes beyond mere survival such as making art, language, music and building settled communities characterised by laws and moral codes. Daniel Dennett (2018) puts the case powerfully in noting:

Dawkins’s title The Blind Watchmaker (1986) nicely evokes the apparently paradoxical nature of these [evolutionary] processes: on the one hand they are blind, mindless, without goals, and on the other hand they produce designed entities galore, many of which become competent artificers (nest-builders, web-spinners, and so forth) and a few become intelligent designers and builders: us (p.37).

A notable product of such design is the construction of communities governed by laws and ethical codes characterised by altruism and co-operation which rise above the brute competitiveness of the evolutionary instincts and impulses.

Jeremy Griffith (2017) shows how the cruder forms of Social Darwinism which misinterpreted ideas about the struggle for existence were gradually replaced by ideas which demonstrated how moral virtues such as altruism were more beneficial to human society than selfish competition. Dawkins (2017) explains, in what evolutionary psychologists call the environment of evolutionary
adaptedness (EEA) it is plausible that—even in a world of fundamentally selfish entities—“those individuals that co-operate turn out to be surprisingly likely to prosper” (p.58). He goes on to note that:

Brains as big as ours … can actively rebel against the dictates of the naturally selected genes that built them. Using language, that other unique gift of the ballooning human brain, we can conspire together to devise political institutions, systems of law and justice, taxation, policing, public welfare, charity, care for the disadvantaged. We can invent our own values. Natural selection gives rise to these only at second remove, by making brains that grow big. From the point of view of the selfish genes our brains raced away with their emergent properties, and my personal value system regards this with a distinctly positive sign (p.61).

Moral philosophers offer a similar story about the evolution of altruism and co-operation in human society and Harris (2010) sums up the position succinctly in observing:

Clearly, our selfish and selfless interests do not always conflict. In fact, the well-being of others, especially those closest to us, is one of our primary (and, indeed, most selfish) interests. While much remains to be understood about the biology of our moral impulses, kin selection, reciprocal altruism and sexual selection explain how we have evolved to be, not merely atomized selves in thrall to our self-interest, but social selves disposed to serve a common interest with others (pp.56-57).

All such arguments would lend further support to Singer’s original thesis, and this case can be reinforced by anthropological and historical/cultural research evidence contained in the work of Christopher Boehm (2012) and Jeremy Lent (2017) on the origins of morality in early societies. Lent described in fine detail how human cultural evolution—and crucially our mores, morals and legal conventions—was irrevocably shaped by the move from hunter-gatherer to agrarian forms of life. Beginning with the Natufian civilisation in the Eastern Mediterranean, settled communities arose in Jordan, Syria and the Lebanon in which tribes started to plant and store grain seeds, build permanent houses, and construct legal conventions concerned with property rights. As Lent summarises such developments:

the agrarian worldview transformed the hunter-gatherer’s sense of nature as a giving environment into one of a cosmos demanding far more from its human participants, giving birth to a world filled with the existential anxiety that has remained with us ever since (p.104).

Boehm’s monumental anthropological research work on moral origins traces the evolutionary development of hominids in seeking to explain how genetic and cultural factors combined to shape the emergence of co-operation, generosity and altruism. The central thesis is that:
prehistorically humans began to make use of social control so intensively that individuals who were better at inhibiting their own antisocial tendencies, either through fear of punishment or through absorbing and identifying with their group’s rules, gained superior fitness. By learning to internalize rules, humankind acquired a conscience … (2012, p.17).

In commenting upon the move from hunter-gatherer to agrarian settled communities described by Lent (2017), Boehm illustrates graphically how—through the suppression of alpha male behaviour through punishment and social ostracism—evolutionary adaptations to social and economic changes led to a move from a “wolflike or apelike ‘might is right’, fear-based social order to one also based on internalizing rules and worrying about personal reputations” (2012, p.176).

Within mainstream philosophy, Phillip Pettit, covers the same area as these evolutionary/anthropological accounts in constructing a “story” about the birth of ethics (2015). Pettit prefers to call his account a “story” on the grounds that—since we can never know for sure how ethical concepts and behaviour actually originated—any attempt to describe such origins must necessarily be imaginative, speculative and, like fictional writing, fundamentally creative. Using the counterfactual device of imagining a society without ethics (he makes an analogy with the “conjectural history of money” which is explained by its emergence from the inefficiencies of the barter or similar systems), the story aims to:

provide a naturalistic genealogy of how ethical talk could have arisen, in particular a genealogy under which ethical judgments play a role in registering bona fide aspects of the world in shaping our responses to that world.

The aim is to vindicate ethics, taken literally or realistically, in naturalistic terms. And the plan is to achieve that aim by explaining how we, the products of a natural and cultural evolution could have come to develop notions of desirability to refer to aspects of the options we face, to shape our choices between those options, and to determine our fitness to be responsible for what we do (pp.214-215).

The overriding idea highlighted throughout Pettit’s story is that there is a natural movement from the ethical language of avowals (such as pledges, promises, conventions, etc.) to active commitments and behaviour in accordance with such moral language. Such conventions emerge because only those ethical concepts supported by appropriately responsible behaviour would survive and remain conducive to the maintenance of social order and functioning.

2.3. Contemporary Ethical Trends

Many of the positive advantages such a naturalistic ethics which emerge from cultural/evolutionary developments in terms of societal cohesion and general well-being are reported in recent sociological and psychological work. In a number of writings, Oliver James (2007, 2008) has argued that levels of emo-
tional distress in industrialised, urbanised societies are much higher for English-speaking countries such as Britain, United States, Canada and New Zealand than they are in other nations such as France, Spain, Belgium, Japan and the Scandinavian states. Using the World Health Organisation (WHO) definition of emotional distress to include illnesses such as “depression, anxiety, substance abuse and impulse disorder” James (2008: p. 10) contends that—contra recent fashionable notions about genes—such distress has little genetic causation but is directly linked to both parental upbringing and the impact of “selfish capitalism” which expounds radically materialistic values in conjunction with bringing about a deterioration of income levels and working conditions for millions of ordinary people in mainly English-speaking countries over the last thirty years or so. Gerhardt (2010) presents similar arguments in her survey of the “selfish society” brought about by neo-liberal economic policies. Addictive and mindless consumption connected to growth for its own sake (or rather for the sake of a minority of rich capitalists) has brought us to the brink of disaster. She expresses this in graphic terms in saying that, over the last few decades, many people in the developed world have been:

Like children let loose in the sweet shop, we have gorged ourselves on everything we could get hold of, blissfully unaware of the true cost of our activities. We have been careless or ignorant of the impact of our behaviour on the poorest and most powerless inhabitants of the planet, on our own children, and on the environment itself (p.17).

In a similar vein, Wilkinson and Pickett (2010) have demonstrated the impact of such careless self-interest on the world’s richest nations in indicating direct correlations between inequality of income and levels of mental illness, addiction, rates of imprisonment, levels of trust and the general health and well-being of nations. In all cases the data are unequivocal: “most of the important health and social problems of the rich world are more common in unequal societies” (p.173). In more recent work (Wilkinson, & Pickett, 2018), the researchers present a summary of the main findings of their earlier work as a preface to offering more detailed and deeper explanations of the impact of inequality on societies. They remind us that their research:

showed that the populations of societies with larger income differences tend to have worse health: lower life expectancy and higher rates of infant mortality, mental illness, illicit drug use and obesity. Greater inequality also damages social relationships: more unequal societies experience more violence (as measured by homicide rates) and higher rates of imprisonment; people trust each other less and community life is weaker. Inequality also damages children’s life chances; more unequal societies have lower levels of child well-being and educational attainment, more teenage births and less social mobility (2018, p.xvii).

In the new work they explore how the psychological effects of inequality in-
crease anxiety and lead to emotional disorders and higher levels of mental illness. Echoing the perspectives offered by Pinker and Dawkins referred to earlier, Wilkinson & Pickett note that:

Our evolution has not been driven only by selective forces in the natural environment. Survival has long been about more than our ability to escape predators ... The social environment, and our relationships with others, have also been selective forces (ibid., p.116).

In a similar vein, applying recent developments in cognitive neuroscience to learning theory, Goshwami (2008), argues that “learning is social” and explains that:

We have social brains. The wealth of studies of infant and animal cognition are showing more and more clearly that the complex mammalian brain evolved to flourish in complex social environments (p.391).

Thus, to return to Singer’s original agent in the shallow pond experiment, we might say that her social brain—constructed by generations of evolution—was in large part responsible for her efforts to save the child from drowning. Such acts—we might call them acts of kindness or compassion—are now supported by extensive psychological evidence that points to important reciprocal benefits for both givers and receivers. As Phillips & Taylor (2010) note: “academic studies of ‘what makes people happy’ show kindness registering much higher on the happiness scale than self-focused behaviour” (p.1). Being kind to others has significant and long-lasting positive effects on mental well-being (Hall, 2017) thus—in addition to the moral arguments for helping others dominant in both secular and religious ethical systems—the traditional maxim that “virtue is its own reward” seems to have much to recommend it.

2.4. Literary and Aesthetic Perspectives

Megan Laverty (2019) has noted the fact that “virtue ethicists assume that literature is a legitimate and, in some cases, a superior enactment of philosophical enquiry” (p.576). As Martha Nussbaum (1990) argues, “certain truths about human life can only be fittingly and accurately stated in the language and forms characteristic of the narrative artist” (p.5). Laverty emphasises the educational importance of the bildungsroman “the novel of self-formation” and highlights a number of prominent writers who have utilised this genre including Goethe, Dilthey, Rousseau and Woolf ((Laverty, 2019: pp. 576-577). Phillips and Taylor’s “short history of kindness” (2010, pp.15ff.) makes reference to a range of poets and novelists who have been concerned to emphasise the benevolent aspect of human nature. Rousseau’s Emile or Education is mentioned in particular for its identification of fellow feeling with natural instincts. The arch-pessimist philosopher, Schopenhauer (2014 edn), had described the human condition as one of endless suffering brought about by an insatiable and blind striving will resulting in an ethics in which:
the appropriate form of address between man and man ought to be, not monsieur, sir, but fellow sufferer, compagnon de misères. However strange this may sound it corresponds to the nature of the case, makes us see other men is a true light and reminds us of what are the most necessary of all things: tolerance, patience, forbearance and charity, which each of us needs and which each of us therefore owes (p.18).

Rousseau’s (1966 edn.) more optimistic interpretation of the human condition causes him to recommend the cultivation of compassion as a key component in the moral education of his ideal pupil Emile. As he writes:

At sixteen the adolescent knows what it is to suffer, for he himself has suffered … when his imagination is kindled by the first beginnings of sensibility, he begins to perceive himself in his fellow-creatures, to be touched by their cries, to suffer in their sufferings. (p.183).

Thus results in the maxim that “it is not in human nature to put ourselves in the place of those who are happier than ourselves, but only in the pace of those who can claim our pity” (ibid., p.184).

Such other-regarding virtues are expressed graphically throughout the writings of Dickens who proposes forms of Christian kindness as a panacea for a whole panoply of human ills and evils. From the heartless factuality of the pettifogging schoolmaster, Mr Gradgrind, in Hard Times, to the cruelty and injustice of the social system portrayed in Oliver Twist and David Copperfield, Dickens depicts the triumph of the superior characteristics of charity, compassion and kindness in human affairs. And of, course, charity and compassion are prominent in the transformation of Scrooge in what is perhaps Dickens’ most famous portrait of the power of kindness in A Christmas Carol.

Underpinning kindness and compassion is the move from self-regarding to other-regarding dispositions, and this is highlighted in both consequentialist (utilitarian) and deontological (duty) ethical systems. It is also prominent in Eastern philosophical/spiritual traditions such as Buddhism, and the novelist, Iris Murdoch, combined both Eastern and Western perspectives in her writings about the perennial human quest to attain moral goodness through “unselfing” (Hyland, 2020). Unselfing is viewed as both a goal and as a process “wherein one learns to see, and cherish and respect, what is not himself” (Murdoch, 2003: p. 17). Olsson (2018) suggests that central to this process is the concept of attention which is described by Murdoch (1998) as a “just and loving gaze directed upon an individual reality” (p.327). It is here that the links between education, morality and unselfing are brought into prominence. As Murdoch writes in her well-known thesis which emphasises the sovereignty of good over other concepts:

If I am learning, for instance, Russian, I am confronted by an authoritative structure which commands my respect. Love of Russian leads me away from myself towards something alien to me … The honesty and humility
required of the student—not to pretend what one does not know—is the preparation for the honesty and humility of the scholar who does not even feel tempted to suppress the fact which damns his theory (ibid., p. 373).

But it is not just the qualities developed through such an attitude to learning but the very process of attending to something beyond and independent of our own egos and concerns which makes learning of such significance. As Olsson argues, the attention that Murdoch recommends invites us to move beyond our obsessions as “narrating creatures” by becoming “wrapped up in something” which involves the “experience of spontaneity, immediacy, being touched and moved”. Such a process may be “induced by a careful attention, freed of requirements, and by the concrete experience of the world” (2018, p.173). Murdoch (2003) agrees with both Kant and Schopenhauer that the pursuit of aesthetic knowledge and experience is a moral enterprise which may help to overcome the egotistic impulse by “inducing, at least a temporary, state of selflessness” (p.179). As she goes on to observe:

Any artist, or thinker, or craftsman knows of crucial moments when an aggregate of reflection and skill must now be pressed a little harder so as to achieve some significantly better result … Ideas break the narrow, self-obsessed limits of the mind. The enjoyment and study of good art is enlarging and enlightening in this way. We may add to this that as mathematics “stands for” any high intellectual discipline, we may, without breaking faith with Plato, suggest that the carpenter “stands for” any careful attentive self-forgetting work or craft, including housework, and all kinds of nameless “unskilled” fixings or cleanings or arrangings which may be done well or badly (ibid., 179-180).

The idea of particular forms of learning and training of the mind as means of escape from the self and suffering was a principal feature of Schopenhauer’s philosophy (1995 edn.) and is, of course, also central to Buddhist teachings, especially in the increasingly popular mindfulness strategies informed by Buddhist philosophy (Hyland, 2011; Ergas, 2019). It is possible that, like Schopenhauer, Murdoch gained insights about the importance of particular forms of training attention from such Buddhist sources since there are remarkable parallels between—as Olsson and others have noted (Mole, 2006)—the centrality of the role accorded to selfless attention in her general philosophy and the idea of Buddhist mindfulness as the “self-regulation of attention so that it is maintained on immediate experience … an orientation that is characterised by curiosity, openness and acceptance” (Bishop, 2004: p. 232).

The notions of kindness and helping others as supreme virtues are also to be found in the most famous and well-loved poetry. In this realm, the romantic nature poems of Wordsworth are paradigm illustrations, and perhaps the most well-known example is the sublime nature-lover’s philosophy of Tintern Abbey (Wordsworth, 1987: pp. 129-130) in which we read the lines:
Though absent long, /These forms of beauty have not been to me
As is a landscape to a blind man’s eye … In hours of weariness sensations
sweet
Felt in the blood, and felt along the heart,/And passing into my purer mind
With tranquil restoration—feelings too/Of unremembered pleasure: such,
perhaps,
As may have had no trivial influence/On the best portion of a good man’s
life,
His little nameless unremembered acts of kindness and of love.

Perhaps the most famous poem of all in this sphere—beloved by practi-
cioners of Eastern spirituality—is that entitled Kindness by Naomi Shihab
Nye (https://onbeing.org/poetry/kindness/)

Before you know what kindness really is
you must lose things,
feel the future dissolve in a moment
like salt in a weakened broth.
What you held in your hand,
what you counted and carefully saved,
all this must go so you know
how desolate the landscape can be
between the regions of kindness.
How you ride and ride
thinking the bus will never stop,
the passengers eating maize and chicken
will stare out the window forever.

Before you learn the tender gravity of kindness,
you must travel where the Indian in a white poncho
lies dead by the side of the road.
You must see how this could be you,
how he too was someone
who journeyed through the night with plans
and the simple breath that kept him alive.

Before you know kindness as the deepest thing inside,
you must know sorrow as the other deepest thing.
You must wake up with sorrow.
You must speak to it till your voice
catches the thread of all sorrows
and you see the size of the cloth.

Then it is only kindness that makes sense anymore,
only kindness that ties your shoes
and sends you out into the day to mail letters and purchase bread,
only kindness that raises its head
from the crowd of the world to say
It is I you have been looking for,
and then goes with you everywhere
like a shadow or a friend.

2.5. Coda: Shallow Pond

Singer’s Shallow Pond thought experiment about our human instincts and dispositions in relation to kindness, compassion and helping others has been shown to be immensely powerful when supplemented by social scientific, philosophical and aesthetic sources. The importance of the genetic/evolutionary underpinnings of the altruistic human traits and dispositions have been graphically illustrated (alongside with, to a lesser degree, basic selfishness and fear-inspired greed) in the recent near-universal reaction to the Coronavirus/COVID-19 pandemic which has swept the world since January 2020 (https://people.com/human-interest/positive-stories-from-coronavirus-crisis/).

Ranging from unprecedented government support for firms and employees, global public acclaim for healthcare and other key workers, to the spontaneous blossoming of neighbourhood community projects, the crisis seems to have brought out the sort of behaviour recommended and celebrated in both the aesthetic and philosophical/sociological ethical accounts outlined above. The naturalistic ethics which characterise our impulses to save the child drowning in the pond are the product of generations of cultural and evolutionary forces which continue to work towards the functioning of communities committed to human flourishing and well-being.

3. Philippa Foot’s Trolley Problem

Foot’s original thought experiment was outlined in an article on abortion and the double effect principle. She notes:

Suppose that a judge or magistrate is faced with rioters demanding that a culprit be found for a certain crime and threatening otherwise to take their own bloody revenge on a particular section of the community. The real culprit being unknown, the judge sees himself as able to prevent the bloodshed only by framing some innocent person and having him executed. Beside this example is placed another in which a pilot whose aeroplane is about to crash is deciding whether to steer from a more to a less inhabited area. To make the parallel as close as possible it may rather be supposed that he is the driver of a runaway tram which he can only steer from one narrow track on to another; five men are working on one track and one man on the other; anyone on the track he enters is bound to be killed. In the case of the riots the mob has five hostages, so that in both the exchange is supposed to be one man’s life for the lives of five. The question is why we should say, without hesitation, that the driver should steer for the less oc-
cupied track, while most of us would be appalled at the idea that the inno-
cent man could be framed (1967, p.3).

The principle of double effect comes into operation due to the fact that, in
choosing the lesser of two evils, we cannot but avoid immoral acts. However, in-
tentions are crucial in these cases, the pilot does not intend to crash the plane no
more than the tram driver is aiming to kill a man working on the track. Unlike
the judge in Foot’s example, they are presented with a fait accompli—if they
don’t act, more people will be killed than if the agents choose other courses of
action.

Foot’s trolley problem has achieved unprecedented fame and celebrity within
philosophical ethics leading to a specialised domain known as “Trolleyology”
(Danaher, 2014). Danaher expresses the key dilemmas in his outline of the two
“classic” versions of the problem (p.1):

**Switch:** A trolley car is hurtling out of control down a train track. If it con-
tinues on its current course, it will collide with (and kill) five workers who
are on the track. You are standing beside the track, next to a switch. If you
flip the switch, the trolley will be diverted onto a sidetrack, where it will col-
lide with (and kill) one worker. Do you flip the switch?

**Footbridge:** A trolley car is hurtling out of control down a train track. If it
continues on its current course, it will collide with (and kill) five workers
who are on the track. You are standing on a footbridge over the track, next
to a very fat man. If you push him off the footbridge, he will collide with the
trolley car, slowing it down sufficiently to save the five workers. He, how-
ever, will die in the process. Do you push the fatman?

In psychological experiments 87% of respondents said they would flip the
switch but only 31% would be willing to push the fat man (ibid., p.2). It seems to
be that our moral intuitions make us reluctant to engage personally with obvi-
ously immoral actions whereas remote action at a distance is apparently more
acceptable. Variations on the original experiment have involved making the one
worker on the track a young person (or even a close relative) whereas the five on
the other track are old people and the resulting responses vary predictably.

It is worth pointing to a key difference between this thought experiment and
the Shallow Pond discussed above. Obviously, Singer’s example is more realistic
but, more significantly, it taps into our natural moral instincts about
other-regarding values in a way in which the Foot experiment does not. Pushing
fat men off railway bridges is, after all, quite some way from normal, everyday
experiences whereas saving people from drowning does happen occasionally.
Considering all this, therefore, against the background of Wilson’s appeal for re-
alism in moral thought experiments, it might be suggested that “Trolleyology” is
simply too fantastic and far too remote from real life. Some moral philos-
ophers—notably R.M. Hare (1981) who was a lifelong proponent of a nuanced
form of utilitarianism—have suggested just this, and expressed indignation that
such fantastic examples can be used to support arguments against the principles
of everybody counting for one and considering the greater interests of the ma-
jority in moral judgments. He suggests that our intuitions are not designed to
cope with fantastically rare experiences; he comments that:

Undoubtedly critics of utilitarianism will go on trying to produce examples
which are both fleshed out and reasonably likely to occur, and also support
their arguments. I am prepared to bet, however, that the nearer they get to
realism and specificity, and the further from playing trains—a sport which
has had such a fascination for them—the more likely the audience is, on re-
fection, to accept the utilitarian solution. I am thinking of their examples in
which trolleys hurtling down the line run out of control and, therefore, have
to be shunted into various alternative groups of unfortunate people. I have
myself, when helping to build a railway, seen trolleys run out of control,
and therefore find the unrealism of the examples very obvious (1981,
p.139).

Bernard Williams, on the other hand (a critic of utilitarianism), viewed our
moral intuitions—defined as "spontaneous convictions, moderately reflective
but not yet theorized, about the answer to some ethical questions" (1990,
pp.94-95)—as legitimate subjects for consideration by ethicists. Similarly, Joshua
Greene (2014) has examined the Trolley problem and similar experiments in
great detail and views them as being extremely valuable in illuminating the deep
psychological mechanisms of moral choice. They illustrate, in the moral sphere,
elements of the fast and slow thinking identified by Daniel Kahneman (2011). In
this sense, our intuitions are examples of System 1 thinking—which “operates
automatically and quickly”—and deeper moral values and beliefs would thus be
examples of System 2 thought which “allocates attention to the effortful mental
activities that demand it … often associated with the subjective experience of
agency, choice and concentration” (pp.20-21). Kahneman shows how both sys-
tems are required for the demands of contemporary living, and there is a ne-
cessary complementarity to their interactions in response to different activities.
Hare (1981) similarly shows how intuitive and critical, rational moral thinking
are not “rival procedures” but “elements in a common structure, each with its
part to play” (p.44). In order to illustrate this process—and respond to the de-
mands of Wilson, Hare and others—it is worth seeking to enhance the potential
realism of the Trolley scenario by means similar to those employed above in re-
lation to Singer’s example.

3.1. Social Science Perspectives

The key principle underpinning the “double effect” scenario of the Trolley prob-
lem is that of sacrificing a lesser number of lives in order to save a greater num-
ber, what may be described—as mentioned above—as justifying an unethical act
by reference to the lesser of two evils. There would be quite a number of putative
historical examples of this principle in action. A well-known case involves
Winston Churchill’s sacrifice of the 51st Highland Division to allow for the evacuation of the British Expeditionary Force (BEP) from Dunkirk in May/June 1940. The Division was ordered by Churchill to fight on against German forces which massively outnumbered them, resulting in the whole division being either killed or captured (10,000 men from the division were imprisoned for the rest of the war; Macpherson, 2017). Thus, any moral justification here—and this was indeed offered in narratives around the “miracle of Dunkirk” (Latimer, 2003)—were ultimately couched in terms of sacrificing 10,000 men to save 400,000 who might survive to subsequently continue the war with Germany.

This historical case is, however, significantly different from the Foot narratives. Churchill’s decision was not made in haste, on the basis of intuitive instinct, but after reflecting on a number of alternative possibilities for saving the BEP. Moreover, if there had been no miracle of Dunkirk, the decision might have been recorded as a war crime (as, indeed, the bombing of Dresden in 1944 is so described by certain historians). The old adage that history is written by the victors clearly applies to this and many other similar historical examples of sacrificing the few for the many. Take the dropping of the atomic bomb on Hiroshima in 1945 for example. The annihilation of 125,000 Japanese citizens in Nagasaki and Hiroshima in August 1945 is usually justified by the argument that this was the only way to end a war which otherwise would have resulted in the deaths of many hundreds of thousands of soldiers. Yet, as Nathan Robinson (2016) notes, many politicians, soldiers and historians have claimed that the Japanese were ready to surrender anyway, and that the horrible deaths of so many Japanese citizens cannot be measured against the deaths of active soldiers. As Robinson reports, Eisenhower, General MacArthur and Herbert Hoover all “believed the bombings to be horrendous and a mistake” (p.1).

It may well be that utilitarian arguments which seek to justify the sacrifice of human lives in terms of the greater good will always seem controversial and unsound when such tragically large numbers of people are involved as in the cases outlined here. Moreover, such examples cannot be untangled from conflicting views about the morality of warfare, particularly so in the light of historical revisionism. Where such arguments—and, thus, the feasibility of the double effect principle underpinning the Foot narrative—seem to be far more plausible is when they are applied to the public allocation of resources and decision-making in areas such as bio-ethics, urban planning and medical practice (see Hare, 1999).

3.2. Medical Ethics and Coronavirus

Medical ethics provides many fruitful instances for the application of principles based on the greatest good and the maximisation of the needs and interests of the many rather than the few. As Peter Mack (2004) argues, in healthcare systems which are constantly seeking to evaluate the benefits of ever-expanding procedures for prolonging life within a limited cost framework the value of
utilitarian ethics is indisputable. As he concludes:

The moral economic agent in action is not only individual, but social, wanting not only health provision but sharing them, not only pricing health services, but valuing them, self-interested but yet bonded to the greater social good (p. 71).

The controversial politics surrounding the contemporary Coronavirus health crisis provides some graphic illustrations of the complexities surrounding the application of utilitarian moral judgments. Parker & Mirzaali (2020) have recently analysed the key ethical issues in the current debate. Both Italy and the USA have clearly adopted the utilitarian principle of maximising benefits for the largest number of patients resulting in the allocation of resources favouring “those patients with the highest chance of therapeutic success” (p.1). Moreover, even though the deontological approach of impartial duty is officially advocated by the General Medical Council in Britain and similar codes across Europe, actual practice in the COVID-19 emergency cannot but avoid the rationing of resources. Thus, older people have tended to be treated less favourably than younger ones, and many patients suffering from cancer and heart problems have had treatment deferred in favour of COVID sufferers.

This state of affairs merely brings into sharper focus the inevitability of the differential allocation of treatments and services against the background of limited and finite economic and supply resources. It is easy to overlook that this means the sacrificing of some lives to save others to a degree which is no less stark than the decisions made by agents in the Trolley Problem. As Manelli (2020) explains, prioritisation in these situations is inevitable but:

Prioritisation does not mean that one life is more valuable than another, as all lives are equally valuable. But when resources are not enough to save all those in need, prioritisation involves allocating resources such that they are more likely to save the most lives. This method allows priority treatment of those who are more likely to benefit from the scarce resource—admission to an ICU in this case—and to recover quickly with a positive outcome, which in turn allows the next in line to benefit from the treatment in question (p.1).

Moreover, such prioritisation along generally utilitarian lines—whether this takes place in medicine, the provision of education or social welfare—is bound to disadvantage some sections of the population against others. But, as with the Trolley Problem, we may justify the clear immorality of disfavouring the lives and life chances of some people by the ethical principle of promoting the greater good.

In terms of sacrificing lives to save others, the many examples of self-sacrifice would, arguably, meet with more immediate moral approval since such acts of supererogation—unlike those of agents in the Trolley state of affairs—do not require any loss or endangering of lives other than the person making the
self-sacrifice. Karl Smallwood (2013) cites the examples of the Russian scientists working at the Pavlosk seed bank during the 1941 siege of Leningrad who chose to starve rather than eat the plants being grown for their fellow citizens. Another example he reports is that of Takeshi Miura and Miki Endo, who were two government risk management workers at the time of the 2011 Japanese tsunami tasked with warning and directing the public to safety. When the ten metre wave hit, both Takeshi and Miki stuck their posts and kept broadcasting, using their last words to direct townsfolk to safety. History is full of such examples—from actions beyond the call of duty during wartime, to the deprivations suffered by those campaigning for workers’ rights in Victorian Britain, to the tribulations of suffragettes fighting for justice for women in the early 20th century—and many of them involve individual acts of moral supererogation which appear to directly contradict human evolutionary survival instincts (Mihara, 2017).

3.3. Literary and Aesthetic Perspectives

It should be noted at the outset that within this domain the idea of sacrificing lives to save a greater number of others generally tends to take the form of self-sacrifice (as in the historical examples outlined above) or through the cultivation of compassionate virtues which allow us to cope with the terrible contingencies of life which make such grief and suffering a feature of the human condition. In this latter sphere the wonderful poems of Rumi are worth noting. For instance, in The Guest House (Helminski, 2000: pp. 187-188) we read the following:

This being human is a guest house.
Every morning a new arrival.
A joy, a depression, a meanness,
some momentary awareness comes
as an unexpected visitor.

Welcome and entertain them all!
Even if they are a crowd of sorrows,
who violently sweep your house
empty of its furniture,
still, treat each guest honorably.
He may be clearing you out
for some new delight.
The dark thought, the shame, the malice.
meet them at the door laughing and invite them in.

Be grateful for whatever comes.
because each has been sent
as a guide from beyond.

—Jellaludin Rumi
In terms of a compassionate response to life’s vicissitudes, it is well worth attending to the wonderful poems of Buddhist teacher, Thich Nhat Hanh. The concluding verses of Please Call Me by My True Names, for instance, sum up beautifully the awful contingencies of human pain and suffering in the world (Hanh, 2002: p. 72).

I am the child in Uganda, all skin and bones, my legs as thin as bamboo sticks. And I am the arms merchant, selling deadly weapons to Uganda.

I am the twelve-year-old girl, refugee on a small boat, who throws herself into the ocean after being raped by a sea pirate.

And I am also the pirate, my heart not yet capable of seeing and loving.

I am a member of the politburo, with plenty of power in my hands. And I am the man who has to pay his “debt of blood” to my people dying slowly in a forced-labor camp.

My joy is like Spring, so warm it makes flowers bloom all over the Earth.

My pain is like a river of tears, so vast it fills the four oceans.

Please call me by my true names, so I can hear all my cries and laughter at once, so I can see that my joy and pain are one.

Please call me by my true names, so I can wake up and the door of my heart could be left open, the door of compassion.

In terms of self sacrifice, perhaps the most famous example in literary history is that of Sydney Carton in Dickens’ A Tale of Two Cities. In the concluding, rather pompous and melodramatic, monologue by Carton as he faces death on the scaffold during the French Revolution:

I see … long ranks of the new oppressors who have risen on the destruction of the old, perishing by this retributive instrument … I see the lives for which I lay down my life, peaceful, useful, prosperous … It is a far, far better thing that I do, than I have ever done; it is a far, far better rest that I go to than I have ever known (1970 edn, pp.436-437).

Of course, Dickens’ politics, like many British people of the time, were coloured by the fear that the destruction of the French monarchy and the establishment of a secular Republic would spread to England. Carton’s sacrifice must be understood in this light and, of course, its significance would not be shared by those people labouring under centuries old poverty and injustice. In the maximisation of interests and needs which favour the many rather than the few, it will always be necessary to take into account the precise characterisation of the many and what is exactly in their interests. It is precisely in highlighting such considerations that the Trolley Problem performs important philosophical tasks.

Other literary examples of sacrificing the one for the many raise similar issues and concerns. Winston Smith’s campaign against Big Brother in George Orwell’s 1984, for instance, represents an atypical anti-hero who opposes the totalitarian regime in the hope of securing a more widespread revolt against the authoritarian thought control of his society. Somewhat less nuanced, more unequivocal
examples of altruistic sacrifice in the interests of the many would be that of Dorothea Brooke in George Eliot’s *Middlemarch* who gave up a privileged bourgeois life to care for the poor and disadvantaged, and the painters and decorators in early 20th century England who, in spite of constant hardship and fear of dismissal, struggled to support each other and their craft as described by Robert Tressell in *The Ragged Trousered Philanthropists*. Such altruistic behaviour—and, of course, the struggle between good and evil which underpins all such sacrifice—has been a common trope in literature since the birth of the novel in the 17th century and continues to inspire writing at all levels (Brie & Rossiter, 2010).

### 3.4. Coda: Trolley Problem

The surreal and other-worldly nature of Foot’s Trolley Problem tends to render the search for moral realism in this case more difficult than that of the Singer thought experiment which rests much closer to natural instincts and intuitions. However, the examples from history, medicine, social science and literature outlined above have served to illustrate both the potential force and widespread application of the dilemmas posed when humans have to balance lives and values against each other. The problem brings out clearly the tension between the fast thinking of our intuitions (it is wrong to take any lives since we are all of equal value) and the slower, more critical and reflective, level of moral thought in which we have to act to save as many lives as possible. Such considerations highlight the intricacy and complexity of moral judgments and help us to come to terms with the fact that some ethical problems do not lend themselves to simplistic black-and-white decision-making. The illumination of these grey areas is an important philosophical and therapeutic process.

As noted above, the current COVID-19 health crisis brings out many of these dilemmas and provides a real-world graphic illustration of the importance of different levels of thinking outlined in various ways by Hare and Kahneman, and demonstrates the irrationality of always expecting simplistic solutions which accord with our intuitions about basic duties and values. Of course, we would naturally like to save all the lives we can in the Coronavirus epidemic. As the Amnesty International organization expresses this: “No one’s safe until everyone’s safe” ([https://www.amnesty.org/en/get-involved/covid-19/](https://www.amnesty.org/en/get-involved/covid-19/)). This captures perfectly our fundamental moral values and intuitions but—on more critical reflection—we must see that not everyone can be safe and that health services have to prioritise resources and services in a way which favours some lives over others. Moreover, in terms of the self-sacrifice sentiments noted earlier, citizens globally have been surrendering freedoms and rights (of course, not always willingly!) in the interests of the many. The compassionate dispositions are naturally forthcoming as illustrated in the universal support for health and key workers who are putting themselves in harm’s way to maintain essential services.
4. Conclusion: Enhancing Moral Education and Learning/Teaching in Philosophical Ethics

It has been suggested that both the Singer and Foot thought experiments can be substantially reinforced in terms of the realism recommended by Wilson through the incorporation of supporting data and examples from literature and social science. The fact that they are cogent and powerful tales—even in their most abstract and minimalist forms—is revealed by their immense popularity in the academic literature. We may ask, therefore, why they are not always efficacious in achieving their aims and why they need any further support. The observations of G.J. Warnock (1967) are relevant here:

That moral argument is not more effective than we find it to be is probably attributable to the cross that all arguments have to bear: an argument offers reasons to people, and people are not always reasonable (p. 72).

The supporting material aggregated above is intended to enhance the moral realism—and hence the efficacy, reasonableness and potency—of the thought experiments. As such, such material might be used to support learning/teaching in moral education and ethics at both school and post-school levels, perhaps in conjunction with the secular values incorporated in the currently popular mindfulness principles and practice (Hyland, 2013, 2014).

As noted at various points above, the values uncovered in both the Singer and Foot stories are rooted in our basic instincts emerging from the genetic and cultural evolutionary history of humankind. There are connections here with Donald Hoffman’s (2019) arguments concerning the fact that our “reality” is constructed on the basis of evolutionary fitness. The central thesis is that “Fitness-Beats-Truth” which states that:

Evolution by natural selection does not favour true perceptions—it routinely drives them to extinction. Instead, natural selection favours perceptions that hide the truth and guide useful action … Space, time and physical objects are not objective reality. They are simply the virtual world delivered by our senses to help us play the game of life (pp.xiv-xv).

Hoffman, thus, has to tackle the problem of whether and how it may be possible to see through the “virtual” world to achieve a glimpse of “reality”. The solution to this problem is advocated as “conscious realism” which posits the notion—starting with the most fundamental and incorrigible facts of human consciousness—that as conscious agents we participate in a social network of other conscious agents who “combine to create more and more complex agents” (ibid., p.199). Thus, Hoffman presents us with an idealist variant of the panpsychism which, for example, Galen Strawson (2016) proposes to solve the so-called “hard problem” of consciousness which consists in the difficulty of accounting for our thoughts, images, feelings—and generally what it is like to be us—in a world which science tells us contains only physical/material objects.
If we accept the objective reality of consciousness as advocated by Hoffman, Strawson, Harris (2014) and others, the moral values identified in the issues discussed above can be seen to reinforce our genetic and evolutionary cultural heritage. If reality is constructed by a social network of conscious agents, the existence and efficacy of common values—respect for life, compassionate dispositions towards the suffering of fellow conscious beings—would be exactly what would be expected to predominate universally. Our ancestral heritage in this lineage of conscious beings is what serves to construct and maintain collective moral convergence and inter-subjective ethical consensus.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References


Harmondsworth: Penguin.


