

# 4.

## THE GOLDEN RULE IN KANT AND UTILITARIANISM

DANIEL BERTHOLD  
BARD COLLEGE

From the vantage point of the history of ethical theory, there can be little doubt that in the modern period two philosophies stand out as by far the most important in terms of the scope of their influence: Kantianism and utilitarianism. These two philosophies offer fundamentally contrasting ethical theories: one emphasizes the intentions behind our actions, the other the consequences; one seeks the justification for moral action in duty for its own sake, the other in the maximization of human happiness; one claims to establish absolute ethical imperatives, the other states that all ethical action is situational and all rules provisional; one is unrelenting in excluding desire, interest, and emotion from moral deliberation, while the other sees these as essential. Kant is as remorseless in his rejection of “doctrines of happiness”<sup>1</sup> as effecting “the euthanasia of

---

<sup>1</sup> Kant’s comments about the “doctrine of happiness,” or “eudaimonism,” begin around 1780, some nine years before publication of Jeremy Bentham’s *Principles of Morals and Legislation*, which is often taken to mark the origin of utilitarianism. But Bentham was by no means the first to speak of the principle of “the greatest happiness for the greatest number.” This exact phrase, as well as a philosophical theory which defended the maximizing of happiness as the primary goal of ethics, was already present in the writings of such philosophers and political theorists as Francis Hutcheson (*Inquiry Concerning Moral Good and Evil*, 1725), Cesare, Marquis of Beccaria (*On Crimes and Punishments*, 1764), and Joseph Priestly (*The First Principles of Government and the*

all morality”<sup>2</sup> as John Stuart Mill is in his attempt to expose how Kant “fails, almost grotesquely,” to deduce any reliable moral duties from his abstract theory.<sup>3</sup>

Yet for all their differences, Kantianism and utilitarianism have one thing in common: they are both regularly described as seeking to incorporate the insight of the golden rule, that “in everything, do unto others what you would have them do to you” (Matt 7:12), and that “Thou shalt love thy neighbor as thyself” (Gal 5:14).<sup>4</sup> Kant’s central principle that the imperatives of morality be “universalized” so as to apply to all, and that in all our actions we respect the other as an “end in itself” and never merely as a means to our own desires, and utilitarianism’s stress upon maximizing the interests and happiness of the greatest number, clearly seem resonant with the spirit of the golden rule.

Kantianism and utilitarianism are often presented as endeavoring to correct certain problems entailed by the vagueness of the golden rule (since the rule does not actually specify what we should do), and as offering contrasting reformulations of the golden rule in purportedly more adequate ways.

---

*Nature of Political, Civil and Religious Liberty*, 1768) – all of whom Bentham credited with influencing his own doctrine, and all of whose ideas Kant was familiar with.

<sup>2</sup> Kant, *The Metaphysical Elements of Ethics* (1780), tr. Thomas Kingsmill Abbott (London: Kongmans, Green, 1889), 19. Henceforth MEE. The English translation of this work has proven extraordinarily difficult to get hold of. References are thus to the online text provided by Marxists.org (2003) at <http://www.marxists.org/reference/subject/ethics/kant/morals/ch01.htm>. See also Kant, *Foundations of the Metaphysics of Morals* (1785), tr. Lewis White Beck (Indianapolis: Bobbs-Merrill, 1969), 395-96. Henceforth FMM. Page references are to the standard Königliche Preussische Akademie der Wissenschaften edition of *Kants gesammelte Schriften* (Berlin: G. Reimer, 1902-1942).

<sup>3</sup> Mill, *Utilitarianism* (1861), ed. Samuel Gorovitz (Indianapolis: Bobbs-Merrill, 1971), 15. Henceforth U.

<sup>4</sup> The formulation of the golden rule most commonly found in the gospels, to “do unto others ...,” is often traced back to the Torah formulation of “love thy neighbor,” given Christ’s frequent association (twice in Matthew, 19:19 and 22:39; twice in Mark, 12:31

There is certainly something to this characterization, yet in what follows I will show that in each case, the reformulation fails to capture something important about the rule. In Kant's case, what is lost is the commandment to *love* (love, and indeed all emotive content, is stripped away in Kant's philosophy as a "pathological" element threatening the purity of moral duty). In the case of utilitarianism, the rule's suggestion of the sanctity of each individual "other" seems to be diminished by the aggregative character of the principle of utility, by which we must calculate the best course of action for the greatest number, inevitably at times requiring the sacrifice of individual interests. Similarly, utilitarianism's exclusion of intentions or motives as essential for determining the moral worth of an action, and the primary importance it places on pleasure and happiness, often seem at odds with the spirit of the golden rule. The importance of studying Kant's and utilitarianism's reformulations of the golden rule is thus two-fold: as reformulations, each calls attention to a certain elusiveness of the rule which invites the attempt to give a theoretical framework for interpreting its power; yet in seeing how these theoretical frameworks each purchase greater explanatory value for the rule at the cost of certain values which seems central to it (whether it be love or the irreducible worth of the individual), they invite us to return to the rule with a greater sensitivity to its core values.

I will begin by sketching out a number of problems that are commonly agreed to attend the golden rule, as a starting point for examining Kant's and utilitarianism's reformulations. I will use Freud's trenchant critique of the rule in his *Civilization and its Discontents* as a point of departure, not because Freud poses these problems in a particularly profound way – indeed his remarks are mainly polemical – but because he

---

and 12:33; and once in Luke, 10:27) of the commandment to "do unto others" with the

couches his criticisms in an historically influential psychological theory of human nature. This is important, since Kant's and utilitarianism's responses to the golden rule differ largely on grounds of psychology. While the utilitarians straightforwardly base their moral philosophy in a psychological theory of human nature – so that their replies to Freud will be on the same terrain – Kant in fact outright rejects psychology as having any important role in developing an ethics.<sup>5</sup> Freud's objection that proponents of the sort of love enjoined by the golden rule operate in an entirely “unpsychological” fashion – which he takes to be decisive grounds for rejecting the rule<sup>6</sup> – would be taken by Kant, odd though it may seem, as a mark of the *value* of this approach, since for him, this sort of love is not “pathological” (not really an emotion at all), but simply “disinterested benevolence” or “philanthropy,”<sup>7</sup> which is only “very improperly” called love (MEE, 19).

After presenting Freud's criticisms, I will turn to the philosophies of Kant and Mill, tracing out their contrasting projects of recasting the golden rule. Kant's position, we will see, is at least initially more complex, since he explicitly raises serious objections to the rule – unlike Mill, who has nothing but praise for it. Hence if Kant's philosophy can legitimately be seen as a reformulation of the golden rule at all (as I think it can), it can only be in such a way as to effect quite radical changes. The position of utilitarianism,

---

Law, the commandment to love thy neighbor, in the Torah (Lev 19:18).

<sup>5</sup> See, e.g., FMM 388, where Kant argues that we must “carefully purify” moral theory “of everything empirical so that we can know how much pure reason can accomplish” without recourse to such sciences as anthropology or psychology. See also Kant's *Critique of Practical Reason* (1788), tr. Lewis White Beck (Indianapolis: Bobbs-Merrill, 1975), 7-8 (Akademie pagination, see n2 above).

<sup>6</sup> Freud, *Civilization and its Discontents* (1930), tr. James Strachey (New York: Norton, 1961), 90. Henceforth CD.

<sup>7</sup> Kant, *The Metaphysical Principles of Virtue (Part II of The Metaphysics of Morals)* (1797), tr. James Willington (Indianapolis: Bobbs-Merrill, 1964), 450. Henceforth MPV. Page references are to the standard Akademie edition (see n2 above).

especially that of John Stuart Mill, initially seems more obviously in harmony with the golden rule, and while there are no points of straightforwardly serious conflicts, there are, I will argue, important commitments of utilitarianism which seem at odds with the spirit, if not the letter, of the biblical Law to love thy neighbor.

- *Freud*

In *Civilization and Its Discontents*, Freud mentions the version of the golden rule that commands us to “love thy neighbor as thyself” twice directly (56, 90), and several times indirectly.<sup>8</sup> His discussion of the rule is placed in the context of a general theory of the instincts, which in *Civilization and Its Discontents* is framed in terms of his conviction that “the meaning of the evolution of civilization” is “the struggle between Eros and Death,” the instincts of love and of destruction (69). As an instinct, Eros is straightforwardly libidinal or genital in character: love is based on the “overwhelming sensation of pleasure” afforded by sexual satisfaction (29). Fundamental to the sexual instinct is that it is directed towards an *individual* “object”; love is thus intrinsically particular and discriminatory.

The love of one’s neighbor that the prophets and Christ describe as the Law is, on the contrary, a duty to abandon the particularity of the sexual relation and to love all equally. Such a love is “aim-inhibited,” as Freud puts it: the “natural,” instinctive aim or goal of Eros – the particular other of sexual desire – must be inhibited or repressed so that the erotic instinct may be sublimated or displaced onto all others without distinction, a displacement which requires “far-reaching mental changes in the function of love” (CD 48f).

On the one hand, Freud sees this radical transformation of Eros as serving the interests of civilization, since civilization “aims at binding members of the community together” while the sexual instinct is exclusive (55f). The success of human civilization indeed largely rests upon the extent to which it can successfully effect the repression of sexual love and encourage aim-inhibited love. The golden rule, then, that “Thou shalt love thy neighbor as thyself,” is the basic ethic of human civilization. On the other hand, such a deflection of the original libidinal nature of the erotic instinct can only be experienced as a painful transgression against human nature and is bound to cause hostility to civilization: “it is impossible to overlook the extent to which civilization is built upon a renunciation of instinct, [and therefore] how much it presupposes precisely the non-satisfaction (by suppression, repression, ...) of powerful instincts” (44), a suppression which ensures that “civil society is perpetually threatened with disintegration” (59). Indeed, apart from “a small minority” – saintly types like Francis of Assisi – who are capable of superhuman efforts of instinctual renunciation (49), “the commandment, ‘Love thy neighbor as thyself,’ ... is impossible to fulfill” (90).

Freud’s view of the impossibility of achieving the demands of the golden rule clearly rests upon his theory of human nature. Love is by its very nature rooted in an instinct that is genital, particular, and fundamentally aimed at satisfying the narcissistic pleasure of the self. Moreover, the intrinsically egoistic character of love is accompanied by a second instinct, Thanatos, or the death instinct, which experiences the other as a constant threat to the satisfaction of the self’s own pleasure. The death instinct ensures that however desperately civilization or “ethics based on religion” (CD 90) may seek to

---

<sup>8</sup> Indirect references are all places where Freud speaks of “aim inhibited love,” which he

disarm Eros of its narcissistic origin in order to combat our instinctive distrust of the other, it will always come hard up against the stark truth of human nature, “*Homo homini lupus*” – man is a wolf to man (59). There is a “primary mutual hostility of human beings,” against which the “endeavors of civilization have not so far achieved very much”:

[Civilization] hopes to prevent the crudest excesses of brutal violence by itself assuming the right to use violence against criminals, but the law is not able to lay hold of the more cautious and refined manifestations of human aggressiveness. The time comes when each one of us has to give up as illusions the expectations which, in his youth, he pinned upon his fellow men, and when he may learn how much difficulty and pain has been added to his life by their ill-will. (59)

In face of this bleak psychology, the command of the golden rule to love thy neighbor seems almost silly, a childish act of closing one’s eyes to reality and just wishing or fantasizing. Freud would have been delighted to know of Kierkegaard’s (entirely approving) reading of the golden rule as assigning us the duty to love “with closed eyes” to all the “deficiencies and imperfections” of one’s neighbor.<sup>9</sup> Freud adopts a tone of bemused incredulity in speaking of this commandment:

‘Thou shalt love thy neighbor as thyself.’ [This ideal] is known throughout the world and is undoubtedly older than Christianity, which puts it forward as its proudest claim. . . . Let us adopt a naïve attitude towards it, as though we were hearing it for the first time; we shall be unable then to suppress a feeling of surprise and bewilderment. Why should we do it? What good will it do us? But, above all, how shall we achieve it? How can it be possible? My love is something valuable to me which I ought not to throw away without reflection. . . . [If my neighbor] is a stranger to me and if he cannot attract me by any worth of his own or any significance that he may already have acquired for my emotional life, it will be hard for me to love him. Indeed, I should be wrong to do so, for my love is valued by all my own people as a sign of my

---

associates with the commandment to love thy neighbor (e.g., CD 49f, 56, 59, 65, 68).

<sup>9</sup> Kierkegaard, *Works of Love* (1848), tr. David and Lillian Swenson (Princeton: Princeton U P, 1946), 131-32.

preferring them, and it is an injustice to them if I put a stranger on a par with them.

... Civilization pays no attention to all this; it merely admonishes us that the harder it is to obey the precept the more meritorious it is to do so. But anyone who follows such a precept in present-day civilization only puts himself at a disadvantage *vis-à-vis* the person who disregards it. ... At this point the ethics based on religion introduces its promises of a better after-life. But so long as virtue is not rewarded here on earth, ethics will, I fancy, preach in vain. (CD 56, 90)

Freud's psychology thus presents four challenges to any ethical theory that is grounded in an incorporation or reformulation of the golden rule:

1. Most generally, the golden rule is psychologically unrealistic – it requires a severe repression of the instinctive basis of love, which is sexual, discriminatory, and egoistic;
2. It devalues love by failing to discriminate between those whom I actually have affection for (or desire) and those I do not – it makes me “throw away” my love by giving it to everyone, with my eyes closed;
3. It is unjust, since for the most part people are “wolves” who do not deserve my love; and
4. It is utterly impractical, since it puts us at a disadvantage with respect to those (the great majority) who do not act likewise.

Each of these challenges are ones that Kant and Mill seek to meet, and we may add two others which Freud does not explicitly mention but which are frequently raised as key objections to the golden rule, and which are certainly consistent with Freud's general skepticism about the rule:

1. The commandment to love thy neighbor is simply too vague to yield any concrete directives for action. As Joel Marks puts it, “the rule offers no guidance regarding right and wrong behavior whatever.”<sup>10</sup>
2. The idea of “doing unto others as you would have them do to you” raises fundamental questions about the meaning of desires, wants, and values that the rule remains silent about: does the rule mean that I should treat others according to my own subjective wants, thus imposing my own standpoint on the other? Or should I treat others according to my understanding (however inevitably imperfect) of what *they* would want?<sup>11</sup> In either case, there will be unavoidable moral dilemmas, since either I will be imposing my own desires and values on others who may find them to be distasteful or even abhorrent, or else I will be forced to treat the other according to values I may find to be distasteful or actually abhorrent. This might be called the George Bernard Shaw objection; as Shaw quips, “do *not* do unto others as you would they should do unto you, [for] their tastes may not be the same.”<sup>12</sup> More fundamentally, why should what we would *want* done to us have any ethical significance at all, since many human wants so clearly conflict with common notions of what is right?<sup>13</sup>

---

<sup>10</sup> Joel Marks, “The Golden Rule Redux,” *Philosophy Now* 49 (Jan/Feb 2005).  
<http://www.philosophynow.org/index.htm>.

<sup>11</sup> See Karl Popper: “The golden rule is a good standard which can perhaps even be improved by doing unto others, wherever possible, as *they* want to be done by.” *The Open Society and its Enemies*, 2 vols. (Princeton: Princeton U P, 1966), 2: 386.

<sup>12</sup> George Bernard Shaw, *Man and Superman* (Cambridge: The University Press, 1903), Preface, “Maxims for Revolutionists.”

<sup>13</sup> I may want men to be given privilege over women; I may want all people trying to enter the country illegally to die at the border; etc. This problem is also at the heart of a series of logical absurdities that are supposed to follow from the golden rule, for example the “masochism objection” – the masochist would be obligated to treat others sadistically,

- *Kant*

Kant's moral philosophy is based on the conviction that ethical action must be grounded in a respect for duty for its own sake. This means, in part, that it is good *intentions* that count, and not the consequences of the action: to value an action because of its good consequences would be to make morality into a purely contingent and fickle affair, since the consequences of action are often out of our control. We may believe our action will benefit others – I join up to fight in a war I believe will liberate people from oppression – but it winds up harming them – the war only replaces one oppressive regime with a worse one. Moreover, the idea of “duty for its own sake” means that ethical deliberation must seek to remove any influence from our inclinations, desires, feelings, and emotions, since again we have little control over them: we are not free to feel this or that – we just *do* feel it. Yet ethics requires freedom. If we were not free (as we are not in our emotional lives, what Kant calls our “pathological” lives, from the Greek *pathos*, or feeling), it would not make any sense to say that we are responsible for our actions. But without responsibility, the very idea of ethical commands would be meaningless. Further, feelings are utterly subjective, and to ground moral action on such inclinations would doom us to ethical relativism. This, as we will see, is at the heart of Kant's objections to utilitarianism (or what for him is “happiness theory”), which both rejects intentions in favor of consequences and is based on feelings of pleasure and pain.

What is left, then, is for duty to be based entirely upon *reason*, which for Kant is universal, the same in us all. As a great enthusiast of the Enlightenment – the *Siècle de*

---

since that's what he desires himself; or the “thief objection” – as Alan Gewirth puts it, “If you always ‘do unto others as you would have them do unto you,’ a thief might say to the

*lumières*, the time of human history poised to witness what the Earl of Shaftesbury called a "mighty light which spreads itself over the world"<sup>14</sup> – Kant was convinced that by detaching ourselves from the sources of subjectivity (feelings, desires, emotions – the realm of darkness obscuring the sources of *Aufklärung*, the “lighting up” of reality by reason), reason is able to yield absolutely universal and necessary imperatives of duty. Indeed, the concept of universality is at the heart of all genuine ethics: it suggests the ideal of moving beyond the boundaries of our own self-interest and the particularity of our own subjective desires to embrace a common perspective on human well-being. The centerpiece of Kant’s moral theory, his categorical imperative, is precisely the codification of this concept of universality: “act only on that maxim [or principle of action] through which you can at the same time will that it should become a universal law” (FMM 421). Much of the notoriously excruciating technical work of Kant’s moral philosophy is devoted to demonstrating that all specific duties (do not lie, act benevolently towards others, etc.), each of which must pass the test of “universalizability,” can be derived from this single imperative.

It might seem as though Kant’s attitude towards the golden rule would be quite favorable. His categorical imperative, after all, is a way of generating duties that will apply to everyone, just as the rule involves the commandment to step outside the narrow confines of our personal point of reference to a life lived in reference to our community with our neighbor. All of the technical shenanigans of Kant’s philosophy might be seen

---

judge, ‘you wouldn’t want to go to prison. How can you send me to prison?’” Gewirth, "The Golden Rule Rationalized," *MidWest Studies in Philosophy* 3 (1978): 133-147.

<sup>14</sup> Shaftesbury to Jean Le Clerc, 6 March 1706, in *The Life, Unpublished Letters, and Philosophical Regimen of Anthony, Earl of Shaftesbury*, ed. Benjamin Rand (1900) (London: Routledge, 1992), 353.

as attending to the work of making the basic insight of the golden rule more rigorous and less subject to the criticism that it is simply too vague to be reliable.

There is indeed a sense in which Kant's ethical theory can be seen to "use the golden rule as inspiration,"<sup>15</sup> but it is a sense in which the rule undergoes radical transformation. The transformation is so radical, in fact, that it has lead one commentator to claim that "Kant succeeded ... almost in invalidating the golden rule and disqualifying it from future discussion in ethics."<sup>16</sup> Let us begin with a passage in which Kant refers to the golden rule as "banal": "Let it not be thought that the banal *quod tibi non vis fieri*, etc. [... *alteri ne faceris*: what you would not wish done to yourself, do not to another]<sup>17</sup> could here serve as guide or principle, for it is ... restricted by various limitations. It cannot be a universal law, because it contains the ground neither of duties to one's self nor of the benevolent duties to others (for many a man would gladly consent that others should not benefit him, provided only that he might be excused from showing benevolence to them). Nor does it contain the ground of obligatory duties to another, for the criminal would argue on this ground against the judge who sentences him" (FMM 430, n14). Thus far, Kant is only raising common objections to the rule – the vagueness objection (the golden rule is not a reliable guide) and the absurdity objection (the judge

---

<sup>15</sup> Bryant Abrahamson, "The Golden Rule," <http://www.inquiringminds.org/newsletter/0504/golden-rule>.

<sup>16</sup> Hans Reiner, *Duty and Inclination: The Fundamentals of Morality Discussed and Redefined with Special Regard to Kant and Schiller*, tr. Mark Santos (The Hague: Martinus Nijhoff, 1983), 274.

<sup>17</sup> This is the negative or prohibitory form of the golden rule, which has a distinguished history. The Roman emperor Alexander Severus had these words written in letters of gold on the walls of his palace long before the phrase "golden rule" was coined (see Philip Schaff, *History of the Christian Church* [Peabody, MA: Hendrickson, 1996], 2: ch. 2, sect. 21. This negative formulation can also be found in Confucius (*Analecets* 15:24), in Hillel (*Shabbat* 31a), in St. Benedict (*Regula Sancti Benedicti* 61:14), and many others.

should not sentence the criminal; see n13 above). But these objections are compatible with the view that Kant is nevertheless “inspired” by the golden rule, and merely seeks to recast it in a non-“banal” way that will overcome vagueness and logical absurdities.

Kant’s objections to the rule are in fact much more profound and far-reaching, and center on his complete rejection of emotion and desire as components of moral deliberation. Most simply, we can see that Kant must oppose any formulation of moral duty which asks us to treat others as we would *want* to be treated – or for that matter as *they* would want to be treated. The categorical imperative is merciless in forbidding any reference to wants, either my own or those of any other. This means, for example, that Kantian ethics rejects the idea that our duties to others in any way rest upon consent: whether the other agrees or not, or gives approval or not, is entirely beside the point. If reason tells us that an act is our duty, we must do it regardless of what I want or the other wants. She may not want to hear the truth about her medical condition – she has untreatable cancer – but to conceal the truth is always wrong: I could not, Kant believes, consistently will that concealing the truth in situations where the truth might be disagreeable to another be made into a universal law, since then we could never trust anyone to tell us the truth (FMM 422).

That Kant’s categorical imperative abstracts away from the desires, the happiness, and even the consent of others means that the nature of the other as conceived in the golden rule is fundamentally altered. In the Leviticus formulation of the rule, we are told that “When an alien lives with you in your land, do not mistreat him. The alien living with you must be treated as one of your native-born. Love him as yourself, for you were aliens in Egypt” (Lev 19: 33-34). But for Kant, in moral life there are no aliens, no native

country, no Egypt – there is no real “otherness,” since all differences of time and place, of culture and custom, are effaced in our moral deliberations. The golden rule’s attunement to the differences between self and other is for Kant a sure sign of its subjectivism and therefore its inadequacy as a foundation of moral philosophy. A genuinely universal moral system requires that there be no strangers, but only rational agents, and no foreign lands but only the land of reason, a “realm of ends” (FMM 433) where we are all the same. Crucially missing from such a rational kingdom is the need for the *empathy* that seems so much at the heart of the golden rule, for empathy implies otherness, and otherness dooms us to moral relativism and a “pathological” foundation of ethics.

The radical character of Kant’s rewriting of the golden rule can perhaps best be seen in his comments about the duty to love others. In his *Metaphysical Elements of Ethics*, Kant writes that “Love is a matter of feeling, not of will or volition, and I cannot love because I will to do so, still less because I ought; ... hence there is no such thing as a duty to love” (MEE, 19). Thus Kant exactly shares Freud’s criticism of the golden rule that the duty to love is utterly “unpsychological,” since, as Kant says (uncannily anticipating Freud), “we must make the sad remark that our species, alas! is not such as to be found particularly worthy of love when we know it more closely” (MEE, 19). As for Freud, so too for Kant it is utterly unrealistic to say that we nevertheless have a duty to feel affection for those who are not worthy of it – our affections cannot be commanded. Kant’s move is to de-psychologize love, and recast it as “disinterested benevolence.” Only in this recasting can we meaningfully be said to have a duty to “love”: “For love, as an affection, cannot be commanded, but beneficence for duty’s sake may; even though

we are not impelled to it by any inclination – nay, are even repelled by a natural and unconquerable aversion. This is practical love and not pathological, a love seated in the will ... and not in tender sympathy.” It is only in this sense that we can “understand those passages in Scripture in which we are commanded to love our neighbor, even our enemy” (FMM 399) – an understanding which replaces affection with reason, love with disinterested beneficence, and empathy (“tender sympathy”) with a practical procedure for determining duties.

We are now in a position to see how Kant would reply to the challenges Freud poses for the golden rule.

*1. The golden rule is psychologically unrealistic – it requires a severe repression of the instinctive basis of love, which is sexual, discriminatory, and egoistic.* As we have just seen, Kant entirely agrees that the rule is psychologically unrealistic. We simply cannot command love, which, Kant agrees with Freud, is intrinsically egoistical. Kant solves this problem by depsychologizing love, making it into a rational principle of disinterested beneficence that *can* be commanded.

*2. The golden rule devalues love by failing to discriminate between those who I actually have affection for (or desire) and those I do not – it makes me “throw away” my love by giving it to everyone, with my eyes closed.* I suspect that Kant would agree with Freud on this point as well. In a fascinating passage from his *Metaphysical Principles of Virtue*, Kant seems to be anticipating Freud’s idea that the golden rule both requires an “enormous inflation” of love (CD 90) – since it must be given to everyone – and yet entails a basic devaluation or cheapening of love, precisely because it must be spread so thin. Kant writes that “benevolence in the form of a general love of mankind is

extensively, to be sure, the greatest possible benevolence” (cf. Freud’s “enormous inflation”), “but intensively (in degree), the smallest possible. If I say that I take an interest in the welfare of this man merely in accordance with the general love of mankind, the interest I take in him is the smallest that can be” (MPV 451). For Freud, this reduction of love to the “smallest possible” intensity by the golden rule’s requirement that we love indiscriminately (just close your eyes and love) is damning evidence of the depreciation of love. Kant might well agree, insofar as love is understood here as an affection. But once love is recast as disinterested benevolence, the “smallness” of love becomes for Kant its great virtue: such love *must* be small – indeed ideally dwindle away into nothingness – *as an emotion*, since the duty of benevolence is derived from reason alone, setting aside all affective attachments (or aversions) to the object of “love.”

3. *The golden rule is unjust, since for the most part people are “wolves” who do not deserve my love.* This is also true for Kant, insofar as he sees our species as largely unworthy of our affection. Again, it is only when the basis of duty is moved from affection to reason that we can (and must) extend beneficence even to those for whom we are “repelled by a natural and unconquerable aversion.”

4. *The golden rule is utterly impractical, since it puts us at a disadvantage with respect to those (the great majority) who do not act likewise.* This criticism will of course be as relevant to Kant’s own philosophy as to the golden rule, since whether it is based on reason or love, on duty for its own sake or a consideration of what others would want us to do, on the categorical imperative or empathy, acting so as to insure the welfare of others may indeed be disadvantageous with respect to self-interest. Kant never seeks to evade this fact, and even seems to pride himself on the purity of his moral system by

acknowledging that it may require a burden of unhappiness.<sup>18</sup> “But a rational being,” Kant writes, though he scrupulously follow [the demands of the categorical imperative], cannot for that reason expect every other [person] to be true to it; ... that is, he cannot count on its favoring his [hope] of happiness. But still [the imperative] ... remains in full force, because it commands categorically, ... without any end or advantage to be gained by it ... .” This alone is what gives “dignity to humanity” (FMM 438-9). Freud is right, then: neither the prophets nor Christ nor Kant deign to argue that their commandments are “practical” in the sense of being advantageous to us. These commands may well require sacrifice, hardship, and unhappiness. To ask, as Freud does, “what good will it do us” (CD 56) to love my neighbor or obligate myself to her welfare, is simply to misunderstand the nature of morality from the ground up.

As for the remaining two common challenges to the golden rule, that it is too vague and that it produces unavoidable moral dilemmas since there will be inevitable conflicts between what I would want done and what the other would want me to do, Kant’s transformation of the golden rule into a doctrine of disinterested benevolence allows him to give quite straightforward answers. We have seen that Kant calls the rule “banal” precisely because it cannot “serve as guide or principle; ... it contains the ground neither of duties to one’s self nor of the benevolent duties to others.” More interesting, though, is to consider Kant’s response to a possible defense of the rule against the charge of vagueness, namely that in places like the Book of John, we are in fact given direction about how to enact the rule. In John 13: 34 and 15: 12, for example, Christ tells us to

---

<sup>18</sup> See Kant’s discussion in his FMM of how we must disregard any consideration of the “usefulness” or “fruitlessness” of our actions for our own interests, and how duty must be independent of any hope for happiness (e.g., FMM 394-96).

“Love one another; *as I have loved you*, so you are to love one another.” Thus the rule does have a content: we know how to love by studying the ways in which Christ loves others. For Kant, though, this does not solve the problem: Christ’s love will be a bad model, because it is “pathological,” based on empathy and affection, which are utterly subjective. More generally, for Kant we must be on guard against theologizing duty (basing moral obligations on scripture, for example); such a foundation is contingent upon dogmas developed at certain times and places in human history. On the contrary, theology (if we must have it at all) should be grounded on reason. Religion should be confined “within the limits of reason alone,” to cite the title of a book that got Kant into hot water with the censor, not to speak of a reprimand (and threat) from the King himself.<sup>19</sup>

Finally, as for the criticism that the golden rule yields moral dilemmas (the conflict between my own desires and values and those of the other), we have seen that Kant’s philosophy seeks to short-circuit the very possibility of such dilemmas by

---

<sup>19</sup> *Religion Within the Limits of Reason Alone (Die Religion innerhalb der Grenzen der blossen Vernunft)* (1793). In 1788, two years after Frederick the Great’s death, the new king, Friedrich Wilhelm II, appointed Johann Christoph Wöllner to the post of minister of cultural and educational affairs. Wöllner quickly issued a censorship edict prohibiting irreligious beliefs from being “propagated so as to cause others to err or falter in their faith.” It was not long before Kant came to the attention of the censor with his essay “On the Struggle of the Good Principle with the Evil for Mastery over Mankind” (1791), and its publication was forbidden. Kant nevertheless somehow managed to convince the Dean at the University of Jena to publish the article as one of four essays in a new book, *Religion within the Limits of Reason Alone*. Shortly thereafter, Kant received a letter personally signed by the King: “Our most High Person has for a long time observed with great displeasure how you misuse your philosophy to undermine and debase many of the most important and fundamental doctrines of the Holy Scriptures and Christianity; how, namely, you have done this in your book, *Religion with the Limits of Reason Alone*. ... If you continue to resist you may certainly expect unpleasant consequences to yourself.” See Ernst Cassirer, *Kant’s Life and Thought*, tr. James Haden (New Haven: Yale U P, 1981), 376-380.

removing their source, the subjectivity of human desire. When duties are no longer determined by what we would *want* others to do, but by reason alone, which is universal, the golden rule is transformed from a “banal” guidebook to human inclinations into a science of ethics.

- *Utilitarianism*

Utilitarianism as an ethical theory is the doctrine that the moral worth of an action is to be judged by the extent to which it maximizes the pleasure or happiness – and minimizes the pain or unhappiness – of the greatest number of people (Bentham, PML 17; Mill, U 18).<sup>20</sup> Utilitarianism is straightforwardly based on a psychology of human desire, thus setting it upon an entirely different foundation from Kant’s ethical theory. It is the *fact* of the human desire for pleasure or happiness – a fact based upon “the natural constitution of the human frame,” as Bentham puts it (PML 19) – that establishes the “proof” of utilitarianism as a moral philosophy. “Human nature is constituted,” Mill writes, “as to desire nothing [that is not] ... a means of happiness” and “conducive to pleasure” (U 39), and “the sole evidence it is possible to produce” that we *ought* to seek to maximize pleasure is simply “that people do actually desire it” (U 37; and see Bentham, PML 19). I will leave aside the question of whether this “proof” commits what philosophers call the naturalistic fallacy (or the is/ought fallacy), namely that one cannot validly conclude anything about what one ought to do from what one does do, but only note in passing that this is one of the grounds for Kant’s rejection of “happiness theories.”

We still need to examine how the utilitarian psychology of desire accords with the golden rule’s ideal of loving one’s neighbor, but at least we know that utilitarianism, by

restoring a psychological account of human motives, will be able to address the experience of love as an emotion, unlike Kant, for whom love becomes a disinterested duty of reason. Furthermore, utilitarianism restores the other as a concretely embodied and situated self whose desires and interests and values must be accounted for. If for Kant the otherness of the other vanishes, since in his “kingdom of ends” everyone is treated alike (there being only one human reason), for the utilitarian it is the actual desires of others – whether we find them to be “rational” or not – that we must take seriously in seeking to bring about the greatest happiness for the greatest number. Kant is right that there is no single definition of happiness, which for him is decisive grounds for rejecting it as a foundation for ethics. We each of us have our own subjectively determined happiness – what gives pleasure to me may be distasteful to you. But it is just this fact of human difference that utilitarians seek to affirm in developing a moral theory that does not allow the self to impose its own conception of value on another.

Utilitarianism thus incorporates the golden rule’s commandment that we do unto others as we would have them do to ourselves *under the interpretation that* what I would have them do to me is to recognize *my* wants and desires; hence I too must do unto the other from the perspective of what *she* wants. While this affirmation of otherness allows utilitarianism to escape what for them is the ghostly, apparitional, fantastic account of the “other” in Kantianism – where there are no “aliens,” no native-born and no strangers from Egypt, but only homogeneous rational agents – clearly it must also face the problem of moral relativism which Kant’s philosophy is designed to solve. The tidiness of Kant’s replies to objections about the golden rule will not be possible for the utilitarian. For

---

<sup>20</sup> Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (1789),

example, Freud's basic criticism that the golden rule is based on an "unpsychological" account of love – which Kant answers so easily by removing all "merely" psychological questions from discussion as misunderstandings of ethics – will have to be engaged by utilitarianism on the grounds of psychology itself: it will have to provide an alternative psychology of desire to the one that leads Freud to see the golden rule as an absurdity and utterly impractical.

Mill explicitly states that utilitarian ethical theory directly shares the insight of the golden rule: "In the golden rule of Jesus of Nazareth, we read the complete spirit of the ethics of utility. To do as you would be done by, and to love your neighbor as yourself, constitute the ideal perfection of utilitarian morality" (U 24). Right at the outset, though, we should come to terms with a criticism that is often posed against utilitarianism, namely that it bases morality upon self-interest. If this were so, then although utilitarianism could still yield obligations to others and perhaps fulfill the letter of the biblical Law, it would seem to replace the empathetic perspective of the golden rule with an egoistic perspective that conflicts with the rule's spirit.

This criticism is most clearly applicable to Bentham, who writes in his *Book of Fallacies* that "in every human breast ... self-regarding interest is predominant over social interest; each person's own individual interest over the interests of all other persons taken together."<sup>21</sup> Thus Bentham would seem to agree with Freud's insistence that love is naturally egoistical. His account of love is even structurally consistent with Freud's theory of the development of the ego from its "primary narcissism" (the original state of the psyche where all libido, or pleasure, is turned inward to the self) to "anaclitic"

---

in *The Utilitarians* (Garden City, NY: Anchor, 1973). Henceforth PML.

or “object”-love, where libido is directed onto an other as the object of pleasure. In Freud’s account, the move from self-love to love of the other is intrinsically bound up with the tragic recognition that the self cannot satisfy itself but needs another to do so: the other, then, is not loved as an end in itself, to use Kant’s terminology, but as a means to one’s own pleasure. Object-love is but displaced narcissism.<sup>22</sup> While Bentham, of course, does not share the apparatus of Freud’s account of the instincts, he does place the experience of pleasure (and the avoidance of pain) as the “two sovereign masters” of us all: “It is for them alone to point out what we ought to do” (PML 17). And the only place in his *Principles of Morals and Legislation* where he speaks of love for others (as opposed to the many references to love of power, glory, ease, liberty, and etc.), which he refers to variously as “brotherly love,” “philanthropy,” “charity” – and once as “compassion” – he is clear that its motive is an experience of pleasure (the “pleasure of sympathy”) for the self (PML 109f).

While Mill is himself sometimes read as justifying the concern for others’ happiness egoistically, according to the happiness it affords the self, the overwhelming spirit of his moral philosophy is one of insisting on the concern for others as irreducible to self-interest. Thus he harshly criticizes Bentham’s psychological egoism,<sup>23</sup> and speaks often of the “nobility” of self-sacrifice and renunciation, of “resigning entirely one’s own portion of happiness” in order to increase the happiness of one’s “fellow creatures” (U 23-24). He is aggravated to distraction by the criticism of utilitarianism that it is an

---

<sup>21</sup> Bentham, *The Book of Fallacies* (London: Hunt, 1824), 392-3.

<sup>22</sup> Freud, “On Narcissism” (1914), in *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, 24 vols., ed. James Strachey (London: Hogarth, 1953ff): vol. 14: 73-102.

egoistic philosophy: “I must again repeat what the assailants of utilitarianism seldom have the justice to acknowledge, that 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” (U 24, emphasis added).

Since regard for the other can even entail the complete renunciation of one’s own prospects for happiness, it cannot (always) be motivated by a calculation about what is in one’s own best interest, but must be something like the empathetic love commanded by scripture. And indeed it is immediately following the above passage where Mill insists that it is “not the agent’s own happiness” which is at the heart of utilitarian ethics that he refers to the golden rule of Jesus of Nazareth as encapsulating “the complete spirit of the ethics of utility.” The motive for promoting the welfare of others may certainly be based on a calculation of “the greatest happiness” for Mill, but ideally it should be “a direct impulse” and “sentiment” which affirms what is most “beautiful” and “exalted [in] human nature” (U 24).

For all the undeniable closeness of Mill’s portrayal of utilitarianism to the golden rule, three questions should be asked before we turn to summarize how utilitarianism would respond to criticisms of the rule. These questions are not intended to disprove Mill’s contention that the moral philosophy of utilitarianism is in essence the same as Jesus’ teaching of love, but to open ways of seeing how Bentham’s and Mill’s recasting of the rule involves giving it a certain tonality or shading which may or may not satisfy students of the scriptures.

---

<sup>23</sup> See especially his “Remarks on Bentham’s Philosophy” (1833), in *Collected Works*

First, then, the fact that utilitarianism determines moral action based upon an aggregation of the interests of everyone who will be affected means that we will regularly face situations where the happiness of some must be denied in order to assure the happiness of the majority. The now famous “runaway trolley car” thought experiment is a good, if extreme, example. A runaway trolley car is hurtling down a track. In its path are five people who will be killed unless you, a bystander, flip a switch that will divert it onto another track, where it will kill one person who is standing there.<sup>24</sup> Simply summing up the various interests of all involved would appear to justify flipping the switch, killing the one person to save the greater number.

While no formulation of the golden rule in the Bible gives any direct aid in interpreting how one’s duty to love others would help us decide in such cases, the spirit of the message seems to count the lives of all to be infinitely precious. If this is true, then the golden rule would seem to be much closer to the Kantian ethic (in this respect, at any rate), in which we can never sanction choosing the life of one over any number of others. Every person is an “end in itself,” and his or her worth is not calculable in quantitative terms. One simply cannot “aggregate” the value of lives.

The utilitarian might reply that such choices are tragically necessary in the real world. However improbable the trolley car example, we all know that we are sometimes faced with often excruciatingly difficult moral dilemmas. In the extreme case of the runaway trolley car, a choice to kill one to spare five does not imply any lack of empathy for the one who is killed or any indifference to the tragedy of the choice. If the utilitarian

---

(Toronto: University of Toronto Press, 1985), vol. 10.

<sup>24</sup> See, for example, Judith Jarvis Thomson, “The Trolley Problem,” *Yale Law Journal* 94 (1985): 1395-1415.

were in a particularly devilish mood, she might even point out how conveniently the authors of the various scriptures in which the golden rule is commanded never address the possibility of moral dilemmas (any more than Kant does, for that matter). Kant and the scriptures present an entirely idealized world, our utilitarian might continue, a “kingdom of ends” immune from tragedy and horror, the eschatological vision of a “holy city, new Jerusalem,” where “God shall wipe away all tears ... and there shall be no more death, neither sorrow nor crying, neither shall there be any more pain” (Rev 21:1-2, 4). But this is not our world, and utilitarianism at least faces up to providing a means to decide in tragic cases. It will be imperfect, since moral dilemmas can rarely be solved simply (see U 30), but it will be guided by the effort to allow for the most happiness for as many as possible, which is all that realistically can be asked of the golden rule. I note here that while this answer seems perfectly respectable to me, I am not at all sure that the authors of the scriptural passages in question would be (or should be) convinced.

A second question has to do with the consequentialist character of utilitarianism. In typically anti-Kantian style, the utilitarians believed that the intentions behind our actions contribute precisely nothing to the moral worth of the action, which is determined entirely by whether the consequences add to or subtract from the happiness of the greatest number (see Bentham, PML 87f; Mill, U 25). Certainly Bentham and Mill acknowledge that intentions have a lot to do with the virtue (or lack thereof) of the actor (e.g., U 26-27), but it is the consequences alone that are important for assessing the moral value of the action itself. Bentham goes so far as to ask us to consider a man whose “motive [is] ill-will; call it even malice, envy, cruelty; it is still a kind of pleasure that is his motive: the pleasure he takes at the thought of the pain which he sees, or expects to see, his

adversary undergo. Now even this wretched pleasure, taken by itself, is good, ... as good as any other that is not more intense.” If the consequences of this intended action somehow avoid harm – perhaps the malicious man is ignorant that his victim craves nothing more than to have pain inflicted upon him – the act itself is not wrong (PML 323 n92).

Surely this view is hard to reconcile with the spirit of the golden rule. Again, none of the passages in scripture which enjoin us to treat others as we would be treated by them, or to love our neighbor, explicitly comment about either intentions or consequences, but it seems far-fetched in the extreme to imagine Christ telling his disciples to just weigh up the consequences of their actions to determine if they have acted rightly. It is the inner state of the one who loves or does not love that seems most important – his soul, his spirit, his commitment and resolve to live a life of love.

The third and final question I wish to raise about the relation of utilitarianism to the golden rule has to do with its most fundamental principle, that all people desire happiness, that “nature has placed mankind under the governance of two sovereign masters, pain and pleasure,” and that “the standard of right and wrong ... [is] fastened to their throne” (Bentham, PML 17). This is a notoriously self-proving claim, since anyone who protests that she (a Kantian, let us say) may well desire to be happy but nevertheless seeks to act *not* with the purpose of maximizing the odds of achieving happiness (either her own or that of others), but simply in order to fulfill her duties, which often require the sacrifice of happiness, can be dismissed quickly by the retort that clearly acting according to her sense of duty is what makes her happy.

Bentham and Mill both use this maddening retort, and act as though it were unproblematic (e.g., Bentham 20; Mill 15). But is it? Suppose one were to argue that when we look at the life of Christ, and the way he loved those who were closest to him, his disciples, he was not concerned about their happiness at all; in fact he loved them in such a way that they were destined to live lives of great suffering and pain. The utilitarian could of course reply that, aha! but the pain of this life was understood as an unfortunate prerequisite for a future life of salvation and eternal bliss. Yet it seems to stretch credulity to read the demands Christ's love placed on others to be based on a reward system. To love Christ is to accept the possibility of suffering, to be prepared if necessary to give up one's concern for contentment and pleasure. Is it really a great merit or sign of profound insight into human psychology on the part of the utilitarians that they can reply that even here, the renunciation of contentment is, for whatever reason, evidently what makes such individuals happy?

These questions are hardly definitive in creating a gap of unbridgeable difference between utilitarianism and the golden rule. The very vagueness of the rule and the openness to interpretation of the theology which frames it means that there will be plenty of room for disagreement with my particular reading of "the spirit of the golden rule" – and thus with my reservations about how exactly it can be mapped onto the contours of utilitarian philosophy. And despite my own reservations, I believe that Mill was noticing a real affinity when he spoke of the golden rule as the "ideal perfection of utilitarian morality." There is nothing "banal" about the rule for utilitarianism, as there is for Kant, for whom the rule is mired in romantic notions of "pathological" love. On the contrary, it

speaks to what the utilitarians see as the heart of moral life, the “desire to be in unity with our fellow creatures” (U 34).

We may now sum up the responses of utilitarianism to the various challenges posed for the golden rule by Freud and others.

1. *The golden rule is psychologically unrealistic – it requires a severe repression of the instinctive basis of love, which is sexual, discriminatory, and egoistic.* The utilitarians have a fundamentally different reading of egoism than Freud does. They are certainly as aware of the narrow, aggressive, and narcissistic potentialities of egoism as is Freud, but they see the egoism of human nature to be complimented by a capacity of reason to understand that self-interest is inextricably bound up with the interests of others. Acting solely for one’s own satisfaction is seen as a failure of reason, and ultimately a failure of moral and social education. Mill writes that “education, ... which [has] so vast a power over human character, should use that power as to establish in the mind of every individual an indissoluble association between his own happiness and the good of the whole” (U 24).

Freud’s most frequent comments on education, on the contrary, define its mission to be what he calls an “education to reality,” disabusing the young of their illusions that they can count on the good will of others and that there is a God who cares about justice in the world.<sup>25</sup> Bentham and Mill, for all the disappointments they experienced in their respective careers advocating for liberal social reforms, remained optimistic about human nature and human society. “No one whose opinion deserves a moment’s consideration,” Mill writes (and one can just imagine Freud cringing from his grave), “can doubt that

most of the great positive evils of the world are in themselves removable and will ... be in the end reduced within narrow limits” (U 23). The golden rule, doing unto others as we would have them do to us, expresses precisely the value of the interdependence of self and other which is essential for human happiness, and which is deeply rooted in our nature as both desiring and social beings.

2. and 3. *The golden rule devalues love and is unjust; most people do not deserve my love.* Again, there is a basic difference of sensibility between Freud and the utilitarians over the nature of human nature. Most others *do* deserve my love, for the utilitarian, since their happiness is essential for my own. Certainly “the stranger” whom the book of Leviticus bids us to welcome, and whom Freud sees as offering me no prospect of satisfaction, is a key figure for Mill. Mill is eloquent in his insistence that it is precisely otherness and difference which is necessary for the health of society: only when we open ourselves to different opinions and different ways of being do we fully understand and gain perspective on our own.<sup>26</sup> To love one’s neighbor, then, is not a devaluation of love but its augmentation: the “alien” enriches us.

3. *The golden rule is utterly impractical, since it puts us at a disadvantage with respect to those (the great majority) who do not act likewise.* There is no doubt about it, this will sometimes be true. But again, civil society, and therefore my own happiness which is dependent upon it, require that we move beyond the almost paranoid distrust of others that Freud’s preoccupation with the other as “wolf” traps us in.

---

<sup>25</sup> *The Future of an Illusion* (1927), tr. W. D. Robson-Scott (No place of publication given: Institute of Psycho-Analysis, 1928), 85-86. And see CD 59.

<sup>26</sup> See Mill, *On Liberty* (1859), in *Utilitarianism, Liberty, and Representative Government* (New York: Dutton, 1951), especially chapter 2, “Of the Liberty of Thought and Discussion,” 102-52.

In terms of the further challenges to the golden rule we have alluded to, that it is hopelessly vague and that it condemns us to inevitable moral dilemmas generated by the disparity between what I would want done and what others would want me to do, the quintessentially pragmatic temperament of utilitarianism counsels us to accept these problems as an unavoidable consequence of living in an imperfect world, while giving us a means to do the best that can be expected. Yes, the golden rule is vague, and while utilitarianism offers certain tools for greater specificity and practicality – showing us how “doing unto others” involves an openness to the other’s conception of what makes her happy, and providing a method for weighing and balancing competing interests – in the end, moral life is often messy. There is no *deus ex machina* solution; the Kantian yearning for an absolutely reliable algorithm for producing immaculately “universal and necessary” imperatives of action is a fantasy that, despite its allure, is simply comical when applied to the real world.

One might even say that the notorious vagueness of the golden rule is part of its great virtue, since it recognizes that ethical deliberation is always situational and therefore ambiguous, uncertain, and to some extent obscure. The particular “circumstances with which [actions are framed] ... must necessarily be taken into account before anything can be determined relative to” the rightness or wrongness of an act, Bentham writes. What is right to do can never be answered in abstraction from the situation: “it can never be known whether [an act will be] beneficial, or indifferent, or mischievous apart from a knowledge of the circumstances” (PML 78f).

This unavoidably situational character of ethical choice necessitates the unavailability of neat and tidy answers. As Mill puts it, speaking of moral dilemmas,

“there exists no moral system under which there do not arise unequivocal cases of conflicting obligation. These are real difficulties, the knotty points both in the theory of ethics and in the conscientious guidance of personal conduct. They are overcome practically, with greater or with less success, according to the intellect and virtue of the individual” (U 30). The golden rule, and the principle of utility which seeks to reshape it into as usable a rule of conduct as possible, by advocating a way of life that affirms the solidarity of our own well-being with that of others, is simply the best we can do.

- *Conclusion*

In closing, I would like to point out a peculiarity of moral theory, that one often finds stridently conflicting theories advocating the very same actions. In Matthew 25, Christ gives an explanation of what it means to love one’s neighbor. Such love is diametrically opposed to the Freudian criterion of love: “What good is it for me?” It demands that we reimagine love beyond the comfortable limits of personal attraction to the other, and find precisely in the abjectness of the other – the other in his helplessness and need, who may have nothing to offer me by way of recompense – the occasion for experiencing my own humanity, what we have seen Mill to call the “most beautiful and exalted possibility of human nature.”

In Matthew, Christ tells us that

I was hungered, and ye gave me meat; I was thirsty, and ye gave me drink; I was a stranger, and ye took me in;

Naked, and ye clothed me; I was sick, and ye visited me; I was in prison, and ye came unto me.

... Verily, I say unto you, inasmuch as ye have done it unto one of the least of these my brethren, ye have done it unto me. (Matt 25:35-36, 40)

For the utilitarian, such action is what Bentham calls a “well-advised benevolence” (PML 117), and what Mill, less prosaically, sees as an exalted impulse to contribute to the happiness of others. For Kant, there is no question of basing such actions on “tender sympathies,” which are utterly unreliable guides to moral action, yet there is certainly a moral duty of “disinterested beneficence” grounded in the recognition of every other, including the hungry, the naked, the sick, and the incarcerated, as an end in itself. For the Christ of Matthew, and for the prophets whose articulation of the Law Christ invoked, the obligation of love is based simply on a radical experience of empathy for those – all of us – in need of love.

The differences of nuance between Kant, utilitarianism, and Christ and the prophets are important, since they provide alternative conceptions of the motives for love – and hence different perspectives on the human condition. But all are committed to the basic insight of the golden rule, that our humanity is inextricably bound up with moral obligations to the other, whether in the face of the other we see the mirror of our own dignity as rational agents (Kant), the ineluctable difference of human desire (utilitarianism), the voice of the stranger (Torah Leviticus), or the cry of need of the abject (Christ).

### *Works Cited*

- Bentham, Jeremy. *The Book of Fallacies*. London: Hunt, 1824.
- , *An Introduction to the Principles of Morals and Legislation* (1789). In *The Utilitarians*. Garden City, NY: Anchor, 1973.
- Cassirer, Ernst. *Kant's Life and Thought*. Tr. James Haden. New Haven: Yale U P, 1981.
- Freud, Sigmund. *Civilization and its Discontents* (1930). Tr. James Strachey. New York: Norton, 1961.
- , *The Future of an Illusion* (1927). Tr. W. D. Robson-Scott. No place of publication given: Institute of Psycho-Analysis, 1928.
- , "On Narcissism" (1914). In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, 24 vols. Ed. James Strachey. London: Hogarth, 1953ff. 14: 73-102.
- Gewirth, Alan. "The Golden Rule Rationalized." *MidWest Studies in Philosophy* 3 (1978): 133- 147.
- The Holy Bible*. King James Version.
- Kant, Immanuel. *Foundations of the Metaphysics of Morals* (1785). Tr. Lewis White Beck. Indianapolis: Bobbs-Merrill, 1969.
- , *The Metaphysical Elements of Ethics* (1780). Tr. Thomas Kingsmill Abbott. London: Kongmans, Green, 1889. References are to the online text provided by Marxists.org (2003) at <http://www.marxists.org/reference/subject/ethics/kant/morals/ch01.htm>.
- , *The Metaphysical Principles of Virtue (Part II of The Metaphysics of Morals)* (1797). Tr. James Willington. Indianapolis: Bobbs-Merrill, 1964.
- Kierkegaard, Søren. *Works of Love* (1848). Tr. David and Lillian Swenson. Princeton: Princeton U P, 1946.
- Marks, Joel. "The Golden Rule Redux." *Philosophy Now* 49 (Jan/Feb 2005). <http://www.philosophynow.org/index.htm>.
- Mill, John Stuart. *On Liberty* (1859). In *Utilitarianism, Liberty, and Representative Government*. New York: Dutton, 1951.
- , "Remarks on Bentham's Philosophy" (1833). In *Collected Works*. Toronto: University of Toronto Press, 1985: vol. 10.

- , *Utilitarianism* (1861). Ed. Samuel Gorovitz. Indianapolis: Bobbs-Merrill, 1971.
- Popper, Karl. *The Open Society and its Enemies*, 2 vols. Princeton: Princeton U P, 1966.
- Rand, Benjamin, ed. *The Life, Unpublished Letters, and Philosophical Regimen of Anthony, Earl of Shaftesbury* (1900). London: Routledge, 1992.
- Reiner, Hans. *Duty and Inclination: The Fundamentals of Morality Discussed and Redefined with Special Regard to Kant and Schiller*. Tr. Mark Santos. The Hague: Martinus Nijhoff, 1983.
- Schaff, Philip. *History of the Christian Church*, 8 vols. Peabody, MA: Hendrickson, 1996.
- Shaw, George Bernard. *Man and Superman*. Cambridge: The University Press, 1903.
- Thomson, Judith Jarvis. "The Trolley Problem." *Yale Law Journal* 94 (1985): 1395-1415.

# 5.

## **PHILOSOPHICAL REFLECTIONS ON THE GOLDEN RULE**

**JEFFREY WATTLES**

**KENT STATE UNIVERSITY**

Is the Golden Rule truly golden? May we take it our rule of living? I believe so, but only if it is broadly and flexibly conceived, embracing interdisciplinary and global perspectives. It needs to nourish itself on developments in diverse cultures of the ancient world, as well as on modern social science, philosophy, and religion. The Golden Rule needs to symbolize living in truth and beauty, as well as goodness. This paper offers a simplified way to organize thinking about the Rule, some levels of meaning in terms of which to listen to, question, and integrate the multiplicity of cultures and disciplines.

The Golden Rule—Do to others as you want others to do to you—is intuitively accessible, easy to understand in its simplicity, communicating confidence that the agent can find the right way. It functions as a summary of one’s moral tradition. The Rule most commonly expresses a commitment to treating others with consideration and fairness, predicated on the recognition that others are like oneself.

The Golden Rule serves the needs of educated and uneducated people alike, and stimulates philosophers to codify its meanings in new formulations. Given the equal, basic worth of each individual, the rule implies a requirement of consistency; as Samuel Clarke put it, “Whatever I judge reasonable or unreasonable for another to do for me; that, by the same judgment, I declare reasonable or unreasonable, that I in the like case should do for him.” In addition, the Rule carries implications for social, economic, and political matters. In one form or another, with interpretations that differ and overlap, the Rule is a precious word in the shared language of humankind.<sup>i</sup>

The Rule’s simplicity enfolds complexity, diverse sequences of levels of meaning, which have been uncovered in cultures and academic disciplines wherever the Rule has been studied. The specific cultural and disciplinary challenges which have been met with the help of the Golden Rule have each added meanings to the Rule that become, through education, the common heritage of humankind. In ancient China, the challenge was to move from an ethics of social-ethical conformity to a life of moral spontaneity. In ancient Greece, the challenge was to separate a sophistic principle of social acceptability from norms of philosophic reason. In ancient Judaism, the challenge was to summarize a complex tradition in a way that highlighted essentials. In the New Testament presentations of the teachings of Jesus of Nazareth, the challenge was to integrate, first,

realism about the consequences of right and wrong conduct, second, an emerging summary of morality, and, third, an ideal of spiritually transformed living. In medieval thought the challenge was to distinguish the level of conduct that could be expected of persons without faith from the level that could be expected of those with faith. In the early modern philosophy of Kant and others, the challenge was to separate interpretations of the Rule that were vulnerable to counterexamples from reformulations that would capture the Rule's implicit rationality. Those efforts of reformulation have been taken to new heights in recent analytic philosophy, while European philosophers including Levinas and Ricoeur have deepened our comprehension of empathy and our awe in the presence of the Other. In twentieth-century psychology, Lawrence Kohlberg linked developmental levels of moral reasoning with levels of meaning in the Golden Rule; and Erik Erikson gave interpretations of the Rule to fit the sequence of challenges that people face during the course of life. These summary statements indicate the Rule's ability to illumine paths of progress through diverse cultural and personal problems.

We thus glimpse the *life* in the Golden Rule. It cannot be captured in a static interpretation, since it engages the thoughtful doer in a process of growth. To follow it to the end is to move from egoism to sympathy, to sharpen moral intuition by reason, and to find fulfillment beyond duty-conscious rule-following in spontaneous, loving service. The unity of the Rule, amid its wide diversity, is its function as a symbol of this process of growth. Whoever practices the Golden Rule opens him- or herself to a process of change. Letting go of self to identify with another individual, or with a third-person perspective on a complex situation, or with a divine paradigm, allows a subtle and gradual transformation to proceed, a transformation with bright hope for the individual

and the planet. The Rule begins by suggesting that the way the agent wants to be treated can function as a standard of conduct; but by placing the other on a par with the self, the Rule engages the agent in approximating a higher perspective from which the kinship of humanity becomes evident. To pursue this higher perspective leads to the possibility of encountering the divine and the realization that every step along the forward path is illuminated by the Creator.<sup>ii</sup>

A bit of reasoning will show the way to a series of levels of interpretation of the Golden Rule. The levels mark a unified process of growth that integrates psychology, philosophy, and religion. The levels also furnish questions for comparative cultural study.

**The level of sympathy.** The Golden Rule expressed as a rule of sympathy runs, *“Treat others with sympathy and consideration, as you want others to treat you.”* Children as young as eighteen months can imagine themselves in another person’s place (for example, feeling jealous of another child on the mother’s lap). Children that young can also behave sympathetically, seeking to relieve the distress of another child (for example, bringing a toy to a child that is crying or going to get that child’s mother). As children grow up, one of the early lessons is to be considerate of others’ feelings. Ironically, children are taught not to hurt others sometimes by parents who ask harshly, *“Is that the way you want others to treat you?”* Tones of threat or condemnation accompanying the message may become associated in the child’s mind with the Golden Rule itself and with morality generally. Erik Erikson warned about the suppressed rage that can be instilled in children by parental moralism. The wise parent learns loving and sympathetic ways to teach consideration for the feelings of others.

Placing sympathy as the first in a sequence of ascending levels makes it necessary to insist that if a Golden Rule of sympathy were universally followed, it would put an end to the indifference and cruelty that mar our world. An expanded concept of sympathy is enriched by various cultures and disciplines. For example, sympathy is a biologically prompted emotion (e.g., mammalian mother-love) which, at its height, may be guided by reason and enlightened by spiritual compassion. Sympathy is not left behind as we develop.

Nevertheless, sympathy by itself is not a sufficient guide to morality, since sympathy can be short-sighted. Pity, which can develop from sympathy, is often unhelpful and loses sight of the dignity of the other person. False sympathy leads parents to shield children from the consequences of their actions. Nor are sympathy and pity reliable guides to political wisdom; sensational media easily rouse these emotions, but to engage the public in thoughtful consideration of problems is a different kind of enterprise.

The need to join sympathy with reason becomes more evident as we consider a practice widely associated with the Golden Rule: imagining oneself in another person's situation. The practice is beneficial insofar as it often rouses us from complacent self-centeredness and brings to mind what we know and imagine of the other. Research has found, moreover, that persons tend to be more effective in helping those in need when they do not merely fill their minds with the impression of the other person suffering, but rather imagine themselves in the other person's *situation*, seeing the world through their eyes, which conduces to a less emotional and more practical approach to the person and problem at hand.<sup>iii</sup>

But the imaginative role-reversal also has limits. First, it is neither necessary nor sufficient for moral discernment. Intuitive empathy may be fully adequate without any deliberate effort of imagining. In addition, merely *imagining* oneself in the other person's situation may not yield an adequate comprehension of the other. There are cognitive limits to imagining, which gives no magic access to new knowledge, though it does give us a chance to recall and give weight to what we already know about the other person. The more the other is a stranger different from the agent, the less information is carried by the act of imagination, but a sensitive effort to imagine can remind one of one's ignorance. It may be necessary to get to know the other through personal acquaintance, the study of history, psychology, and sociology, exposure to the arts, and so on. Moving beyond imagination to dramatic role playing in which persons act out the roles of nurse and patient, for example, can be very educational; and one discovers how important it is to get feedback from the recipient to confirm or correct one's intuitive understanding. Still, even a good understanding of the other does not suffice for moral clarity or moral motivation. And viewing oneself through the eyes of others can be overdone, as becomes clear in W. E. B. DuBois's 1903 classic, *The Souls of Black Folk*.

After the Egyptian and Indian, the Greek and Roman, the Teuton and Mongolian, the Negro is a sort of seventh son, born with a veil, and gifted with second-sight in the American world—a world which yields him no true self-consciousness, but only lets him see himself through the revelation of the other world. It is a peculiar sensation, the double-consciousness, this sense of always looking at one's self

through the eyes of others, of measuring one's soul by the tape of a world that looks on in amused contempt and pity.<sup>iv</sup>

The final caveat about imagining oneself in another's situation is implied in Joseph Conrad's novel, *Heart of Darkness*, which portrays the risk of empathizing with someone who is monstrously immoral.

**The level of reason.** If the Golden Rule is to function as the rule of living, it must symbolize a broad range of virtues indeed, including those that could be grouped under the heading of scientific living. Moral conduct is not only well intentioned; it is also intelligent, looking well into the future. It is common in ethics to observe that one needs to be aware of the facts of the situation to decide rightly, and countless applied ethics texts give one-paragraph situation descriptions in order to set up the problem to be solved. How would we prefer to be treated: on the basis of an intuitive and inadequate understanding or on the basis of a reasonable understanding of our long-range welfare? When we go to a physician, we expect the latter, so far as our physical welfare is concerned. A career counselor should have information about career prospects for the range of fields that the client is considering. A well-informed parent knows something of developmental psychology. The desire to show goodness apart from truth is empty. Utilitarianism is sometimes faulted for requiring decision-makers to achieve impossible calculations of future effects of alternative policies of action, but the theory should be credited for bringing decision-makers face to face with one component of responsible deliberation. "Deliberate slowly, act quickly," admonishes Aristotle, and the Golden Rule should not be embraced as a short-cut when there is time and need for deliberation.

And Kant tells agents to do all the good they can, which requires far more than consulting the pure, apriori deliverances of moral reason; and if his third major formulation of the categorical imperative in the *Groundwork* were adjusted in the light of his understanding of historical and political evolution, it might read, “Act so that the maxim of your action, if generally followed, would conduce to progress toward an advanced civilization.” Clearly such a principle calls for historical wisdom.

In addition to the need for Golden Rule conduct to be empirically well-informed, there is a need for the Rule to come into its own as a properly moral principle: “*Treat others in the light of moral reason, as you would have others treat you.*” Thus, a summary list of levels of meaning of the Golden Rule can be further articulated by developing distinctions such as that between scientific and moral reason. When Aristotle talks about the beautiful or noble deed, when Kant talks about the awesome fact and value of human dignity, when Mill appeals to what has been called an ideal observer’s perspective for weighing the interests of all beings affected by a course of action, when Hegel portrays human freedom becoming real through family life, economic activity, and political participation, we see moral reason coming into its own. Kant and Mill both regarded the Golden Rule as a popular maxim that approximates what their theories were designed to articulate philosophically. The Rule engages us in the exercise of moral reason, and as we discover problems in texts of Aristotle, Kant, Mill, Hegel, or other thinkers, we will inevitably find the need to add or subtract, to give a fresh restatement or at least a contemporary interpretation, even if we value a particular text as scripture.

Our very concept of a moral principle is expanded by a historical review of the meanings of the Golden Rule. The first set of meanings comes from Confucian tradition.

A moral principle can provide the key to—or symbolizes—some central virtue of character. A moral principle may be grouped with others in a cluster that expresses the core of morality. A moral principle can express a norm of relationships. Living in accord with moral principle is essential to the flourishing of social groups. One can use a moral principle throughout one's lifetime. Since a lifetime involves growth to maturity through various stages, one's experience of the practice of the principle will undergo transformation in the process. Accordingly, the meanings of such a principle may be multiple and non-static. Not only does one pursue different objects, but the very structure of one's motivation changes, too. Since, at the summit of moral maturity, one's conduct of relationships is not primarily motivated by restraint in conscientious conformity to a rule, a moral principle leads beyond itself, leads to loving spontaneity, bringing a certain transcendence of the moral standpoint from which it was initially undertaken. The next two meanings come from ancient Greco-Roman culture: a moral principle may orient the agent in a way that is generally consistent with prudence born of social-psychological insight; and there is more to a moral principle than a maxim of prudence: justice and fairness are expressed (or at least not violated) in a moral principle. Classical Jewish thought adds four more meanings: A moral principle may function as a summary connoting the best of one's tradition. Divine conduct can be the paradigm for the practice of a moral principle. A moral principle can govern one's relations with all people, and all people may be expected to abide by a moral principle. The New Testament adds that the necessary prerequisite for fruitful and sustained living in accord with moral principle is spiritual transformation by grace through faith. From medieval and early modern thought we can add that a moral principle may be inherently accessible in the mind of every

person. A moral principle can express an eternal pattern. A moral principle may express the best of human sentiment. A moral principle is rational at least in that it is intelligible and provides a guide to moral reflection. And—what is most problematic for the Golden Rule—a moral principle has no counterexamples; it provides a necessary and /or sufficient basis for sound moral judgment. During the past century, the following meanings have become explicit. A moral principle may express the practical import of central religious truth. Moral principles never contradict each other. A moral principle expresses the universalizability and prescriptivity implicit in the use of moral language. The ideal moral principle may make minimal assumptions and permit the derivation of maximal practical consequences. In particular, a moral principle may be a universally available standard (i.e., one with no religious presuppositions). Religious meanings and values associated with a moral principle need not be explicit in order for productive moral discussion to operate. And finally, a moral principle can enhance sensitivity to the problems of domination over the other, moralistic presumption, and the tendency of action to inflict passivity, if not suffering, on the recipient.

One of the ironies of the Golden Rule is that, at the hands of a certain kind of reason, which Hegel would call understanding rather than reason (*Verstand*, not *Vernunft*), dozens of objections have been raised, complaining that the standard implied by the Rule was too low, too high, or too vague. In quest of an “ethical theory” that would give necessary and sufficient conditions for moral judgment, some way to complete the sentence, “An action is right if and only if . . .” it became popular to reject the golden rule because of counterexamples that could be alleged. In response, one is forced to reject, reformulate, or re-interpret the Golden Rule.

Let's consider two objections. First, the Golden Rule naively presupposes that we are alike; as George Bernard Shaw wrote, "Don't treat others as you want others to treat you—their tastes may be different." On the surface, this objection may be handled by considering how we treat a friend at the ice-cream store: we don't order the flavor we prefer, but the flavor our friend prefers. In other words, we generalize the Rule to an intuitively appropriate degree. Put more deeply, this objection points to the need to treat others with due regard for their individual uniqueness and for any relevant specific characteristics, sex, racial composition, social and economic class, age, condition of health, culture, education, and so on.

The second objection generates another level of meaning in the Golden Rule. The objection alleges that the Rule tells a person with base desires to treat others so as to gratify the same desires. After Augustine's example of a person who wants to get drunk with a companion and Alan Gewirth's example of a person who wants to sleep with his neighbor's wife, the current version of the objection consorts with perversity. "What if a sadomasochist were to treat others as he wants others to treat him?" An adequate reply includes several points. First of all, there is a remarkable recursive feature to the Golden Rule which enables it to rebound resiliently from objections. I do not want to be treated in accord with a rule interpreted to yield bad results, so I must reinterpret the rule to use it. Second, the critic is neglecting to ask what the rule means, for example, in the teaching of Confucius, Hillel, or Jesus, and is asking only what its words can be made to mean if abstracted from every context. Third, the Golden Rule is addressed to *you*, not to a hypothetical third person, and it presupposes a certain level of moral maturity in the agent. Fourth, this is indeed one interpretation of the Rule—the lowest one; however, the

Rule has a moral flavor associated with it that is inconsistent with such applications. Nevertheless, we human beings do project distorted desires onto others who sometimes do share those desires. The Golden Rule interpreted on the level of the flesh sometimes describes us far better than we are ready to acknowledge. In addition, people have acted abusively when attempting to apply the Golden Rule. And we sometimes do conform to moral norms mainly because of self-interest. Fifth, if a sadomasochist were to make a sincere effort to apply the Golden Rule in every area of life, he would probably grow in the sympathy and self-respect presupposed by the Rule. Let it be made clear that the normal physical desires are part of the Creator's provision for us, and they have a genuine role to play in the functioning of superbly integrated men and women. It is only the distortion and wrong expression of those desires that is here under discussion.

There are many reasons to seek a level of meaning in the golden rule beyond the level of moral reason. Sometimes even the most exhaustive study of circumstances, principles, and patterns does not suffice to indicate what is to be done, and the mind seeks for higher wisdom. Moral reason points beyond itself, since it is committed to choosing the best course of action, *all things considered*. In attempting to take an ideal perspective in a complex situation, it is natural for the human mind to reach for an *actual* ideal, for God. Moreover, even when *what* is to be done is clear, the graciously spontaneous *way* to act may remain beyond one's powers. If moral decision and action are to be wholehearted, they must draw on the full range of the personality, not only the mind, and must respond to the fullness of the other person, including the spiritual dimension. In addition, many people need religious motivation to fuel the engine of moral conduct. Indeed, a reasonable, do-your-best ethic is misleading if grace is needed to do your best.<sup>v</sup>

**The level of love.** The ideals implicit in the Golden Rule culminate here. Love is characterized as the true and beautiful and good way of relating as members in a universal family. The Golden Rule is the principle of the practice of the family of God. The only time in history when the Golden Rule seems to have approached the status of a slogan for a mass movement was during the late nineteenth and early twentieth centuries in the United States, when there was widespread allegiance to the gospel of the fatherhood of God and the brotherhood of man. This trend within liberal Protestant Christianity combined with America's earlier religious heritages to give the Golden Rule a special place in the minds and hearts of untold numbers of persons. It is well to recall the functions performed by the combined concepts of the fatherhood of God and the brotherhood of man. The teaching implies the unity and personality of God. The concept of God as parent preserves the thought that God transcends the believer yet suggests that God is close, that we may experience the divine presence. The sequence of components in the slogan implies the primacy of the relationship of the individual with God. The phrasing does not privilege any religion. The brotherhood of man is the social consequence of the individual's relationship with the Creator. Talk of brotherhood addresses the special challenges of modernity, to dissolve the forces that tear the fabric of humankind: religious intolerance, nationalism, racism, sexism, economic and political injustice, and so on. And "the fatherhood of God and the brotherhood of man" sets forth those components of religion with the universal logic of family life.<sup>vi</sup> The teaching of this movement can be simply put. The Golden Rule is the principle of the practice of the family of God.

What relations obtain between a Golden Rule of brotherly and sisterly love and a Golden Rule of moral reason? From the perspective of modern ethics in general and Kant in particular, any religious interpretation would be no more than a supplement, adding meaning and motivation. Religion, moreover, must take care not to interfere, since bad religion betrays moral reason. What is sought is the power of an integrating alignment of spiritual faith and moral reason. But Kant did not acknowledge the fragility of moral reason, its need for a religious foundation. Kant testifies with all his heart and mind and soul and strength: “Now I say that man, and in general every rational being, exists as an end in himself and not merely as a means to be arbitrarily used by this or that will.” Kant proclaims the infinite dignity of the humanity in each person—empirical evidence seemingly to the contrary notwithstanding. His affirmation is an article of faith. At the same time he is agnostic about ever being able to tell whether we truly act from that motivation which alone lifts us above our material nature. Thus what we can lucidly certify is nothing actual, only our *potential* for rational self-governance. Because our own conduct so often betrays our birthright as sons and daughters of God, it is very easy to lose sight of our dignity as creatures with moral reason, which is God’s gift to us all. Kant’s celebration of the infinite worth of the humanity “that indwells us” is a rationalist amalgam, launched by Plato, assimilating reason and spirit.<sup>vii</sup> But if we are actually the family of God, if the divine spirit actually within each of us, and if each unique and mysterious personality is actually a masterpiece of the Creator’s art, then we have more robust grounds for affirming one another and more leverage for the labor of growth. There was an unacknowledged and partly mutilated religious layer beneath Kant’s rationalism. What remains common is that ethicists assume, without argument, the equal,

basic worth of persons.<sup>viii</sup> I propose that there is a religious ground for the flourishing of moral reason: the source of human dignity is our common origin in the Creator. One may hazard the hypothesis that without a religiously based sense of human kinship, the drive to assert and actualize equality on other grounds courts error and frustration. In a world where nationalism, racism, and domestic violence are so widespread, given the logical and practical fragility of the affirmation of human dignity, moral reason is wise to acknowledge a religious basis for the recognition of the infinite worth of the individual.

There is another level of love, another practice of the family of God, and that is to love another with the love that comes from God. This goes beyond brotherly and sisterly love to parental love, fatherly and motherly love. Leviticus 19 gives as the word of God, “Be holy as I am holy.” Matthew 5 teaches, “Be you perfect as your Father in heaven is perfect.” In various ways, the human being is called to be like God in the measure possible for a mortal. A parent is willing to do far more than a brother or sister is willing to do (in the typical case that I assume here). The other person is seen as a divinely created, infinitely loved, spiritually indwelt, son or daughter of God.

Divinely parental love goes beyond reciprocity in one sense. Anthropologically speaking, the expectation that the good we do to others merits good returned to us is well-nigh universal in human communities. Even an unconscious expectation of receiving good in return may cast a shadow on our altruism, but this darkness departs in the sunrise of the motivation of divine love. This is not to suggest that God does not desire communion with us, only that his blessings are not given with strings attached, requiring that we know, on each occasion, whence they come and formally express, on each occasion our gratitude. Nevertheless, reciprocity in another sense persists, for the Golden

Rule invites the agent to extend to others the same attitude of service that one would welcome as the recipient of someone else's divinely parental love in the same kind of situation.

In God, whose spirit dwells within, we find the source of boundless love, and to God we turn for wisdom to guide our expression of love. When that wisdom meets that motive, the summit of the Golden Rule is reached.

To express the Golden Rule in terms of sympathy, reason, and love needs three comments. First, it is not a hermeneutical tool that can be assumed to apply in any obvious way to the Golden Rule's diverse expressions, many of which cannot be simply classified. In such cases this sequence can often simply guide inquiry into what dimensions may be present.

Second, it might seem that moral reason has been quite transcended here by the agent flooded with grace, but this is not the case. Rather, there is something like a game of leapfrog going on between our growth in moral reason and our growth in spiritual experience. In this game, the first child goes forward by leaping over the child just in front of him. Then the first child kneels down, and the second one leaps over him. Every philosophic advance prepares a new religious advance, and vice-versa.

An agent seeking to find and do the will of God may use three principles for prayer that integrate well with the best of Kantian moral reason.

1. In seeking to know the will of God, one should not expect prayer to replace an intelligent study of the situation and an exhaustive effort to make the necessary adjustments. One does not expect the Deity to do one's homework, and one needs to be prepared carefully to work out the details of what is to be done. Scientific and

philosophic thinking have a role to play in the prayer process. They are taken to the limit, not marginalized, by the responsible person seeking the will of the Creator of a universe where, despite mystery and uncertainty, one may observe dependable laws of matter and mind.

2. One prays for the growth that the problematic situation requires, for assistance in the quest for perfection. Such religious hope is already part of the Kantian program.<sup>ix</sup>

3. At the summit of the prayer process, having done one's utmost to think through the problem, one opens the mind to receive a higher wisdom. As one's prayer life matures, there increasingly comes an influx of truth, beauty, and goodness, a new perspective on the situation. However, one does not relinquish the responsibility to assess the meanings and values of what comes to mind during this time. If energies from the subconscious stream in, fine—we can use these energies. If new ideas and impulses arise, they will be reflectively reviewed before the decision is made. Given the dangers of fanaticism, it is safer to risk rejecting a spiritual input than to mistake a personal impulse for the divine lure. Thus Kant's requirement of autonomy is not compromised. Responsible freedom and the dignity of human personality are preserved in this relation between the Creator Father and the creature son or daughter. Reason—our best thinking—is enhanced, not overturned, in the process. This integration of emerging reason and emerging spiritual experience protects against fanaticism.<sup>x</sup>

The last comment highlights one more phenomenon of the life of the Golden Rule: it does not draw attention to itself, but directs attention into the texture of what is to be done and the person who is to be served. Rule consciousness vanishes in self-forgetful service. One does not need continually to remind oneself of the rule of living. Just as

empathy often makes it unnecessary deliberately to put oneself in the other person's shoes, so the agent who lives the Golden Rule is increasingly immersed in sympathy, long-range prudence on the other's behalf, brotherly or sisterly solidarity, moral reason, and parental love infused by God. Repeated work with this sequence or any worthy path of thoroughness returns one to the way of simplicity. In 4<sup>th</sup> century B. C. E. China, Mencius put it thus: "A noble man steepes himself in the Way (*tao*) because he wishes to find it in himself. When he finds it in himself, he will be at ease in it; when he is at ease in it, he can draw deeply upon it; when he can draw deeply upon it, he finds its source wherever he turns." Truly, the life in the Golden Rule, the rule of living, derives from the Way that resides therein.

---

<sup>i</sup> This paragraph is adapted from the conclusion of Jeffrey Wattles, *The Golden Rule* (New York: Oxford University Press, 1996), pp. 189-90. This conclusion, key paragraphs from the other chapters, and selected chapters are available online: <http://www.personal.kent.edu/~jwattles/goldrule.htm> . I am grateful to Philip Rolnick for suggestions and conversation on an earlier version of this paper.

<sup>ii</sup> *The Golden Rule*, loc cit.

<sup>iii</sup> *The Golden Rule*, p. 119.

<sup>iv</sup> W. E. B. Dubois, *The Souls of Black Folk* [1903] 1990, pp. 8-9.

<sup>v</sup> *The Golden Rule*, p. 182.

<sup>vi</sup> *The Golden Rule*, p. 91-92. Since the heyday of the gospel of the Fatherhood of God and the brotherhood of man, critics have challenged each of those core concepts. If Jesus' concept of sonship with God can be interpreted as embracing a plurality of

---

meanings, and if the Father concept of God can be grasped not only as metaphor but also as analogy, Christianity could regain a synthesis of these key elements. See

<http://www.personal.kent.edu/~jwattles/unifam.htm> and

<http://fp.dl.kent.edu/jwattles/cgospel.htm> .

<sup>vii</sup> Hinduism speaks of the *atman*, the eternal spirit self (*Bhagavad-Gita*, chapter 2).

Mahayana Buddhism speaks of the Buddha nature within (*Mahaparinirvana Sutra*).

Judaism speaks of “the spirit in man, the candle of the Lord, searching all the inward parts” (*Proverbs* 20.27). Jesus said, “The kingdom of God is within you” (*Luke* 17.21).

Islam speaks of God as “closer to you than the vein of your neck” (*Qur’an* 50.17).

Although Steven T. Katz, for example, has argued in *Mysticism and Religious Traditions* that contextual differences make it problematic to posit the same reality coming to expression in these various texts (, it is also possible to hypothesize that they are indeed giving voice to a shared insight.

<sup>viii</sup> For a discussion of ten examples of ethicists who give no grounds for the thesis of equal basic human worth, see Louis Pojman, “Are Human Rights Based on Equal Human Worth?” *Philosophy and Phenomenological Research* (52.3) (1992), 602-622.

<sup>ix</sup> Immanuel Kant, *Metaphysics of Morals*, The Doctrine of Virtue, #21; *Religion within the Limits of Reason Alone* (New York: Harper & Row, 1960), p. 47).

<sup>x</sup> The discussion of Kant and prayer is adapted from my conference presentation, which appeared as “Religious Experience, Fanaticism, and Kant” in R. Blake Michael, ed., *The Ohio Academy of Religion Scholarly Papers 2002* (Delaware, Ohio: Ohio Wesleyan University: 2002), 27-34 (<http://www.personal.kent.edu/~jwattles/relexprk.htm>) .

## 6.

# THE GOLDEN RULES OF RELIGION

DAVID SLOAN WILSON

BINGHAMTON UNIVERSITY

STATE UNIVERSITY OF NEW YORK

### **Religion, Evolution, and the Revival of Functionalism**

Religion is often defined in terms of belief in supernatural agents, but Emile Durkheim offered a different definition in *The Elementary Forms of Religious Life* (1912/1995:44).

A religion is a unified system of beliefs and practices relative to sacred things...which unite into one single moral community called a Church, all those who adhere to them.

Durkheim's definition emphasizes the functional nature of religion, in which supernatural agents and all else sacred figure as proximate mechanisms, the means to an end of a united community. More generally, Durkheim helped to initiate the tradition of functionalism in the human social sciences, which interprets all cultural systems as

largely adaptive for their members. As he put it, “In all its aspects and at every moment of history, social life is only possible thanks to a vast symbolism (1912/1995: 233)”.

Functionalism flourished during the first half of the 20<sup>th</sup> century but then fell upon hard times. Part of the problem was that group-level functionality was assumed, rather stated as a prediction that could be falsified. Functionalists also had a static view of society that seemed to deny the possibility of change or the importance of individual agency. Finally, holistic conceptions of society seemed unscientific and even mystical, compared to more reductionistic conceptions, which became known as methodological individualism (see Wilson 2002, ch 2 for a review).

Despite these failures, developments in evolutionary biology suggest that the rejection of functionalism in the human social sciences was premature. Natural selection proceeds not only by small mutational steps, but also by social groups becoming so highly integrated that they become higher-level organisms in their own right (Maynard Smith and Szathmary 1995). This phenomenon is called a major transition of evolution and has occurred repeatedly throughout the history of life, perhaps including the origin of life itself as groups of cooperating molecular reactions. The individual organisms of today are not only the groups of past ages, but the groups-of-groups-of-groups of past ages!

The evolution of social insect colonies such as ants, bees, wasps and termites also fall within the paradigm of major transitions. The members of a honeybee colony are physically separate, unlike the cells of a multi-cellular organism, but their activities are so well coordinated for collective survival and reproduction that they qualify as a superorganism. They even possess a collective mind, in which single bees play the same kind of limited role in a social network that single neurons play in a neural network. The concept of a group mind might have appeared unscientific and mystical to those who rejected functionalism in the mid-1900's, but it has been demonstrated in exquisite detail by social insect biologists such as Thomas Seeley and his colleagues (e.g., Seeley 1995).

It is becoming increasingly certain that human evolution represents a major transition of evolution, turning our ancestral groups into the primate equivalent of bodies and beehives. These groups had their share of internal strife (which also exists in bodies and beehives) but they succeeded primarily on the basis of their teamwork in competition

with other groups, other species, and their challenging physical environment. Teamwork included not only physical tasks such as gathering, hunting, warfare, and childcare, but also mental tasks such as perception, memory, and decision-making. The capacities for symbolic thought and the social transmission of behavior, two hallmarks of our species, are fundamentally communal activities that require mental teamwork. Cultural evolution enabled our ancestors to spread out of Africa, inhabiting all climatic zones and occupying hundreds of ecological niches, harvesting everything from seeds to whales. Then the invention of agriculture initiated a positive feedback loop with respect to the production of resources and population growth, leading to our current large-scale societies. Against this background, Durkheim's statement that "In all its aspects and at every moment of history, social life is only possible thanks to a vast symbolism" acquires a modern ring.

### **Implications of the New Functionalism for our Understanding of the Golden Rule**

These broad developments are described in more detail elsewhere (Wilson 2002, 2005, 2007; also visit this website designed to establish Evolutionary Religious Studies as a new discipline: <http://evolution.binghamton.edu/religion/>). Here I will focus on the implications of the new functionalism for our understanding of the Golden Rule. A number of points can be made which are so elementary from a modern functionalist perspective that they are unlikely to be wrong.

- 1) To the extent that religions are adaptive at the group level, they must encourage cooperation and inhibit exploitation among their members. If we equate the Golden Rule very loosely with "in-group cooperation," then it can indeed be said to lie at the core of all religions—to the extent that they are adaptive at the group level. See point 6 for a discussion of when religions are *not* adaptive at the group level.
- 2) For a religion to be adaptive at the group level, it must do much more than exhort its members to be cooperative in an undifferentiated fashion. It must provide a more detailed set of if-then rules for how to behave in particular contexts. The metaphor of society as an organism becomes highly instructive in this regard when taken seriously. Single

organisms are mind-bogglingly complex in their genetic replication machinery, anatomy, physiology, and behavior—and *must be*--to survive and reproduce in their respective environments. If cultural systems are even remotely like single organisms in their functional organization, then they must provide something comparable for their members. There is no single Golden Rule of religion, but Golden *Rules*, and lots of them.

3) One of the most profound implications of the new functionalism is that *human cultural diversity is comparable to biological diversity*. Genetically we are a single species, but culturally we are more like a multi-species ecosystem. Evolutionary biologists don't spend a lot of time studying what all species share in common. Their main challenge is to explain why species are so different from each other. Similarly, the study of religion must go beyond the search for universals to include the interpretation of differences. Not only does a single religion have many Golden Rules, but different religions have different Golden Rules—and *must*--to become adapted to their respective environments. Pointing out that all religions are internally cooperative (to the extent that they are adaptive at the group level) is like pointing out that all organisms have an internally harmonious anatomy and physiology. It is a true statement, but it doesn't get us very far. We need to know about the anatomical and physiological adaptations of particular organisms—and their analogs for religious (and other cultural) systems.

4) The many Golden Rules of any particular religion are not necessarily derivable from a single rule such as the Golden Rule. In his contribution to this volume, philosopher Harry Gensler shows how some applications of the Golden Rule lead to pathological outcomes (“fool's gold”) and attempts a general formulation that avoids these problems. His goal is to make the Golden Rule a single principle that, properly understood, can be used to derive specific behaviors in any particular situation. This is a worthwhile philosophical exercise, but most religions are not designed that way. As an alternative, consider the two proverbs “A stitch in time saves nine” and “If it ain't broke, don't fix it.” They are the opposite of each other, but each prescribes a useful behavior in a given context. Precaution is great when we have the time for it, but sometimes it must take a back seat to more pressing concerns. People who use these proverbs don't puzzle over the

fact that they are inconsistent with each other, or try to create a single general rule from which both proverbs can be rationally derived. They simply invoke the appropriate proverb in the appropriate context. Religious systems are more like collections of proverbs than philosophical systems that strive for consistency across contexts. Cultural systems in general are partially the product of blind variation and selective retention (Campbell ). They work without being consciously designed and without anyone knowing how they work. Alexis d'Toqueville (1835/1990) pointed out that the Mexican Constitution was copied from the American Constitution, but Mexico is not at all like America. Something makes them different, which in our ignorance we call “custom,” that we don't understand but makes a huge difference in how well these societies function as corporate units. The same can be said for the adaptive organization of religions. Moreover, even when elements of religion are consciously designed, they do not necessarily take the form of a system that can be rationally derived from a small set of rules, as I will show for the example of Calvinism below.

5) The Golden Rule is often described as capturing the true spirit of religion, as if religion in its purest sense is about universal brotherhood, even if most religions fall short of the ideal. The contribution to this volume by philosopher Jeffrey Wattles is in this spirit and shows how The Golden Rule might be used as the basis for a global value system of the future. I share Wattles' desire to foster universal brotherhood, but it is also important to distinguish this normative project from the more descriptive project of understanding the nature of religions of the past, present, and near future. Descriptively, it must be acknowledged that most religions are designed to promote, not universal brotherhood, but the collective interests of their members. There is nothing that makes the Golden Rule more essential than other rules prescribed by a religion. It is employed selectively within the group and even more selectively toward outsiders. The details are expected to vary greatly, depending upon environmental circumstances. Once again, an analogy with biological species is instructive. Interactions among species span the full range, from ruthless competition and predation, to coexisting without direct interactions, to obligate mutualisms. We should expect the same spectrum of interactions among cultural systems, each requiring their own sets of if-then rules.

6) The five points listed above are based upon a conception of religion as a group-level adaptation. However, the elements of religion are not invariably adaptive at the group level. They can also benefit some individuals at the expense of others within the same group, benefit cultural traits at the expense of both individuals and groups (similar to a disease organism), or have no function at all, since there is more to evolution than adaptation, as the late Stephen Jay Gould tirelessly pointed out. It is a strength of the new functionalism that group-level adaptation is not assumed, but is predicted to occur only when certain conditions are met. For example, consider the Catholic practice of indulgences and other abuses that led to the Protestant Reformation. They were clearly exploitative practices that benefited some members of the religion (the elites) at the expense of others (the rank and file). More generally, many religions exhibit a cycle of corruption and renewal that is predicted from a fully rounded theory of natural selection as a multilevel process (Wilson 2002, p 182-187). When we consider elements of religion that are not adaptive at the group level, then the importance of the Golden Rule diminishes still further and even disappears altogether.

### **Three Case Studies**

It is beyond the scope of this article to expound upon these six points in detail. Instead, I will briefly illustrate them with three case studies.

**The Origin of Calvinism.** In my book *Darwin's Cathedral* (Wilson 2002, Ch. 3) I chose Calvinism as an example of how a specific religion can be studied from an evolutionary perspective. The origin of Calvinism represents a natural before-and-after experiment. The City of Geneva had expelled the Catholic Church and was trying to function as a corporate unit on the strength of a democratically elected civic government and undifferentiated Protestant zeal. It was failing; factionalism was called "the Genevan disease." Calvin and his colleagues established a more differentiated religious system that helped Geneva function better as a corporate unit and became a model for religious reform elsewhere. What were the specific ingredients that caused Calvinism to work in

this practical sense (what Durkheim called “secular utility”) and what was the role played by the Golden Rule?

Calvinism is especially useful as a case study because Calvin insisted on writing the catechism and set of rules (The Ecclesiastical Ordinances) as a condition for coming to Geneva. He also elaborated on his religious views in great detail elsewhere throughout his lifetime. Thus, if any religion has a chance of qualifying as an internally consistent philosophical system derivable from a few general rules, it is Calvinism.

The first time that the Golden Rule appears in the Calvin’s catechism is in section 9 (out of 33 sections; passages and page numbers are from Hesselink 1997):

The direction in which all the commandments of the law tend, Christ our Lord sufficiently declared when he taught that the whole law was comprised under two heads: “We are to love the Lord our God with our our heart, all our soul, and all our strength. Then we are to love our neighbor as ourselves (15).

Like Rabbi Hillel, who used The Golden Rule to summarize the Torah as briefly as possible, Calvin uses the phrase “Love thy neighbor as thyself” as the briefest possible summary of God’s commandment. The essence of the religion is to promote cooperation among members of the Church, as I stressed in point 1 above. However, just as Jacob Neusner’s contribution to this volume reveals that the Golden Rule was not treated as a foundational principle of Judaism from which specific prescriptions could be derived (and wasn’t even the main point of the parable of standing on one foot!), “love thy neighbor” does very little useful work in Calvin’s catechism. Much more important is the cultivation of a state of mind that combines fear of damnation with gratitude for the prospects of being saved, resulting (if the belief system is taken seriously) in a permanent state of contrition. Calvin also works hard to stress that the human intellect is so feeble that it is impossible to know God’s will and blasphemy to try, paving the way for specific prescriptions that do not require rational justification.

In point 2, I stress that adaptive religions must go beyond generic cooperation and provide specific if-then rules for how to behave in particular situations. As an example,

Calvinism was impressively designed to control deviance in its own leaders, which is reflected in the following passage from the catechism:

Therefore, pastors may dare boldly to do all things by God's Word, whose stewards they have been appointed to be...But let them turn aside from this to their own dreams and figments of their own brains, then they are no longer to be considered to be pastors but rather as pestilential wolves to be driven out. For Christ does not command others to be heard than those who teach us what they have taken from his word (36).

In contrast, Calvin's church was in a subordinate relationship with the civic government and was not in a position to punish secular rulers, which is also reflected in the catechism:

Not only should we behave obediently toward those leaders who perform their office uprightly and faithfully as they ought, but also it is fitting to endure those who insolently abuse their power, until freed from their yoke by a lawful order. For as a good prince is proof of divine beneficence for the preservation of human welfare, so a bad and wicked ruler is his whip to chastise the peoples' transgressions (38).

These two passages are much like the two proverbs described in point 4. They make opposite prescriptions for how to behave, but each is appropriate in their given context. There is no need to make them logically consistent or derivable from a single rule such as the Golden Rule. They merely need to be separately invoked whenever appropriate. The world is full of injustices that cannot be controlled. What better way to cope with them and turn them to advantage than by portraying them as God's way of punishing us for our own sins!

The second passage provides an example of what rationalists love to hate about religion. If all blessings are attributed to God's beneficence and all misfortunes are

portrayed as God's way of punishing us for our own sins, then religion becomes immune from skeptical inquiry. This seems weak-minded from a rationalist perspective, but the rationalist is weak-minded for failing to evaluate religious belief by the appropriate criterion— as a set of specific if-then rules for how to behave. Calvin's religion was highly sophisticated and adaptive when judged by the appropriate criterion, even if it doesn't take the form of a logically consistent system that enables a diversity of behaviors to be derived from a small set of rules.

**The problem with extending the Golden Rule to everyone:** Fast-forwarding to the present, Carlton Pearson can be regarded as a modern-day Calvin attempting to establish his own vision of Christianity—so far without much success. As reported in National Public Radio's *This American Life*

([http://www.thisamericanlife.org/Radio\\_Episode.aspx?episode=304](http://www.thisamericanlife.org/Radio_Episode.aspx?episode=304)) and in his own words (Pearson 2007), Pearson grew up in a Pentecostal community and gained an early reputation for being able to cast out the devil. He attended Oral Roberts University and became a protégé of Roberts himself, who regarded Pearson as his “black son,” capable of reaching an African-American audience. Before long he had become a Pentecostal Bishop, traveling the world and presiding over his own megachurch in Tulsa, Oklahoma. He also became knowledgeable about religious traditions other than his own. Then one day he had an epiphany while watching television with his infant daughter on his lap. A news program showed footage of an African famine, including starving babies with their swollen bellies. Suddenly his own belief that these babies would be sucked into hell if they didn't convert to his particular religious faith struck him as monstrous. In his own mind he had a conversation with God and realized that hell is not a place we might go after we die, but the misery that we create for each other on earth.

Pearson's revelation caused him to question other major tenets of Pentecostalism, such as the inerrancy of the Bible. Unfortunately, when he tried to preach his new gospel, he ended up losing his congregation and being branded a heretic. Slowly he began to acquire a new congregation, in much smaller numbers, of people from a Pentecostal background and who had been victims of intolerance and therefore resonated to Pearson's new “Gospel of Inclusion.”

In point 5, I stressed that The Golden Rule is often described as capturing the essence of religion, as if religion in its purest sense is about universal brotherhood, even if most religions fall short of the ideal. By this criterion, Pearson took a giant step forward with his revelation, so why did he take a giant step backward by losing his congregation? Because fear of disobeying God and certainty about what it means to obey God's will perform vital functions in Pentecostalism, no less than Calvinism. As one person who left Pearson's congregation put it, "We need to find another church that's solid in the word." As one of the youth pastors who stayed with Pearson humorously put it: "What's the best way to get the kids' attention? We'll scare 'em! We'll say, do you like to burn? ... That's how most of us got saved. We chose because the alternative was scary! ... Threat of judgment day sure is easy to pack a church out. That fear factor is definitely effective. If we take away the requirements of coming to church and paying your dues...you can put some guys out of a job!"

In point 4, I stressed that cultural systems can be regarded as like biological species that coexist by inhabiting separate niches. Pearson's new Gospel of Inclusion was poorly suited for his old congregation but well suited for a different set of people who had experienced intolerance. He still occupied a niche, albeit a much smaller one. Another interesting point is that Pearson's Gospel was new only against the background of Pentecostalism. Liberal Protestant denominations gave up believing in hell and Biblical literalism long ago. Yet, when Pearson had his revelation, he didn't convert from Pentecostalism to an existing liberal denomination, but felt the need to create a liberal Pentecostal denomination of his own. In this fashion, religions diversify in parallel in different cultural traditions, much as species diversify in parallel on different islands. None of these fascinating trends can be understood by portraying the Golden Rule as somehow capturing the true spirit of religion.

**Liberal and conservative Protestant denominations as separate cultural species.** A recent analysis by Ingrid Storm and myself illustrates the degree to which different religions can cause their members to behave as if they were different species (Storm and Wilson, 2008). The analysis makes use of a large representative sample of American teenagers, who were studied by Csikszentmihaly and Scheider (2001) to see how young

people prepare themselves to enter the work force. In addition to completing a large number of one-time survey items, data was also taken using the experience sampling method (ESM) which involves being signaled at random times during the day and reporting both external (where you are, who you are with, what you are doing) and internal (what you are thinking and how you are feeling on numerical scales) experience. The ESM is as close as psychological research gets to field studies of nonhuman species in their natural environments.

Our analysis focused on a comparison between students who belonged to liberal vs. conservative Protestant denominations. Thus, everyone was American, a teenager, and belonged to the same major religious tradition. In these respects they were culturally uniform. However, some belonged to liberal denominations (Episcopal, Methodist, Presbyterian and the United Church of Christ) and others belonged to conservative denominations (Southern Baptist, Nazarene, Assemblies of God, Seventh Day Adventists, and Mormon). This particular cultural difference had a transformative effect on their values and behavioral response to their environments, as shown in the following figures from Storm and Wilson (2008).

In figure 1, the survey item “Do you think of yourself as a religious person?” is related to the survey item “In my family, we express opinions even when they differ.” The more liberals agreed with the first item, the more they agreed with the second. The more conservatives agreed with the second item, the *less* they agreed with the second. Their respective religions were pulling them in completely different directions.

In figure 2, the survey item “In my family, I am the one to decide which friends I can spend time with” is related to the survey item “Do you usually feel stressed?” Liberals felt stressed when they couldn’t make their own decisions and became less stressed when provided elbowroom. Even at their most mellow, however, liberals felt more stressed than conservatives, who evidently didn’t require elbow room! This figure also illustrates the important point that even though fear plays a critical role in the belief systems of religious conservatives, they do not spend most of their time in a fearful condition, but rather in the more benign psychological state of feeling saved and in control of their environment.

The differences become even more interesting when we consider the moment-by-moment data from the ESM. Liberals spend considerably more time alone than conservatives. Moreover the positive mood of conservatives is highly dependent upon being in the presence of others, in contrast to liberals, who often prefer being alone, as shown for a sample of variables in figure 3.

In the biological literature, figures 1-3 would be called “norms of reaction,” which describe the phenotype of the organism (y-axis) across a range of environments (x-axis). Different species have different norms of reaction—*and must*—to survive and reproduce in their respective niches. The liberal and conservative teenagers in our sample obviously belong to the same biological species and mingle with each other in the same schools, but their respective cultures cause them to be profoundly different from each other in their norms of reaction at the phenotypic level, which is all that natural selection ever sees.

What are “niches” to which liberal and conservative cultural systems are adapted? Liberals place a high value on individual autonomy and decision-making. Individuals are expected to internalize the norms of their culture and do the right thing on a case-by-case basis after thinking about it. This strategy can be highly successful but can also be costly in the time required for information processing, in making mistakes, and in ignoring successful behaviors winnowed by tradition that work without anyone knowing why they work. Liberalism tends to thrive in safe, stable, affluent societies such as current-day Western Europe.

Conservatives place a high value on obedience to authority. This strategy might stifle creativity but has a number of advantages, such as easing the burden of information processing, retaining successful behaviors winnowed by tradition, and coordinated action. Conservatism tends to thrive in dangerous, unstable, and impoverished societies such as the Middle East. One reason that America is diverse with respect to liberalism and conservatism is because it is so stratified with respect to the safety, stability, and affluence of its citizens.

See Storm and Wilson (2008), Norris and Inglehart (2004), Lakoff (1996), and Jost et al (2001) for a more detailed discussion of liberalism and conservatism, in both their religious and non-religious manifestations. For the purpose of this article, the most important point to make is that the Golden Rule contributes nothing whatsoever to our

understanding of these profound cultural differences. As an undifferentiated appeal to cooperate, “Do unto others” is universal and therefore powerless to explain the nature of religious and cultural diversity.

## Conclusion

Prosocial values such as altruism and the Golden Rule are often portrayed as at the heart of all religions. In a previous symposium and edited volume on altruism across world religions (Neusner and Chilton 2005), William Green began by making the shocking claim that when altruism is defined as “intentional action ultimately for the welfare of others that entails at least the possibility of either no benefit or a loss to the actor,” it is foreign to the imagination of *all* the major religious traditions. Subsequent chapters by scholars of the major religious traditions confirmed Green’s claim. This does not mean that religions are bastions of selfishness, but rather that they do not conceptualize helping others as requiring sacrifice on the part of the individual. As Green put it in his summary statement, “only by a rigid secular calculus is benevolence less benevolent because the actor benefits.”

A similar shock seems to be in store for us in the case of the Golden Rule. It does provide a tidy summary of the core fact that most enduring religions are designed to foster cooperation among their members. It does not appear to serve as a core principle that is consulted to derive specific predictions for how to behave. To understand how religions actually work, we must think about Golden *Rules*, not a single Golden Rule.

## Literature Cited

Csikszentmihaly, M. and B. Schneider. 2000. Becoming Adult: How Teenagers Prepare for the World of Work. New York: Basic Books

Durkheim, E. (1912/1995) The Elementary Forms of Religious Life. New York, Free Press.

Jost, J. T., J. Glaser, A. W. Kruglanski and F. J. Sulloway (2001). Political Conservatism as Motivated Social Cognition. *Psychological Bulletin* 129(3): 339-375

Lakoff, G. (1996). Moral Politics: What Conservatives Know that Liberals Don't. Chicago: University of Chicago Press

Maynard Smith, J. and E. Szathmary (1995). The Major Transitions of Life. New York: W.H. Freeman.

Neusner, J. and B. D. Chilton (2005). Altruism in World Religions. Washington D.C.: Georgetown University Press.

Norris, P. and R. Inglehart (2004). Sacred and Secular: Religion and Politics Worldwide. Cambridge: Cambridge University Press

Pearson, C. (2007). The Gospel of Inclusion: Reaching Beyond Religious Fundamentalism to the True Love of God. Tulsa: Azusa Press/Council Oak Books

Seeley, T. (1995). The Wisdom of the Hive. Cambridge, MA: Harvard University Press.

Storm, I. and D.S. Wilson (2008). Liberal and Conservative Protestant Denominations as Different Socio-Ecological Strategies. Submitted.

Tocqueville, A. d. (1835/1990). Democracy in America. Garden City, NY: Anchor Books.

Wilson, D.S. (2002). Darwin's Cathedral: Evolution, Religion, and the Nature of Society. Chicago: University of Chicago Press..

Wilson, D. S. ( 2005). Testing major evolutionary hypotheses about religion with a random sample: *Human Nature*, 16: 382-409.

Wilson, D.S. (2007). *Evolution for Everyone: How Darwin's theory can change the way we think about our lives*. New York: Delacorte Press.

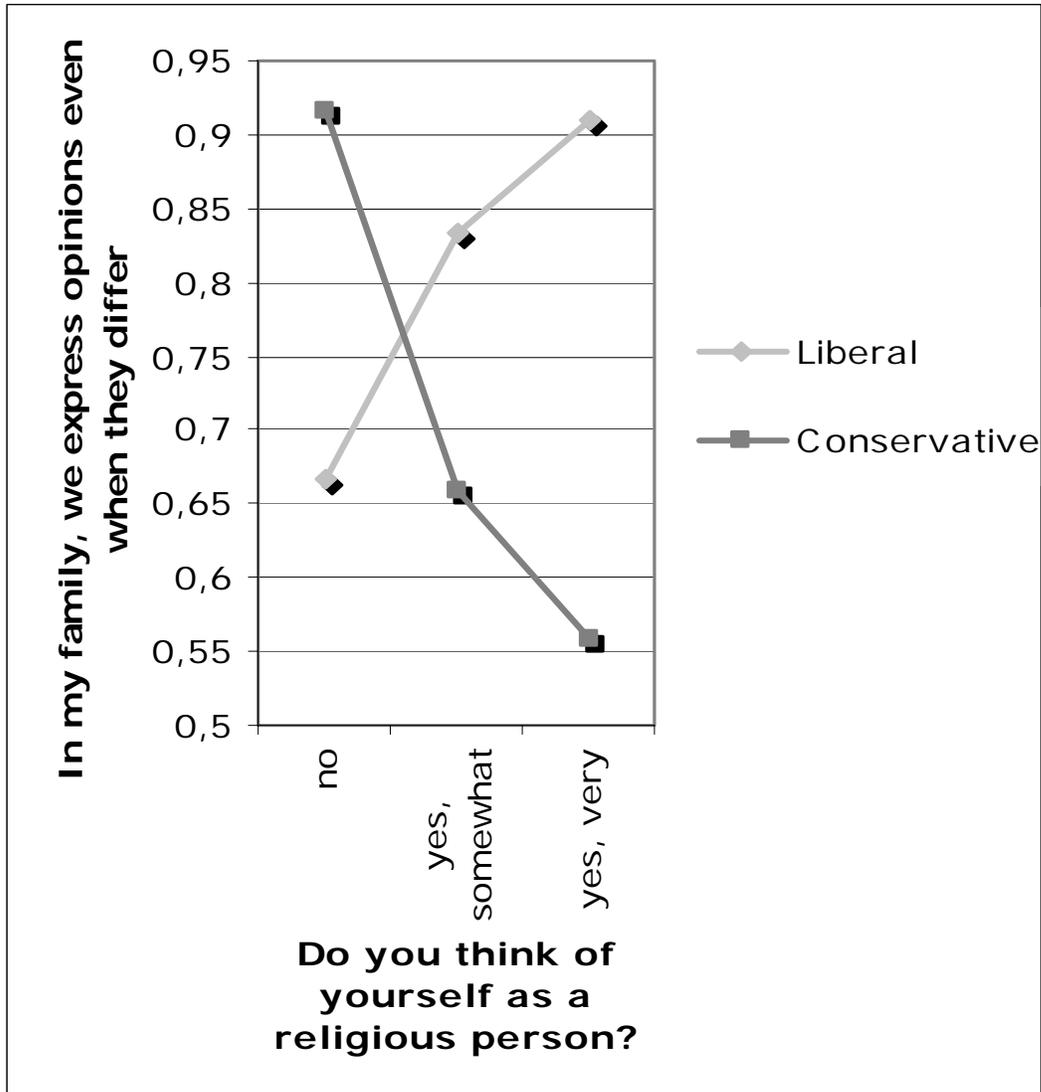


Figure 1. Response of American teenagers from Liberal and Conservative Protestant denominations to two survey items. See Storm and Wilson (2008) for details.

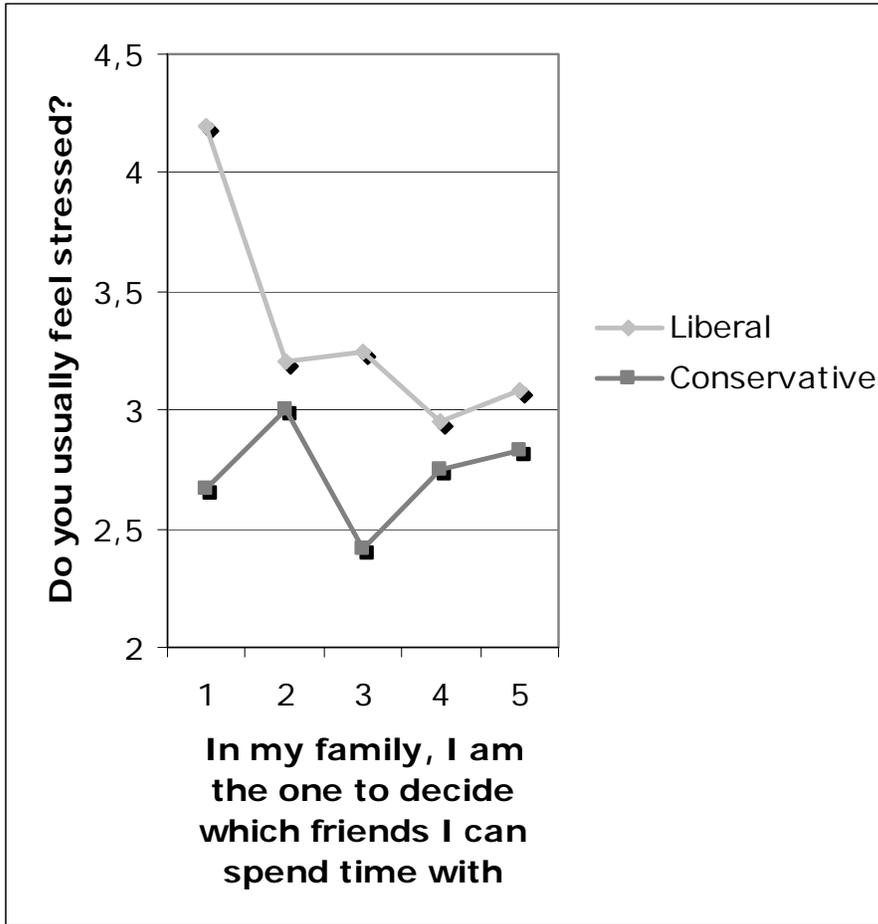


Figure 2. Response of American teenagers from Liberal and Conservative Protestant denominations to two survey items. See Storm and Wilson (2008) for details.

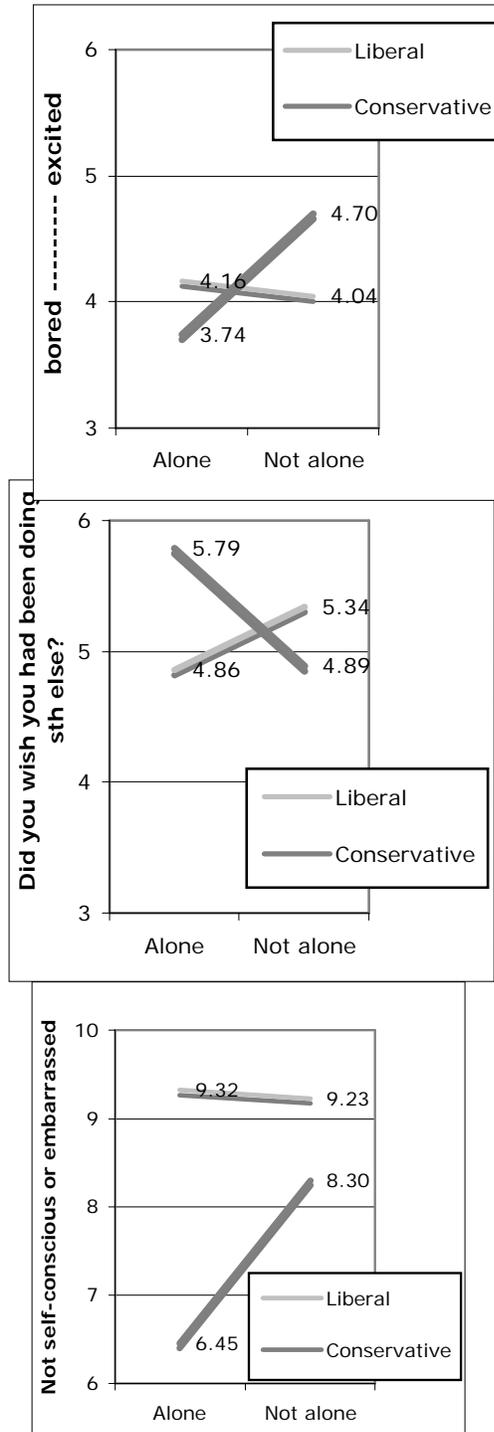


Figure 3. Response of American teenagers from Liberal and Conservative Protestant denominations when asked to assess their mood on a moment-by-moment basis, when alone and in the presence of others. See Storm and Wilson (2008) for more information.

**8.**

**HOW THE GOLDEN RULE CAN LEAD TO  
REPRODUCTIVE SUCCESS:  
A NEW SELECTION BASIS FOR ALEXANDER'S  
"INDIRECT RECIPROCITY."**

**CHRISTOPHER BOEHM  
UNIVERSITY OF SOUTHERN CALIFORNIA**

A UNIVERSAL PREACHING

What kind of animal would think up a “Golden Rule,” and urge its fellows to practice it? Both biologically and philosophically, this is an ultimate question which goes all the way back to Darwin (1871:500). The great naturalist wrote:

To do good unto others—to do unto others as ye would that they should do unto you—is the foundation-stone of morality. It is, therefore, hardly possible to exaggerate the importance during rude times of the love of praise and the dread of blame.

Darwin was speculating about how a species that was evolved through inter-individual competition, which should make only for selfishness, could engage in behaviors that were costly to individual fitness—but helped the group. He went on to hypothesize that as units, groups might be competing through natural selection just as individuals were, and that the groups with more altruists might be winning the race.

These early thoughts of Darwin’s suggest that the golden rule serves as a beacon to guide the behavior of a species that is *morally aware*, and today we know that empirically this type of “injunction” is widespread. It amounts to a simplified, all-purpose preaching about how people should behave, and the message definitely is: “behave as you want others to behave in return,” and not, “behave toward others as they have treated you in the past,” nor, to be sure, “behave just in your own self-interest and that of your immediate family.”

At the level of phenotype, the golden rule is a call for generosity in human social interactions. And the very existence of such a rule poses a profound question for the study of the human genotype, for natural selection should not be able to support generosity that becomes “generic” as opposed to kin-oriented or socially-reciprocated.

It is noteworthy that teachings like the Golden Rule are to be found not just in Christian religion, but in all religions (Campbell 1975, 1983; see also other chapters in this volume). And such *dicta* are, of course, far from being specific to modern (or ancient) civilized cultures, which have institutionalized religions broadcasting carefully crafted messages to their flocks in the interest of promoting prosocial behavior (e.g., Wilson 2002). Indeed, generalized calls for “social harmony” are found worldwide in nonliterate tribes (Sober and Wilson 1998), while calls for individual generosity are a regular feature of small-scale cooperative societies like those of mobile hunter-gatherers (Boehm 1993, 2000).

Cooperation requires reciprocity, and in evolutionary studies both have received a great deal of attention with respect to humans and other highly social species. Underlying this interest, surely, is our own sense of the “animal” within us, to speak nontechnically. It is widely assumed that “counterhedonic” moral rules (see Campbell 1975) are needed to shape the behavior of this animal with its dangerous potential for social disruption (Boehm 1982) and modern society’s academic theoreticians have elaborated this human type of awareness in contexts of philosophy, behavioral or social science, ethology, and evolutionary biology.

As concerned communities (see Flack and de Waal 2000; de Waal 1996), ordinary people always face social problems, and they often make an effort to resolve them harmoniously, in the name of the entire community. This means that scholars from the academy are not the only social theoreticians. Indeed, as anthropologist Paul Radin (1927) told us long ago, our nonliterate peers produce a few exceptional philosophical thinkers who can make the largely-intuitive theories of their fellows beautifully explicit. Nonliterate people understand many of the functional

requirements of their societies, and these include not only the resolution of conflicts before these become too disruptive, but actively bringing about the reform—or elimination—of serious deviants such as serial killers, major cheaters, and thieves. In their quest for social harmony, they also promulgate and promote manipulative ideologies which clearly are designed to improve the fabric of social life in their groups (see Boehm 2000,).

One specific way to improve this fabric is to preach a golden-rule ideology to one's fellows as a way of reminding individuals what their cooperating groups expect of them (see Campbell 1975), and this definitely applies to nonliterate people of the type who evolved our genes for us. I mean, of course, the mobile, egalitarian hunter-gatherers that anthropologists have studied on so many of the world's continents. Compared to ourselves, these people are particularly dependent upon "preaching" to shape their everyday social life, for they refuse to tolerate strong leaders who can boss others around or despoil them.

If established with authority, powerful group leaders like Shaka Zulu can use their definitive political clout to make others behave themselves, or to decisively stop fights (see Service 1975). However, because mobile band-level societies vehemently hold down their alpha types (see Woodburn 1982), when conflicts or acts of deviance arise band members must either try to *influence* others to follow rules which support group cooperation and harmony, or in times of real social crisis they must mobilize themselves as entire concerned communities to deal with the problem. Thus, on an everyday basis foragers' main tools are education and persuasion as opposed to forceful use of authority. And reminding people to "do unto others..." is an important part of this routine process of training and persuasion.

An ultimate biocultural question is, why do people everywhere do such preaching in favor of cooperative and generous reciprocation? Such persuasion surely has been with us since well before the origins of agriculture, so it is likely to be consistent with our genetic nature—even though obviously its particular expression can vary from culture to culture. The immediate message is simply that one should not be stingy in cooperating with others. The larger lesson is that willing reciprocity is functionally important to group life. It is no surprise that group members try to “tweak” this function, for its reinforcement supports cooperative activities that benefit them all.

#### THE EVOLUTION OF COOPERATIVE RECIPROCATION

In 1971 biologist Robert Trivers (1971) published a seminal scientific article about “helping behavior” among animals in general, and this formed one cornerstone of the new interdisciplinary science of “sociobiology” (Wilson 1975). Trivers was intrigued by the fact that individually-costly helping behavior, at least when it is directed at genetic strangers, should not be supported by natural selection because of the “free-rider” problem identified by Williams (1966). The specific obstacle was that genetic altruists would be taken advantage of by genetic nonaltruists, and while altruistic genes might keep on arising through mutation, they would always fade away.

Trivers’ model offered a way out. He suggested that *over time*, if two genetic competitors were to provide each other with exactly equal amounts of assistance, meaning that these donations had perfectly equilibrated long-term costs and benefits, natural selection could keep such a behavior in place. The problem with reciprocal altruism theory is first that actual examples are very difficult to find in nature, and

second that specific mechanisms ensuring such perfect equality of sacrifice and donation over time are difficult to imagine—particularly in a behaviorally labile species that is highly competitive by nature, and, being intelligent, is prone to cheat.

Since members of a cooperating pair are likely to do better than individuals who do not cooperate, Trivers is making a case for biological “mutualism” within the same species: by helping each other equilaterally, both parties are coming out ahead. Human cooperation has received extensive evolutionary study by members of varied disciplines (e.g., Axelrod and Hamilton 1981; Boehm 1997; Bowles and Gintis 2004; Boyd and Richerson 1982, 1992; 2005; Fehr et al. 2002; Gintis et al. 2001; Gurven 2006; Kaplan and Gurven 2005; Hammerstein 2003; Knauff 1994; Nowak et al. 2004; Pepper and Smuts 2000; Ridley 1996; Smith 2003), to the point that a major academic industry has arisen. Economists and others, using college students (e.g., Fehr and Gächter 2002) and sometimes nonliterate people (Heinrich et al. 2001) as their subjects, have been able to explore this problem in experimental settings by making use of game theory, but a solid evolutionary analysis built from data on small foraging societies is needed to provide such investigations with more of an in-depth evolutionary context.

Much of this recent experimental work was inspired by the co-founder of sociobiology, Richard D. Alexander (1987). His *The Biology of Moral Systems* is probably the most enterprising work on the evolution of morals since Darwin’s (1871) *The Descent of Man*, and while Alexander’s fieldwork interests lay in crickets and beetles, his main theoretical interests included human beings as social animals, and more specifically the question of how humans could have evolved the kind of prosociality we exhibit.

Alexander saw evolving human individuals essentially as bundles of self-interest competing to see who could gain the greatest inclusive fitness, and given this paradigm it came as no surprise that people were investing in their own offspring (see also Trivers 1972) and helping their closer relatives. But there was also a mystery, for in everyday life people also help nonkin—and they do so at personal cost and without any equalizing repayment.

Hard pressed to explain the altruistic feelings and impulses which on a commonsensical basis are so apparent in our own species, Alexander felt obliged to turn to a then very controversial group-selection type of theory, just as Darwin (1871) had done before him. By 1987, evolutionary biologists knew a great deal more about how inheritance works than Darwin did, and the consensus was that unless genetically-varying groups had short life spans and therefore high extinction rates, ones comparable to individuals, group selection couldn't make much of a difference to a gene pool (see Wilson 1975). Thus, Alexander either had to posit frequent genocidal warfare as a fixture in our evolutionary past, or else he had to explain genuinely altruistic actions in some other way.

It was appropriate that Alexander (1987) turned anecdotally to hunter-gatherer studies for clues about how our moral way of life might have evolved, and while he recognized Trivers' exacting, dyadic reciprocal altruism model as important, he did not think it could be applied to what humans were doing in their hunting bands. These people obviously were not thinking on a "tit for tat" basis, for in significant ways they cooperated not in dyads or as families—but as entire local groups, which included unrelated families.

A great deal has been published about hunter-gatherers since Alexander wrote in 1987. One thing worth noting is that even though most hunters work hard at their callings and widely share the large game they kill, their contributions are quite unequal—not only at any one point in time (see Kelly 1995) but over their lifetimes, as well. Furthermore, they can move at will from band to band, so sometimes their partners in cooperation will be changing overnight, and then changing again a few months or a few years later when another move is made. Thus, in cooperative enterprises their generosity is extremely unlikely to be exactly reciprocated by those they cooperate with; indeed, Trivers' dyadic model is quite inapplicable to such a complicated and openended scenario.

Trivers' (1971) did make passing reference to what he called "general reciprocity," as a very different system of giving which does not involve precise reciprocation over time, and Alexander saw this as fitting much better with what foragers and other cooperative humans were doing. He also was influenced by the theory of economic exchange of anthropologist Marshall Sahlins (1972), who in *Stone Age Economics* recognized three types of reciprocity in nonliterate people.

Sahlins begins with "generalized reciprocity," which is normally between close kin where there is no definite expectation of return—particularly in the near future. Brothers or sisters, for instance, are prone to help their siblings with little thought of immediate payback. A traditional philosopher might call this mode of generosity "altruistic," insofar as giving without any definite compensation is involved. However, because close kin are involved, a biologist would cite this as an instance of "nepotism." In evolutionary terms, "altruism" is reserved for acts by which donors disadvantage their own fitness in order to help nonkin, whereas nepotism involves no

genetic sacrifice for donors because the donors and the recipients share their genes to a significant degree (see Trivers 1972). Help your close kin, and—*genetically speaking*—you are helping yourself.

Next Sahlins (1972) considers “balanced reciprocity.” In such bilateral exchanges, there is a strong expectation of fair equalization over the short term, as in direct barter or exchanging gifts with nonkin. Such exchanges are likely to take place among familiars in the same community; they may be genetic strangers, but they are *socially* familiar. And finally, there is “negative reciprocity,” which involves overtly self-seeking behavior as in hard bargaining—or even as in cheating to manipulate the system. The name of this third, very competitive exchange-game is simply *gain*, as opposed to making things come out even, and such exchanges are likely to take place between parties who are basically strangers both genetically and socially.

To show how all of this works, Sahlins (1972) diagrams concentrically a sedentary tribal settlement with the family group at the center, within which reciprocity is generalized. Next we move out to the entire immediate community, which engages in balanced reciprocity. Outside this small tribal community is the external world of other, similar agricultural tribal communities, and in dealing with them negative reciprocity is the mode of exchange.

Sahlins’ book was titled *Stone Age Economics*, which brings to mind human foragers. In fact his models of reciprocity are based much more on the way post-Pleistocene agricultural *tribesmen* live and socialize, as opposed to *hunter-gatherers*. Indeed, for true stone-age types the typological fit is far from perfect. For one thing, in spite of their ever-changing personnel (Palmer et al. 1997), socially and economically these mobile hunting bands are in important ways like large families—

even though biologically there are always unrelated kin units to be found in the same band (see Kelly 1995).

Of course, with respect to the acquisition and consumption of plant foods or very small game pervasive, openended sharing does take place just within families (see Kelly 1995), as indicated in Sahlins' model, and it is, in fact, highly "generalized." However, with respect to meat as a very important part of the forager diet, in effect the ownership of large carcasses is shifted from the individual hunter(s) who gained them to the band as a whole (Boehm 2004c). Thus, in this large-game context "generalized reciprocity" becomes operative for the entire local community, and a biologist would have to ask how such behavior could evolve since nonkin are involved.

Without going into detail about exactly how hunting bands share meat (e.g., Lee 1979) Alexander (1987), taking his cues from Trivers and Sahlins, coined the term "indirect reciprocity" to describe typical cooperative activities that were present in small human groups (see also Boyd and Richerson 1989). By this he meant that a person would help someone else who was in need, not necessarily a kinsman, and then expect reciprocation from some third party at some time in the future. This is readily exemplified. When foragers engage in meat sharing within a band, it is quite usual for one or a few hunters to supply much of a band's large game (Kelly 1995)—even though at least five or six active hunters are needed to keep the carcasses coming in regularly enough for *everybody* to have a stable nutrition (see Smith 1991, 2003; Winterhalder 2001; Winterhalder and Smith 1981)). In effect, this meat is taken over politically by the entire group, to be certain that those who actually

obtained a given carcass will not try to favor their own familiars—or use the meat as a commodity for their own political purposes (see Boehm 2000, 2004b).

Foragers will severely punish those who go against the system and try to dominate the meat, or who cheat seriously on the system through stealth and cunning, and on the positive side generic calls to generosity are typical of such groups (Boehm 1993, 2000). As a result their cooperative meat sharing, even though it sometimes involves an ostensibly selfish and “agonistic” *style* (see Peterson et al. 1993), is effective and reasonably equilibrated among a band’s families.

Given how much these people prize large game, we must ask how a merely-persuasive “golden rule” approach could possibly do much good, but keep in mind that every hunter, even the most prolific, is aware of having been a dependent child—in a family that was of necessity dependent on a cooperative system of meat sharing if it was to eat meat at all regularly. He also is aware of the fact that should he become ill, or suffer an accident, or be snake bit, and therefore become unproductive, his kin but also to a significant degree his band will give him support. He also knows that as aging progresses, he and his family may become increasingly dependent on other hunters who are younger—just as he as a younger hunter contributed exceptional effort. He also knows that public opinion is strongly in favor of sharing—and that generous behavior brings a good reputation which will increase the band’s inclination to cooperate with him or assist him.

In looking at this cooperative system and its benefits over time, the overall economic payoffs are obvious—even to someone who at the moment is contributing far more than his family is consuming. They are more obvious still to an average or below-average hunter. In effect, by entering into a cooperative system which is based

on *indirect* reciprocity, an individual is setting up a “safety net,” or, in our fiscal terms, is purchasing an insurance policy (Wiessner 1982, 1996; see also Smith and Boyd 1990).

In characterizing this approach to exchange as “indirect reciprocity,” Alexander (1987) is obliged to question from an evolutionary standpoint how such a system could possibly work with humans being so strongly guided by nepotistic self-interest. And because this is a question that pertains to our genetic nature, it is best answered in terms of the Late-Pleistocene people who evolved this nature for us—the same type of mobile hunter-gatherers we are discussing here.

Alexander understands that humans are a special animal, in that we are *moral* beings. As such, we invariably develop and implement behavior codes which, among other things, call for cooperation within the group—and deliberate damping of behaviors likely to cause conflict. However, cooperation need not necessarily be either altruistic in the genetic sense, or generously motivated in the psychological sense. Indeed, once a group has set its sights on cooperation in theory it can simply use punishment to make sure that individuals conform (see Boyd and Richerson 1992; Bowles et al. 2003; Gintis 2000; Wiessner 2005).

If cooperation were based solely on fear of such frighteningly-decisive group sanctions, instructions like “treat others as you would be treated yourself” would be superfluous. However, most hunter-gatherer cooperation would seem to be “voluntary,” in the sense that people appreciate its benefits, want to be good cooperators, and are amenable to following cultural cues. Golden-rule preaching helps to keep this pattern in place, precisely because humans are set up biologically to internalize cultural norms (Gintis 2004; Simon 1990).

## ALEXANDER AND SIMON'S EXPLANATIONS

Alexander (1987:83) believes that the challenge in explaining moral evolution is to account “for the altruism of moral behavior in genetically selfish terms.” Taking a cue from Darwin (1871) in this context, he reflects deeply upon the reputations of people and how they affect their fitness. Being a generous cooperator brings high prestige, and this high esteem brings tangible reproductive advantages. In this reputational sense, generosity can be its own reward—and even being generous to nonkinsmen can be rewarding because good reputations pay off in fitness.

Alexander believes that if the individual in question is viewed as an especially reliable and trustworthy reciprocator, he will be sought out by others as a partner to cooperate with. Furthermore, groups lucky enough to contain such individuals may succeed better than groups lacking them, and the group advantage will accrue to outstanding, generous reciprocators just as it accrues to any other group member. But whatever the level of selection, it is clear that humans are evolved to maximize their fitness by behaving in ways that make it clear to other group members that they are desirable partners for cooperation.

There is more, for social reputations are amenable to personal manipulation.

Alexander believes that

The consequences of indirect reciprocity... include the concomitant spread of altruism (as social investment genetically valuable to the altruist), rules, and efforts to cheat.... I would not contend that we always carry out cost-benefit analyses on these issues deliberately or consciously. I do, however, contend that such analyses occur, sometimes consciously, sometimes not, and that we are evolved to be exceedingly accurate and quick at making them (Alexander 1987: 97).

The main point is that humans are evolved to make sophisticated, self-serving decisions in the process of reputation building, and that systems of indirect

reciprocity support not only the spread of generous behavior, but the spread of both rules and cheating on rules. Thus, we are evolved not only to cleverly maximize our generous reputations, but to do so in a way that involves giving away as few reproductive resources as possible.

Of course, behaving altruistically at the level of phenotype does not mean necessarily that genuine feelings of indiscriminating generosity are involved as evolved tendencies. According to Alexander, this might be the case if prehistoric group selection was robust. But what is more certain is that evolutionary actors are evolved to engage in other-beneficial behavior just when this will do their own genotype the most good.

Alexander builds his case independently of group-selection possibilities, or the possibility that genuine altruism exists at the psychological level. It is because Late-Pleistocene hunter-gatherers were navigating systems of indirect reciprocity by maximizing their good reputations that we have evolved in the direction of highly-adept reputation building. As a result, we meter out our socially-apparent generosity so as to do ourselves the most good, and we not only make lightning calculations as to when an altruistic flourish will bring us reputational rewards that outweigh the costs, but we also are evolved to warily cheat on our group's rules (see Cosmides and Tooby 1991)—at the same time that we try to get others to follow the rules better than we ourselves do (Alexander 1987). In this context golden-rule preaching makes exceptional evolutionary sense—assuming that the objects of such preaching are in fact vulnerable to such messages.

Alexander introduces “moralizing” (preaching manipulatively to others), as an individual reproductive strategy which improves one's relative fitness within the

group (see also Campbell 1975). Keep in mind that hunting bands are not divided into those who regularly behave as cooperative altruists, those who mainly preach in favor of altruism to manipulate others, and outright selfish cheaters. People play these roles flexibly, as appropriate, and all three capacities can enhance individual fitness.

It is by being disposed to behave in these ways that a species which essentially is selfishly nepotistic can give such an impression of behaving altruistically, and Alexander explains away even saintly types like Mother Teresa as being extreme examples of what a selfishly-oriented system of indirect reciprocity and reputation-mongering can produce at the level of phenotype. He elaborates:

That I regard true altruism, in the sense of genetic self-sacrifice, as rare means that I also regard order and justice, as they now exist, as outcomes of self-interested behavior... I also regard it as crucial that we test this hypothesis thoroughly, and, if it is correct, that we take it directly and plainly into account in our efforts to preserve or restructure society (Alexander 1987:192).

In effect, Alexander is saying that it could help us to improve our social life if we would take a more realistic and necessarily rather cynical position, and concede that our “altruistic” behavior does not really fulfill the ideological definition of being selflessly generous; that is, it is not “genuine” but almost always is based on self-seeking tendencies that lie deep in human nature. Keep in mind that for Alexander there would be no basis for believing that we harbor innate altruistic tendencies—unless hunter-gatherers had been imposing frequent genocide on one another for thousands of generations in the Late Pleistocene, to drive group selection forces to much higher levels than are normally envisaged.

If navigating a judgmental moral society were nothing more than a Machiavellian game of reputation building, and if ostensibly altruistic acts were

merely tokens in that game and were being performed without any truly generous feelings, this would not have stopped bands from developing indirect-reciprocity-based systems of meat sharing. They did so in part because they were sophisticated enough to promote a golden rule mentality. However, one still must ask how we are to explain today's anonymous donations to blood banks or to starving children in Africa, or explain special altruistic acts which are to be found among hunter-gatherers when they respond to other people's emergencies regardless of kin ties being absent. Such costly behaviors are cited regularly to point out that human nature must contain some *genuinely* altruistic feelings—i.e., some innate tendencies to wider generosity. In fact, this common sense observation provides a scientific hypothesis that requires much more exploration.

#### GROUP SELECTION AND GENETIC PIGGYBACKING

In the meantime, are we to believe the scientifically well-established, mathematically modeled conclusions of evolutionary biologists like Hamilton (1964) and Trivers (1971, 1972), or our own commonsensical intuitions about how our genetic nature seems to work? In theory, group selection theory could explain acts of altruistic generosity as long as they are group-helpful, and in evolutionary biology group-selection thinking has recently made a major comeback in the form of multilevel selection theory (see Wilson and Wilson 2007).

Also, as an economist Bowles (2006) has explored the possibilities for group selection in prehistoric humans, through computer simulations. However, the dominant paradigm guiding work in experimental economics—an evolutionary field which is focused on understanding reciprocity, cheating, and cooperation—has been based essentially on *inclusive fitness* assumptions which encompass reciprocal

altruism and mutualism but not necessarily group selection. The evolutionary paradigm they rely upon is beautifully parsimonious (see Trivers 1985; Wilson 1975), and quite properly it has dominated evolutionary biology. It offers great rigor, since it is amenable to exact cost accounting assessments both in modeling and in the field. However, exclusive commitment to such a paradigm requires that humans be counted as having dispositions only to help themselves and their close kinsmen, and this doesn't fit too well with making donations to blood banks.

Nor does it fit very well with the fact that widespread, imperfectly-reciprocated social sharing took place so pervasively among culturally-modern Late Pleistocene hunting humans—even though Alexander has tried to explain this away as a basically a selfish game. Alexander does leave the group-selection door open, which offers a straightforward genetic explanation for sharing among nonkin within the same group. However, in *Sociobiology* Edward O. Wilson (1975) came to the conclusion that even though in theory group selection was possible (and likely) in real life, the conditions for it to operate more than very feebly—let alone robustly enough to fix altruistic genes in gene pools—were very unlikely to occur for any species other than social insects. Alexander (1987) accepts these limitations, but he theorizes that in the Pleistocene, when all people were living in hunting bands, genocidal warfare just *might* have driven group selection strongly enough to support some innate generosity to strangers.

In 1987, when Alexander was writing, to many evolutionists genocidal fighting and frequent band-level extinctions seemed necessary for natural selection to operate significantly at the group level (see also Soltis et al. 1995). A human generation is counted as 25 years for evolutionary purposes, and overall a Pleistocene human life

averaged at best perhaps 35-40 years. If Pleistocene bands were being totally extincted by each other every 25 years on average, then selection taking place between groups could have operated on a par with selection taking place within groups, which of course favored selfish nepotism and nothing more. Although there is no definitive *negative* evidence with respect to genocidal warfare among pre-Holocene hunter-gatherers (see Kelly 2001), there is no direct positive evidence, either; in fact, the first archaeological signs of massacres are to be found only in the Holocene (see Keeley 1996), when in its basics human nature was already evolved. Thus, Alexander knew he was speculating when he theorized about a possible role for group selection.

Subsequently, Sober and Wilson (1998) have talked about trait-group selection, which does not require outright group extinctions, and Pleistocene hunters might well have behaved in ways that made this likely (see Boehm 1997). More recently, Edward O. Wilson and David Sloan Wilson have co-published an article acknowledging that multilevel selection theory (including its group-selection components) provides a needed partial paradigm shift for evolutionary biology (see Wilson and Wilson 2007). However, there remain doubts that this type of theory will provide a full and satisfactory explanation for the biological basis of altruistic behavior (e.g., Boyd and Richerson 2005; de Waal 2006).

A different and potentially complementary theoretical approach involves genetic “piggybacking” (e.g., Simon 1990). Simon’s theory is that as cultural animals humans are extremely indoctrinable (see also Eibl-Eibesfeldt 1982) or culturally *docile*, and that in their groups they are easily able to indoctrinate innately-unaltruistic individuals to be generous to nonkin—at least within the same group.

Simon's theory is that the individual fitness benefits from being innately very culturally *docile*—that is, very good at indiscriminately learning cultural rules *in general*—are very great. They are so great that such a disposition can stay in place in a gene pool in spite of some individual losses from people internalizing and following manipulative cultural cues that influence them to behave altruistically. We are speaking of golden rules here.

Simon's model is *pleiotropic*, in the sense that biologists tell us that some genes have more than one function at the level of phenotype. With Simon's piggybacking model, the main function—cultural docility or indoctrinability—is strongly selected, while a second function—being influenced to behave altruistically—is moderately deleterious to the individual carrying the same gene. Since just one gene supports both traits, these two traits are likely to be inextricable as far as natural selection is concerned, and as long as the net fitness quotient is favorable natural selection can support a preponderantly-useful pleiotropic gene with no help, necessarily, from group selection.

As with Alexander's theory, essentially Simon's hypothesis is premised on a human genetic nature which has been shaped by inclusive fitness alone, and not by group selection. And Simon's model provides yet another, quite different way of "explaining away" altruistic behavior as being merely an effect of "something else." But we must ask: would either Alexander's or Simon's scenario really "work," in promoting wider cooperation, if there were absolutely no innate altruistic tendencies to work with in the first place? We are far from being able to determine such things by examining the human genome, but it seems safe to say that if either of these

processes had some *innate* altruism to work with, even in small amounts, its job certainly would be made much easier.

Alexander seems to have realized this, for in 1987 he was rather going out on a limb when he suggested that group selection, with its potential to straightforwardly support altruistic traits, might have been helping things along. Since then, more advanced versions of group-selection theory (e.g., Sober and Wilson 1998; Wilson and Sober 1994) and recent evolutionary analyses based on hunter-gatherers (e.g., Boehm 1999) do provide a better scenario for bringing in group selection. However, here I would like to introduce yet another theoretical possibility for explaining how human nature could have acquired a genuinely altruistic component.

#### THE EFFECTS OF POSITIVE SANCTIONING

What if the innate generosity that golden-rule *dicta* are designed to reinforce were selected neither through group selection nor pleiotropic piggybacking? Here I present in tentative form a theory which is based on what I shall call “sanctioning selection,” and this new model would appear to pertain solely or mainly to human beings, whose outsized brains permit them to make special inputs into natural selection process.

It has not gone unnoticed that the punitive social sanctioning of human groups can impact on gene pools. Trivers (1971) referred to *moralistic aggression* as a response of small human groups to rule-breakers, and he noted that the punishments meted out could have impacted on gene pools, in that punished traits would have gained a lower statistical representation. Otterbein (1988) continued this line of thinking with respect to capital punishment in nonliterate societies, and Boehm (1999) has suggested that in addition to making our human political nature milder,

many millennia of punishment through punitive group social control also could have acted as a selective force that made individuals better able to control themselves as they tried to conform to social norms. In effect, this provides an explanation for the origin of the human conscience—as a product of group social control (Boehm, in preparation).

In amplifying this interest in punishment's effects, Wrangham (2001) has suggested that capital punishment might have drastically reshaped human nature—to the degree that one might argue that humans were practicing “autodomestication” as, over time, they singled out their most aggressive deviants and killed them just as dog breeders eliminate troublesome pups from litters.

However, none of these arguments provides a new way of explaining *altruistic* traits; all they do is provide a remarkable social mechanism for changing human genetic nature in directions less likely to lead to aggressive deviance that requires punishment. Here, I would like to enlarge the scope of this newer discussion, to include not only the evolutionary effects of this negative sanctioning by groups, but also *positive* group sanctioning---as this has been defined by sociologists (e.g., Turner 2000) in terms of social rewards received either from individuals or from entire groups.

Following the majority of sociologists (e.g., Black 1984), anthropologists like Edgerton (1975) have focused preponderantly on negative or punitive social sanctioning, which includes criticism, ridicule, shaming, ostracism, ejection from the group, and capital punishment. Such behaviors are found in modern societies and in the early civilizations that preceded them (Campbell 1975), and also in the tribal societies that emerged with domestication starting 12,000 years ago and then in the

chiefdoms that followed them. They also are found, in simpler forms, in today's mobile hunter-gatherer societies (Boehm 2000), and the latter provide the best models for what our (similar) predecessors were doing as far back as 45,000 years ago and quite possibly earlier.

Recently, with prehistoric reconstruction in mind I undertook to summarize the kind of *positive sanctioning* that takes place in these same foraging bands today (see Boehm 2002, 2004a). I selected bands of the same type that prevailed in the Late Pleistocene, which may be assumed to have been smallish and composed of multiple families, some of which were unrelated. They also were mobile, in the sense that they changed locales perhaps ten times per year; they were economically independent, in that they earned their own subsistence; they were uniformly egalitarian, in that alpha males were not permitted; and they prized and emphasized meat as a food no matter what proportion of their diet it accounted for (see also Boehm 2002).

These early human bands had the same varied array of punitive negative sanctions as listed above, which were applied against social transgressors who engaged in antisocial patterns like serial killing, flagrant cheating, thieving, and other types of serious deception (Boehm 2000). In addition, they gave approval to prosocial behaviors such as being generous, being politically unassuming, and being reliable as a partner in cooperation (Boehm 2004a), and such approval, as Alexander (1987) said, brought rewards in the form of being chosen as a group leader, or of being viewed as a desirable partner in marriage or cooperation. While Alexander was correct in saying that such reproductive rewards could compensate for fitness losses suffered from behaving generously, this social type of selection process also had

another effect and a major one. *It exerted a direct selection pressure in favor of those innately altruistic traits which were favored by band members.*

Just as negative sanctioning selection shaped our genetic nature to make it less aggressive—and more prone to exhibit self-control in the form of a conscience—positive sanctioning selection supported genes making for altruistic generosity and cooperation. It was able to do so because when mutations that predisposed certain genuinely altruistic behaviors arose, the individuals carrying them gained in relative fitness. This does not rule out Alexander's theory of dissembling, which in my view involves opportunistically "strategizing" one's generous acts. But it tells us where the genetic tendencies contributing to such generous acts came from in the first place. It also tells us that specific types of genetic altruism, in the form of group-approved acts of benevolence which extend beyond kinsmen, can have a straightforward basis in natural selection—but this involves a very special type of natural selection which is shaped not by passively-acting physical or social environments, but by symbolically-communicating groups that deliberately try to implement not only their everyday rules, but also their larger social ideals.

In many instances, these positive and negative sanctioning actions would have had complementary selection effects. For instance, at the same time that undue stinginess or cheating in meat distributions would have invoked group wrath with serious fitness consequences for malefactors, individuals who were exceptionally generous not only would have escaped the punishments inflicted on the stingy, but would have accrued reputational fitness benefits as described above. And both effects would have facilitated smoother cooperation for groups even as they shaped our genetic dispositions.

Alexander surely is right that a substantial amount of conspicuously-generous behavior involves significant opportunistic manipulation in terms of self-presentation, and Simon surely is correct in suggesting that socialization and values-internalization can powerfully guide the phenotypic expression of altruistic behavior. The assumption of both authors is that their theories would work even if there were no altruistic tendencies in our nature. However, positive sanctioning-selection theory provides some actual altruistic tendencies in our genetic makeup for these two processes to work with, and I believe that this enhances both of them as hypotheses. In such a context, a small amount of bona fide altruism might go a long ways.

#### PREHISTORIC POSITIVE SANTIONING?

This novel theory of altruism needs further exploration, obviously, and we shall briefly do so here. We have seen that when natural selection operates in its most basic, Darwinian mode, inclusive fitness should be all that counts—unless some rather weak interdemic or group effects are present. Simon’s (1990) pleiotropic explanation of altruistic behavior fits with this widely-accepted scenario, and basically so does Alexander (1987).

However, when a band of like-minded people comes up with a set of moral values and rules and stays true to them for thousands of generations, and when it spontaneously doles out rewards to exceptional individuals who behave with unusual generosity, this group behavior becomes, in effect, a vicarious selection agency. Because such behavior is guided in a stable way by human preferences, it can have its own separate effect on gene pools (Boehm 1978, 1991, 1996; see also Campbell 19975, 1983).

As long as this special, culturally-based human agency of gene selection remains consistent in its effects, and as long as it has sufficient generations to work with, it can keep on influencing the human genome in ways that are actively guided by group preferences. I do not mean to say that something like “genetic engineering” is taking place, for the people involved were merely trying to improve their immediate level of cooperation; they were not consciously trying to tamper with their gene pool, even though this became an unintended consequence.

What it amounts to, is that our gene pools have been shaped, unwittingly, by very strong human preferences for generous behavior. And those preferences go far beyond nepotism, even though human groups do always approve of nurturing children and helping close relatives. They go beyond nepotism because people naturally cooperate in families but generous cooperation with nonkinsmen is more problematic. In encouraging such cooperation, human bands would have created a well-focused selection pressure in favor of altruistic generosity.

The Late Pleistocene “egalitarian syndrome” I described above was centered on large-game hunting, and developing this activity would have led to a perceptually obvious political need to somehow get rid of or pacify the alpha males who predictably would disrupt equalized meat distributions (Boehm 1999, 2004b). It is very difficult to imagine a large-game subsistence base working efficiently otherwise, particularly in the face of the dangerously unpredictable Pleistocene climatic conditions that prevailed in Africa and Eurasian environments while our genes were evolving. And it is equally difficult to imagine an equitable system of meat distribution’s having existed in what was a hierarchically-oriented species, without

people banding together to assertively suppress the alpha-male system that naturally tended to impose itself.

This assertion that large-game hunting communities were *obliged* to cope with their own alpha-male systems must be further explained. Winterhalder (2001) makes a powerful case for the large-game hunting kills of individuals needing to be shared by an entire band, rather than by just the family of the hunter. For meat to work as a main subsistence item, it must be eaten reasonably regularly, and if a single families were to consume large carcasses on their own, they would be feasting once in a very long while and otherwise would be enduring very protracted meat-famines. However, if at least five families join into a band, having numerous active hunters provides a dramatic “averaging” effect—if and only if the meat is always shared out equally. Thus, in politically-egalitarian bands people could eat adequate amounts of meat quite regularly.

Today, all mobile hunter-gatherers of the Late-Pleistocene type share their meat, and in order to do so they must hold down selfishly-assertive tendencies in their top males, who otherwise would seriously disrupt equalized meat-sharing. The same was true in the Late Pleistocene, for similar dominance tendencies have been with us for millions of years (see Wrangham and Peterson 1996). And if one thinks about preadaptations, ancestrally there were already political coalitions that opposed the domination of top males and effectively reduced their power, so the potential for behavior of this type was already well developed when humans first began to rely on large game at least a quarter of a million years ago (see Boehm 1999).

We may assume that the original egalitarian coalitions which arose in this context were based solidly on self-interest and the use of collective force, and that the

need to share large game was a very likely catalyst in making groups reject alpha behavior so decisively. However, in these earliest large-game hunting communities the counterdominant coalitions would have been nonmoral, for a conscience could develop only *after* such negative sanctioning selection arose. This probably would have taken at least 1,000 to 2,000 generations.

Thus, if hunting was in full force by 250,000 years ago, in theory some version of an early human brain equipped with a conscience could have been evolving by 200,000 to 225,000 years ago, and possibly before humans became fully modern anatomically. Once a conscience arose sometime in the Late Pleistocene, moralistic gossiping and the reckoning of personal reputations would have begun to inform group life, and *positive* sanctioning selection would have gone into operation. It was then that robust selection in favor of generous behavior would have begun, and the agency of selection was humans with their biases in favor of generosity.

It is safe to hypothesize that by the time humans had become both anatomically modern *and* culturally modern, basically their social practices would have become consistent with what is practiced today by people who similarly live in mobile bands, keep down their alpha males, treasure meat as a dietary item, divide their sporadic large-game kills more or less equally within the band—and favor generous cooperators both when they gossip moralistically and when they make social choices.

People became *anatomically* modern over 100,000 years ago, and they became *culturally* modern, as indicated by their lifestyles and artifacts, at latest by 45,000 years ago (see Klein 1999). Thus, we may confidently project an egalitarian syndrome back to that time, and assert that something like a modern type of positive

sanctioning selection was at work by then. This applies to both the Eurasian Upper Paleolithic and a bit earlier the African New Stone Age, and this date is conservative because eventually, with further archaeological evidence, the advent of cultural modernity may well be pushed back some (see McBrearty and Brooks 2000).

This means that as a predictable component of the egalitarian syndrome, positive social sanctioning selection has had at least 1800 generations to work with, in acting as a gene-selection agency which favors precisely the types of generosity that are encouraged and rewarded by hunter-gatherers. Wilson (1978) suggests that it takes about 1000 generations for a major change in a genome to take place, but he is speaking of ordinary natural selection operating at a putatively steady rate. Because positive sanctioning selection is actually *guided*, by very specific human social intentions, it therefore is very strongly “focused.”

Natural environments also may change in ways that “focus,” natural selection on changing or supporting certain traits, but it bears emphasis that when a human group begins to reward fairness and generosity exhibited by individuals who are dividing up large-game carcasses, and generosity of those who are better situated toward those in unfortunate circumstances, the selection rate could be rather high because of the consistency, and because the rewards are reproductively impactful. I should point out that negative sanctioning selection, with its reliance on capital punishment, is likely to operate at a much higher rate even than positive sanctioning selection. Possibly, computer simulations could be devised to test these hypotheses.

If this positive sanctioning-selection scenario has merit, it provides yet another avenue to explain not only why humans practice as much cooperation as we do, but also why we are able to harbor genuinely altruistic feelings toward genetic strangers

outside of our bonded communities—and even beyond the boundaries of our nations or our species. In a highly labile species, these can be seen as “generalized” behavioral expressions of a potential for genetic altruism which originally evolved in the bosom of meat-sharing bands. As we have seen, this potential evolved genetically because band members rewarded individuals who were unusually generous, and thereby unwittingly manipulated their own gene pools. Group selection certainly could have been producing a similar effect, but there is no reason that sanctioning selection and group selection couldn’t have operated in tandem. In fact, this additional theory takes some of the burden off of a potentially over-worked group selection theory, in explaining our obvious (and to many still mysterious) impulses to altruistic generosity.

Alexander and Simon have afforded us profound insights about the basis for moral life in human groups, and basically they have tried to do so inventively, but without deviating from the dominant evolutionary paradigm. In terms of our genetic nature their theories see selfish individuals performing altruistically-benificent acts for others “in spite of themselves.” However, if sanctioning selection (and quite possibly group selection) have been providing us with some limited representation of generous traits in our gene pools, and if such generosity is not restricted to close kin, then these two evolutionary hypotheses about the evolution of altruistic cooperation would become far easier to envisage.

If there has been significant selection for innate altruism, when Alexander’s “Machiavellian” reputation-mongers engage in altruistically-generous deeds because they know this will help them and their families, they are doing more than merely dissembling as a sociopath might. They are putting to use their own limited but

genuine tendencies to be generous, in a strategized way that will maximally enhance their reputations and therefore their fitness. And when Simon's genetically-docile "cultural apprentices" unconsciously absorb the golden-rule messages put out by their parents and their groups, this internalization is made much easier because there are genuinely generous altruistic impulses for this socialization process to work with.

If my hypothesis is accepted, we now have both sanctioning selection and group selection available as mechanisms that can explain the evolution of altruistic traits in human gene pools, and there is no reason to suggest that these processes could not work together. We also have Simon's piggybacking theory and Alexander's theory of indirect reciprocity to assist in this explanation, and likewise these two models are in no way in conflict with each other. Add up the effects of all four of these processes, and there should be a better basis for explaining the intensive cooperation that has prevailed on a moralistic basis in human bands over the past 45,000 years, and probably more.

## DISCUSSION

It is somewhat ethnocentric to speak of *the* Golden Rule, insofar as such rules seem to be humanly universal, be they explicit or implicit. One anthropological conclusion would be that golden rules are simply socio-ideological devices, used opportunistically by manipulative groups or individuals to spur group members in the direction of generosity and nonaggression, and to do so in spite of very strong evolved propensities that favor selfish nepotism.

But what about the earnest "biological cynicism" about our genetic nature which is inherent in Alexander's and Simon's discussions of moral behavior, as well as in treatments like Wilson (1978) and Trivers (1985)? Should we view the golden

rule as a manipulative means of training or pressuring selfish nepotists to behave generously in spite of themselves, and nothing more? It is my view that people without a trace of altruistic feeling in their behavioral repertoire would be unlikely to come up with such a rule in the first place, and that, furthermore, such a rule would be unlikely to work very well on such entities if they were totally selfish genetically.

The alternative I propose is that there may well be a sound social basis for the selection of genetically altruistic traits—at least in moderation—and that this comes in the form of the positive sanctioning selection I have hypothesized, and surely also through some group selection. In thinking about these theoretical developments, it is important to keep in mind that our genetic nature disposes us frequently to states of social ambivalence (Boehm 1999), as when one is torn between a modest but genuine generous impulse, one which is reinforced by the possibility of opportunistically enhancing one's reputation, and a competing desire to behave in a way that is entirely selfish. It was in helping people to resolve such ambivalences in a prosocial direction, that human groups long ago began to come up with their golden-rule exhortations. And because large-scale societies have continued to value cooperation, they have continued to use golden rules to reinforce the prosocial side of our genetic nature.

Thus, golden-rule preaching is deeply embedded in our evolutionary past. It has been such injunctions that have encouraged otherwise-ambivalent individuals to go with the generous side of their nature, insubstantial as it may prove to be, and to reap the reputational rewards that have shaped our gene pools and made us at least moderately altruistic in our natural inclinations.

## REFERENCES CITED

Alexander, Richard D.

1987 *The Biology of Moral Systems*. New York: Aldine de Gruyter.

Axelrod, Robert and William D. Hamilton

1981 *The Evolution of Cooperation*. *Science* 211:1390-1396.

Black, Donald

1984 *Toward a General Theory of Social Control*. Orlando: Academic Press.

Boehm, Christopher

1978 *Rational Pre-Selection from Hamadryas to Homo sapiens: The Place of Decisions in Adaptive Process*. *American Anthropologist* 80:265-296.

1982 *The Evolutionary Development of Morality as an Effect of Dominance Behavior and Conflict Interference*. *Journal of Social and Biological Sciences* 5:413-422.

1991 *Lower-Level Teleology in Biological Evolution: Decision Behavior and Reproductive Success in Two Species*. *Cultural Dynamics* 4:115-134.

1993 *Egalitarian Society and Reverse Dominance Hierarchy*. *Current Anthropology* 34:227-254.

1996 *Emergency Decisions, Cultural Selection Mechanics, and Group Selection*. *Current Anthropology* 37:763-793.

1997 *Impact of the Human Egalitarian Syndrome on Darwinian Selection Mechanics*. *American Naturalist* 150:100-121.

1999 *Hierarchy in the Forest: The Evolution of Egalitarian Behavior*. Cambridge: Harvard University Press.

2000 *Conflict and the Evolution of Social Control*. *Journal of Consciousness Studies*, Special Issue on Evolutionary Origins of Morality (L. Katz guest editor) 7:79-183.

2002 *Variance Reduction and the Evolution of Social Control*. Paper Presented at Santa Fe Institute, 5th Annual Workshop on the Co-Evolution of Behaviors and Institutions, Santa Fe. (Posted at Santa Fe Institute Web Site).

2004a *Explaining the Prosocial Side of Moral Communities*. In, *Evolution and Ethics: Human Morality in Biological and Religious Perspective*. P. Clayton and J. Schloss, eds., pp. 78-100. New York: Eerdmans.

2004b *Large-Game Hunting and the Evolution of Human Sociality*. In, *Origins and Nature of Sociality among Nonhuman and Human Primates*. R. W. Sussman and A. R. Chapman, eds., pp. 270-287. New York: Aldine.

2004c What Makes Humans Economically Distinctive? A Three-Species Evolutionary Comparison and Historical Analysis. *Journal of Bioeconomics* 2:109-135.

Bowles, Samuel

2006 Group Competition, Reproductive Leveling, and the Evolution of Human Altruism. *Science* 8:1569-1572.

Bowles, Samuel, and Herbert Gintis

2004 The Evolution of Strong Reciprocity: Cooperation in Heterogeneous Populations. *Theoretical Population Biology* 65:17-28.

Bowles, Samuel, Ernst Fehr, and Herbert Gintis

2003 Strong Reciprocity and the Conditions for the Evolution of Altruism. [www.Umass.Edu/Preferen/Gintis/Theoretical%20primatology.doc](http://www.Umass.Edu/Preferen/Gintis/Theoretical%20primatology.doc)

Boyd, Robert, and Peter J. Richerson

1982 Cultural Transmission and the Evolution of Cooperative Behavior. *Human Ecology* 10:325-351.

1989 The Evolution of Indirect Reciprocity. *Social Networks* 11:213-236.

1992 Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups. *Ethology and Sociobiology* 13:171-195.

2005 Solving the Puzzle of Human Cooperation. In, *Evolution and Culture*. S. Levinson, ed., pp. 105-132. Cambridge: MIT Press.

Campbell, Donald T.

1975 On the Conflicts Between Biological and Social Evolution and Between Psychology and Moral Tradition. *American Psychologist* 30:1103-1126.

1983 The Two Distinct Routes beyond Kin Selection to Ultrasociality: Implications for the Humanities and Social Sciences. In, *The Nature of Prosocial Development: Interdisciplinary Theories and Strategies*. D. Bridgeman, ed., pp. 11-41. New York: Academic Press.

Cosmides, Leda, and John Tooby

1992 Cognitive adaptations for social exchange. In, *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. J. H. Barkow, L. Cosmides, & J. Tooby, eds., pp. 163-228. New York, NY: Oxford University Press.

Darwin, Charles

1871 *The Descent of Man and Selection in Relation to Sex*. London: John Murray.

Edgerton, Robert B.

1975 *Deviance: A Cross-Cultural Perspective*. Menlo Park: Cummings.

Eibl-Eibesfeldt, Irenäus

1982 Warfare, Man's Indoctrinability and Group Selection. *Zeitschrift für Tierpsychology* 60:177-198.

Erdal, David, and Andrew Whiten

1994 On Human Egalitarianism: An Evolutionary Product of Machiavellian Status Escalation? *Current Anthropology* 35:175-184.

Fehr, Ernst, Urs Fischbacher, and Simon Gächter

2002 Strong Reciprocity, Human Cooperation and the Enforcement of Social Norms. *Human Nature* 13:1-25.

Fehr, Ernst and Simon Gächter

2002 Altruistic Punishment in Humans. *Nature* 415:137-140.

Gintis, Herbert

2000 Strong Reciprocity and Human Sociality. *Journal of Theoretical Biology* 206:169-179.

2004 The Genetic Side of Gene-culture Coevolution: Internalization of Norms and Prosocial Emotions. *Journal of Economic Behavior and Organization* 53:57-67.

Gintis, Herbert, Eric Smith, and Samuel Bowles

2001 Costly Signaling and Cooperation. *Journal of Theoretical Biology* 213:103-109.

Gurven, Michael

2006 The Evolution of Contingent Cooperation. *Current Anthropology* 47:185-192.

Hamilton, William D.

1964 The Genetical Evolution of Social Behavior I, II. *Journal of Theoretical Biology* 7:1-52.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath

2001 Cooperation, Reciprocity and Punishment in Fifteen Small-Scale Societies. *American Economics Review* 91:73-7.

Kaplan, Hillard and Gurven, Michael

2005 The Natural History of Human Food Sharing and Cooperation: A Review and a New Multi-Individual Approach to the Negotiation of Norms. In, *Moral Sentiments and Material Interests: On the Foundations of Cooperation in Economic Life*. H. Gintis, S. Bowles, R. Boyd, and E. Fehr, eds., pp. 75-113. Cambridge: MIT Press.

Keeley, Lawrence H.

1996 *War Before Civilization*. Oxford: Oxford University Press.

Kelly, Robert L.

1995 *The Foraging Spectrum: Diversity in Hunter-Gatherer Lifeways*. Washington: Smithsonian.

Kelly, Raymond C.

2000 *Warless Societies and the Evolution of War*. Ann Arbor: University of Michigan Press.

Klein, Richard G.

1999 *The Human Career: Human Biological and Cultural Origins*. Chicago: University of Chicago Press.

Knauff, Bruce M.

1994 Culture and Cooperation in Human Evolution. In, *The Anthropology of Peace and Nonviolence*. L. Sponsel and T. Gregor, eds., pp. 37-67. Boulder: Lynne Rienner.

Lee, Richard B.

1979 *The !Kung San: Men, Women and Work in a Foraging Society*. Cambridge: Cambridge University Press.

McBrearty, Sally, and Alison Brooks

2000 The Revolution That Wasn't: A New Interpretation of the Origin of Modern Human Behavior. *Journal of Human Evolution* 39:453-563.

Nowak, Martin A., Akira Sasaki, Christine Taylor, and Drew Fudenberg

2004 Emergence of Cooperation and Evolutionary Stability in Finite Populations. *Nature* 428:646-650.

Otterbein, Keith

1988 Capital Punishment: A Selection Mechanism. Comment on Robert K. Dentan, On Semai Homicide. *Current Anthropology* 29:633-636.

Palmer, Craig T., B. Eric Fredrickson, and Christopher F. Tilley

1997 Categories and Gatherings: Group Selection and the Mythology of Cultural Anthropology. *Evolution and Human Behavior* 18:291-308.

Pepper, John W., and Barbara Smuts

2000 The Evolution of Cooperation in an Ecological Context: An Agent-Based Model. In, *Dynamics in Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes*. T. A. Kohler and G. J. Gumerman, eds., pp. 45-76. New York: Oxford University Press.

Peterson, Nicolas

1993 Demand Sharing: Reciprocity and the Pressure for Generosity among Foragers. *American Anthropologist* 95:860-874.

Radin, Paul

1927 *Primitive Man as Philosopher*. New York: Appleton.

Ridley, Matt

1996 *The Origins of Virtue: Human Instincts and the Evolution of Cooperation*. New York: Penguin Books.

Sahlins, Marshall D.

1972 *Stone Age Economics*. New York: Aldine de Gruyter.

Simon, Herbert

1990 A Mechanism for Social Selection and Successful Altruism. *Science* 250:1665-1668.

Smith, Eric Alden

1991 *Inujjamiut Foraging Strategies: Evolutionary Ecology of an Arctic Hunting Community*. New York: Aldine de Gruyter.

2003 Human Cooperation: Perspectives from Behavioral Ecology. In, *Genetic and Cultural Evolution of Cooperation*. P. Hammerstein, ed., pp. 401-427. Cambridge: MIT Press.

Smith, Eric Alden and Robert Boyd

1990 Risk and Reciprocity: Hunter-Gatherer Socioecology and the Problem of Collective Action. In, *Risk and Uncertainty in Tribal and Peasant Economies*. E. A. Cashdan, ed., pp. 167-192. Boulder: Westview Press.

Sober, Elliott and David S. Wilson

1998 *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge: Harvard University Press.

Soltis, Joseph, Robert Boyd, and Peter J. Richerson

1995 Can Group-Functional Behaviors Evolve by Cultural Group Selection? An Empirical Test. *Current Anthropology* 36:473-494.

Trivers, Robert L.

1971 The Evolution of Reciprocal Altruism. *Quarterly Review of Biology* 46:35-57.

1972 Parental Investment and Sexual Selection. In, *Sexual Selection and the Descent of Man, 1871-1971*. B. G. Campbell, ed., pp. 136-179. Chicago: Aldine.

1985 *Social Evolution*. Menlo Park: Benjamin/Cummings.

Turner, Jonathan H.

2000 *On the Origins of Human Emotions: A Sociological Inquiry into the Evolution of Human Affect*. Stanford: Stanford University Press.

de Waal, Frans B. M.

1996 *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Cambridge: Harvard University Press.

2006 *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press.

Wiessner, Polly

1982 Risk, Reciprocity and Social Influences On !Kung San Economics. In, *Politics and History in Band Societies*. E. Leacock and R. B. Lee, eds., pp. 61-84. Cambridge: Cambridge University Press.

1996 Leveling the Hunter. In, *Food and the Status Quest: An Interdisciplinary Perspective*. P. Wiessner and W. Schiefenhövel, eds., pp. 171-191. Oxford: Berghahn Books.

2005 Norm Enforcement among the Ju/'Hoansi Bushmen: A Case of Strong Reciprocity? *Human Nature* 16:115-145.

Williams, George C.

1966 *Adaptation and Natural Selection: A Critique of some Current Evolutionary Thought*. Princeton: Princeton University Press.

Wilson, David Sloan

2002 *Darwin's Cathedral: Evolution, Religion, and the Nature of Society*. Chicago: University of Chicago Press.

Wilson, David S., and Elliott Sober

1994 Reintroducing Group Selection to the Human Behavioral Sciences. *Behavior and Brain Sciences* 17:585-654.

Wilson, David S. and Edward O. Wilson

2007 Rethinking the Theoretical Foundation of Sociobiology. *Quarterly Review of Biology* 82:327-348.

Wilson, Edward O.

1975 *Sociobiology: The New Synthesis*. Cambridge: Harvard University Press.

1978 *On Human Nature*. Cambridge: Harvard University Press.

Winterhalder, Bruce

2001 Intragroup Resource Transfers: Comparative Evidence, Models, and Implications for Human Evolution. In, *Meat-Eating and Human Evolution*. C. B. Stanford and H. T. Bunn, eds., pp. 279-301. New York: Oxford University Press.

Winterhalder, Bruce, and Eric A. Smith, eds.

1981 *Hunter-Gatherer Foraging Strategies: Ethnographic and Archeological Analyses*. Chicago: University of Chicago Press.

Woodburn, James

1982 Egalitarian Societies. *Man* 17:431-451.

Wrangham, Richard

2001 Interview on John Brockman's website, Edge. <http://www.edge.org/3rdculture/Wrangham>.

# GOLD OR FOOL'S GOLD? RIDDING THE GOLDEN RULE OF ABSURD IMPLICATIONS

HARRY J. GENSLER, S.J.  
JOHN CARROLL UNIVERSITY

## I

**What does the Golden Rule say?**<sup>1</sup> It says “Treat others as you want to be treated.” This simple formula often works well, as in “I want others to be polite to me, so I’m polite to them” or “I don’t want others to rob me, so I don’t rob them.” But it can give questionable results when the parties are in *different situations* or have *flawed desires*:

DIFFERENT SITUATIONS: If you want Dr. Davis to remove your appendix, then remove her appendix.

FLAWED DESIRES: If you (in a fit of self-hating depression) want everyone to hurt you, then hurt everyone yourself.

“Treat others as you want to be treated” hides a sea of ambiguities and problems. To see these more clearly, we’ll consider these two formulations of the Golden Rule:

Pyrite-1. If you want X to do A to you, then do A to X.

Pyrite-2. Supposing that if the situation were reversed then you’d want X to do A to you, then do A to X.

Both are called “pyrite,” since they glitter but are ultimately only fool’s gold.<sup>2</sup>

---

<sup>1</sup> The four parts of my paper follow the four questions in William Scott Green’s keynote essay, “Parsing Reciprocity: Questions for the Golden Rule,” at the beginning of this book. My answers are based on and sum up previous works of mine about the Golden Rule, especially my textbook *Ethics: A Contemporary Introduction* (London: Routledge, 1998), chaps. 7–9; my more technical *Formal Ethics* (London: Routledge, 1996); and my logical formalization in *Introduction to Logic* (London: Routledge, 2002), chap. 11. Related still earlier works include a 1977 Michigan doctoral dissertation “The Golden Rule” (available through University Microfilms, Ann Arbor, Mich.); “Ethical Consistency Principles,” *Philosophical Quarterly* 35 (1985): 157–70; and “Ethics is Based on Rationality,” *Journal of Value Inquiry* 20 (1986): 251–64. My <http://www.jcu.edu/philosophy/gensler/goldrule.htm> page on the Golden Rule has useful links.

<sup>2</sup> While both pyrites are formulated positively, similar problems face negative forms like “If you want X not to do A to you, then *don’t* do A to X.” Both forms are logically equivalent if for “do A” we can substitute “refrain from doing such and such.” The positive/negative distinction doesn’t coincide with the more important doing-good/not-harming distinction. Matthew 5:42 says “Give to one who asks, and don’t turn your back on one who wants to borrow”; this is about doing good, whether we use the positive “Give” or

Pyrite-1 can lead to absurdities when you and X are in different situations, as in the appendix case; here are further examples:

- To a person with a somewhat deaf uncle Al: If you don't want Al to speak loudly to you, then don't speak loudly to him.
- To a masochist with a friend Bob who hates pain: If you want Bob to cause you pain, then cause him pain.
- To a broccoli-hating waitress whose broccoli-loving customer Carol just ordered broccoli: If you don't want Carol to serve you broccoli, then don't serve her broccoli.<sup>3</sup>

Pyrite-1 ignores differences in circumstances. It tells us: "If you want others to treat you in a given way in your present situation, then this is how you are to treat them – even if their situation is very different."

Pyrite-1 also can lead to problems when you have flawed desires about how you are to be treated, as in the self-hatred case; here are further examples:

- If you want friends to rob banks with you, then rob banks with them.<sup>4</sup>
- If you want others to encourage you to smoke (which you mistakenly think is a healthy practice), then encourage them to smoke.
- If you desire (as your Nazi society taught you to desire) that you'd be tortured if you were found out to be Jewish, then torture Jews.

Pyrite-1 assumes that our desires about how we are to be treated are fine and give us a flawless guide on how to treat others. But our desires may be seriously flawed.

Pyrite-2 overcomes the *different situations* problem. Pyrite-2 has us ask, not how we want to be treated now, but rather how, if we were in the place of the other person in the reversed situation, we'd then want to be treated:

- To a person with a somewhat deaf uncle Al: If the situation were reversed (and you were hard of hearing), then you'd want Al to speak loudly to you. So by Pyrite-2 you ought to speak loudly to him.
- To a masochist with a friend Bob who hates pain: If the situation were reversed (and you hated pain), then you wouldn't want Bob to cause you pain. So by Pyrite-2 you ought not to cause him pain.
- To a broccoli-hating waitress whose broccoli-loving customer Carol just ordered broccoli: If the situation were reversed (and you loved broccoli and ordered it), then you'd want Carol to serve you broccoli.

---

the negative "Don't turn your back." See my *Formal Ethics* 108–111.

<sup>3</sup> George Bernard Shaw had something like this objection in mind when he wrote "Do not do unto others as you would that they should do unto you. Their tastes may not be the same." See his *Man and Superman* (New York: Wm. H. Wise & Company, 1903), 217.

<sup>4</sup> St. Augustine's objection was roughly "If you want Al to get drunk with you, then get drunk with him." See his *The Lord's Sermon on the Mount*, trans. J.J. Jepson (Westminster, Md.: Newman Press, 1948), 161–62. This objection is the earliest (about 400 AD) mention of Golden Rule absurdities that I know of.

So by Pyrite-2 you ought to serve her broccoli.

So Pyrite-2 is an improvement over Pyrite-1.

Unfortunately, Pyrite-2 has problems when other person has flawed desires:

- Robert wants you to help him to rob banks. If the situation were reversed (and you were in his place), then you'd want Robert to help you to rob banks. So by Pyrite-2 you ought to help him to rob banks.
- Depressed Dora wants everyone to hurt her. If the situation were reversed (and you were in her place), then you'd want Dora and everyone else to hurt you. So by Pyrite-2 you ought to hurt her.

And we have Immanuel Kant's famous case about the judge and criminal:<sup>5</sup>

- You're a judge about to jail Charlie, a dangerous criminal who doesn't want to be jailed. If the situation were reversed, then you wouldn't want Charlie to jail you. So by Pyrite-2 you ought not to jail him.

Pyrite-2 is just a complicated way to say we shouldn't act against someone's desires:

Suppose X wants you to do A to him. If the situation were reversed, you'd want X to do A to you. So by Pyrite-2 you ought to do A to X.<sup>6</sup>

But sometimes we *must* act against someone's desires. Imagine that you must hire either Alice or Betty (but not both) and each wants to be the exclusive hire. Pyrite-2 here gives conflicting directives, depending on which party you switch places with:

- If your situation were reversed with Alice's, you'd want Alice to hire you and not Betty. So you ought to hire Alice and not Betty.
- If your situation were reversed with Betty's, you'd want Betty to hire you and not Alice. So you ought to hire Betty and not Alice.

---

<sup>5</sup> Kant's 1785 footnote about this occurs in his *Groundwork of the Metaphysics of Morals*, trans. H.J. Paton (New York: Harper & Row, 1964), 430. Jeffrey Wattles noted that Kant's footnote "almost completely silenced a previously energetic European tradition of reflection on the Golden Rule for 150 years"; see Wattles's "Levels of Meaning in the Golden Rule," *Journal of Religious Ethics* 15 (1987): 106–29. Wattles later speculated that Kant's formulation of the categorical imperative may have been prompted in part by the need to reform the Golden Rule; see Wattles's *The Golden Rule* (New York: Oxford University Press, 1996), 214. For a good discussion of Kant's footnote, see H.U. Hoche's "The Golden Rule: New Aspects of an Old Principle," *Contemporary German Philosophy*, ed. D.E. Christensen (University Park: Pennsylvania State University Press, 1982), 69–90.

<sup>6</sup> We could avoid this result by adding a rationalizing condition to Pyrite-2, perhaps like "Suppose that if the situation were reversed *except that you are imagined to be rational* then you'd want X to do A to you; then do A to X." The problem with this, however, is that people often need to be treated in certain ways just because they are irrational. For example, I may need to confront X about being irrational. But if I were in X's place but rational then I wouldn't want to be confronted about being irrational; so then, by the revised rule, I shouldn't confront X.

Despite a promising start, Pyrite-2 also falls to fatal objections.

What should we conclude about the Golden Rule? Most moral philosophers dismiss it as a folk principle that falls apart once we express it clearly. Many say the spirit behind the rule is right but better expressed in completely different terms (in perhaps “Do good to others” or utilitarianism or Kantianism). So the Golden Rule isn’t a major force in mainstream moral philosophy<sup>7</sup> and many ethics courses don’t even mention it. Unless the Golden Rule is defended better, dismissing it seems reasonable.

Miners say that pyrite, while “fool’s gold,” is often a sign that genuine gold is nearby; we just have to dig more carefully. And that is my attitude about the Golden Rule. With more searching, we can find a better formulation that resists absurd implications.

## II

**What does the Golden Rule mean?** I suggest this as a clear formulation that is useful in moral thinking and avoids the absurd implications (the box on the right is more precise):

<p>Gold-1:</p> <p>Treat others only as you consent to being treated in the same situation.</p>	<p>Gold-1 forbids this combination:</p> <ul style="list-style-type: none"><li>• I do something to another.</li><li>• I’m unwilling that this be done to me in the same situation.</li></ul>
--	---

I see the Golden Rule as a consistency principle; it tells us not to have our actions (toward another) be out of harmony with our desires (about a reversed-situation action). The Golden Rule is part of a larger family of consistency principles; if we see how such principles work, we can see how to formulate the Golden Rule in a better way. So my approach isn’t “Fiddle with the wording of the Golden Rule until you avoid the absurd implications.” Instead, my approach is “Understand how consistency principles work and then apply these lessons to the Golden Rule.”<sup>8</sup>

The family of consistency principles, to which the Golden Rule belongs, prescribes things like logical consistency among our beliefs, consistency between our moral beliefs and how we live, consistency about our evaluations of similar actions, and consistency between our ends and means. What distinguishes such consistency principles is that they don’t directly tell us what to do or believe; instead, they tell us to avoid inconsistent combinations. These principles can be put into an axiomatic form, whereby many consis-

---

<sup>7</sup> The exception is R.M. Hare’s *Freedom and Reason* (Oxford: Clarendon Press, 1963), which interprets the Golden Rule as a rule for the consistent use of “ought”: if we are consistent in our use of “ought,” then we won’t think that we *ought* to do A to X unless we prescribe (desire) that if we were in X’s place then A would be done to us.” Hare was a major influence on my thinking about the Golden Rule, even though I disagree with his exact approach; see my “The Prescriptivism Incompleteness Theorem,” *Mind* 85 (1976): 589–96; *Ethics* 72–83; and *Formal Ethics* 134–43. My interest in the Golden Rule started when I heard a talk by Hare in Detroit in 1968.

<sup>8</sup> For the meaning of “consent” or “willing” in my principles, see *Formal Ethics* 43–44.

tency theorems (including Gold-1) can be derived from a few consistency axioms.<sup>9</sup> But we don't have time for this here. So we will instead proceed more intuitively, and more by analogy than by deduction. We will focus on what we can learn from three other consistency principles about how to formulate the Golden Rule.

(1) The Universalizability principle can help us with the *different situations* problem:

Universalizability: If an action is right (wrong, good, bad, etc.), then any *relevantly or exactly similar* action would also be right (wrong, good, bad, etc.) regardless of the individuals involved.

Two actions are *relevantly similar* if the reasons why one fits in a given moral category (right, wrong, good, bad, etc.) also apply to the second. Two actions are *exactly similar* if they have all the same properties in common. While no two actions of the actual world are *exactly similar*, the notion applies usefully to hypothetical cases. To test my moral consistency, I can imagine and evaluate an exactly similar action in which the parties are in different places – in which, for example, I'm on the receiving end of the action.<sup>10</sup>

It would be a mistake to drop the “similar situation” qualifier, as in “If Dr. Davis ought to remove my appendix, then I ought to remove her appendix.” Pyrite-1 makes the same mistake about the Golden Rule. Consider the case where my uncle Al is somewhat deaf. Pyrite-1 has me ask, “Do I want Al to speak loudly to me (in my present situation)?” Since I don't (my hearing being normal), then by Pyrite-1 I ought not to speak loudly to Al – which makes little sense. With Gold-1, I'd imagine myself in Al's place (and thus being hard of hearing) in the *reversed situation* – and I'd ask how I desire that I'd be treated in this situation. Presumably, I desire that if the situation were reversed, and I was hard of hearing, then Al would speak loudly to me; so I'd speak loudly to him – which makes more sense.

Gold-1 has me ask “Am I now willing that if I were in X's place in the *reversed situation* then A be done to me?” There are two good ways to take “reversed situation.” I could imagine an *exactly similar reversed situation*, where I have *all* of X's current properties and X has *all* of my current properties.<sup>11</sup> Or, instead of switching all the properties in my mind, I could switch just those relevant to evaluating the act. If I'm unsure whether a property is relevant, I could switch it anyway – just to be safe. This approach has me imagine a *relevantly similar reversed situation*. Either approach works fine.<sup>12</sup>

(2) The Logicality principle can show us a flaw in how Pyrite-2 is formulated:

Logicality: Don't combine logically inconsistent beliefs; and don't believe

---

<sup>9</sup> See the first three works of mine mentioned in the first footnote.

<sup>10</sup> “If it's permissible for me to do A to X, then in the reversed situation it would be permissible for X to do A to me” is an implication of universalizability. I don't regard it as a form of the Golden Rule, since it doesn't relate my *action* toward another to my *desire* about how I am to be treated.

<sup>11</sup> Technically, we should just switch properties that are *universal*, in the sense of being expressible without proper names (like “Gensler” or “Chicago”) or pointer words (like “I” or “now”); see *Formal Ethics* 70.

<sup>12</sup> Sometimes actual situations that we recognize to be relevantly similar can serve as well as imagined hypothetical situations: “Don't combine *acting* to do A to X, *believing* that Y's doing A to you is relevantly similar, and *being unwilling* that Y do A to you.”

something without believing what logically follows from it.<sup>13</sup>

Consider Kant's case where you're a judge about to jail a criminal. You believe this:

(a) All dangerous criminals ought necessarily to be jailed.

Since you're consistent, you also believe what follows from this:

(b) If I were I dangerous criminal, then I ought to be jailed.

Correspondingly, you also *desire* that if you were a dangerous criminal then you be jailed. So you satisfy both Logicality and Gold-1.

It would be a mistake to think that, to be consistent in holding (a), it must also be the case that if you were a dangerous criminal (with a flawed criminal mind) then you'd believe that you ought to be jailed. Pyrite-2 makes a similar mistake when it holds that you cannot justifiably jail the criminal unless it is the case that if you were in his place (and thus had a flawed criminal mind) then you'd desire to be jailed.

The point here is subtle and often misunderstood. The Golden Rule, properly understood, is about *our present reaction to a hypothetical case*. It isn't about how we'd react if we were in the hypothetical case. Compare these two questions:

Gold-1:  
Am I now willing that if I were in X's place in the reversed situation then A be done to me?

Pyrite-2:  
If I were in X's place in the reversed situation, would I then be willing that A be done to me?

Suppose that I have a two-year-old son, little Will, who doesn't understand electricity and keeps putting his fingers into electrical outlets. I try to discourage him from doing this, but nothing works. Finally, I decide that I need to spank him; but would this violate the Golden Rule? In answering this, I should ask the first question – not the second:

Am I now willing that if I were in Will's place then I be spanked?  
  
(This has "willing that if." It's about my present adult desires about a hypothetical case.)

If I were in Will's place, would I then be willing to be spanked?  
  
This has "if" before "willing." It's about the desires I'd have as a small child.

As an adult, I say "I now desire that if I were in this situation then I be spanked." I might add, "I'm thankful that my parents spanked me in such cases – even though I wasn't

---

<sup>13</sup> Logicality and the other consistency principles mentioned here are subject to implicit qualifications; for example, it is assumed that the logical relationships are ones that are or should be clear to us, that our being consistent won't cause Dr. Evil to destroy the world, and so forth. See *Formal Ethics* 19–23 and 108.

pleased then.” Thus I can spank my child without breaking the Golden Rule, since I’m willing that I would have been treated the same way in the same situation.

This distinction is crucial when we deal with someone who isn’t rational – such as one who is drunk and confused, senile, or in a coma. Suppose that a friend at a party is drunk and confused, completely incapable of driving, and yet wants to drive home. You tell him that he cannot drive home – and you’re willing that if the situation were reversed (and you were drunk and confused) then you be told the same thing. You tell your friends “If I ever get drunk and confused in this way, and yet want to drive home, please tell me that I cannot drive home.” In applying the Golden Rule to this case, you need to ask the first question, not the second:

Am I now willing that if I were drunk and confused then I be told that I cannot drive home?

If I were drunk and confused, then would I then be willing to be told that I cannot drive home?

With the second question, drunk and confused desires provide the norm of how you are to treat your friend. Properly understood, the Golden Rule is about our present attitude toward a hypothetical case. To use it correctly, say “I’M WILLING THAT IF” – don’t say “I WOULD BE WILLING.”

(3) This Ends-means principle can teach us about the *flawed desires* problem:

Ends-means: Don’t combine *wanting* to attain end E, *believing* that taking means M is needed to attain E, and *not acting* to take means M.

Suppose that pre-med student Mona violates this principle, and thus has her ends and means out of harmony. What should Mona do? This depends on her situation. Maybe she should take means M (e.g., study harder) or maybe she should give up end E (her goal of becoming a doctor – since maybe she is awful at biology and faints at the sight of blood). Our ends-means principle doesn’t tell her specifically what to do; instead, it tells her that her ends and means are out of harmony and that something has to change.

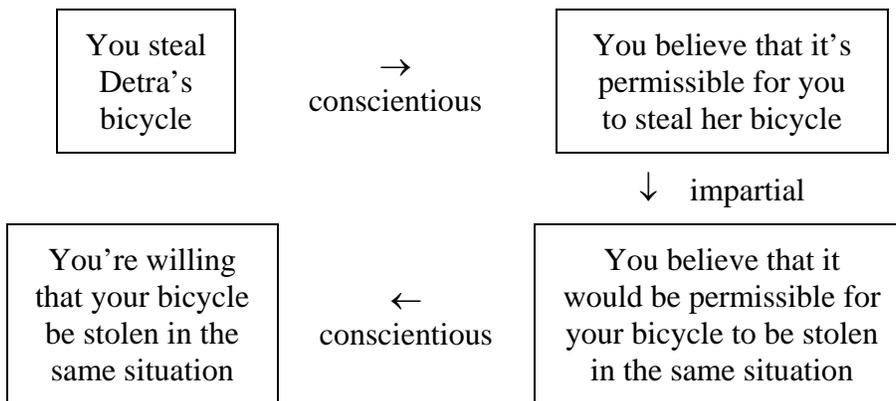
It would be a mistake to take the ends-means principle as a direct guide to action, to assume that your ends (and ends-means beliefs) are perfectly fine and that you only need to act accordingly. Maybe your ends are seriously flawed. Maybe your goal to become a doctor is unrealistic; or maybe your goal to destroy the world is evil. Pyrite-1 and -2 make the same mistake, since they assume that our desires about how we are to be treated are perfectly fine and give us a flawless guide on how to treat others. Our desires can be seriously flawed. Sometimes, in a fit of self-destructive depression, we might want everyone to hurt us; we cannot conclude that our duty then is to hurt others. Or imagine prideful Peter who helps others who are in need but is unwilling that others help him in similar cases; while Peter’s actions are commendable, his desires about how he be treated are flawed. Or imagine racist Rita, who was conditioned by her Nazi society to desire suffering for all Jews and even for herself if she were found out to be Jewish. Fortunately,

as we will see later, there are rational ways to criticize such flawed desires.<sup>14</sup>

Properly understood, the Golden Rule is a consistency principle; instead of telling us what specific actions to do, it directs us to avoid inconsistent combinations. So the rule doesn't replace or compete with regular moral norms, like "It's wrong to enslave others," which tell us directly what to do. Instead, it prescribes consistency – that we not have our actions (toward another) be out of harmony with our desires (about a reversed-situation action). The Golden Rule points out a certain inconsistency in our behavior, which usually comes from disregarding the interests of others.

What sort of *inconsistency* do we have when we violate the Golden Rule? Clearly we don't have a logical inconsistency between beliefs; in fact, what clashes here isn't beliefs but rather actions and desires. But why is it inconsistent to violate Gold-1?

Consistency in a broad sense is more than just logical consistency in beliefs. It also includes, as we have seen, things like consistency between ends and means. More important here are the consistency (which I call *conscientiousness*) between our moral beliefs and how we live (our actions, intentions, and desires) – and the consistency (which I call *impartiality*) between our evaluations of similar cases. Gold-1 is a consequence of these two kinds of consistency. Suppose that you're *conscientious* and *impartial* in these senses, and yet you want to steal Detra's bicycle. Being conscientious, you won't steal her bicycle unless you think this act is permissible (all right). Being impartial, you won't think this act is permissible unless you think it would be permissible for your bike to be stolen in the same situation. Being conscientiousness, you won't think this unless you're willing that your bike be stolen in the same situation. So if you're conscientious and impartial, then you won't steal Detra's bicycle unless you're willing that your bike be stolen in the same situation. Here's a diagram:



So if we're conscientious and impartial, then we'll follow Gold-1: we won't do some-

---

<sup>14</sup> One might suggest a phrasing about what is *correct* (or perhaps *rational* or *mature*) to desire about the reversed situation. While this works, it would be difficult to apply; and it makes the Golden Rule a trivial addition to whatever decides the correctness of desires (see the later Dabney footnote). So I prefer to keep Gold-1 as it is but supplement it with further rationality norms, about things like knowing the facts and rationalizing one's desires, as in the next section. Gold-1 requires these further elements, not in its formulation, but to lead us more effectively into knowing the right or reasonable thing to do. While Golden Rule consistency is an important part of moral rationality, it needs to work together with other elements.

thing to another unless we're willing that it be done to us in the same situation. So if we violate Gold-1, then we violate either conscientiousness or impartiality or both. So if we assume that we ought to be conscientious and impartial, then we can *deduce* that we ought to follow the Golden Rule.

Before leaving this section, we need to consider how the Golden Rule works if three parties are involved. Recall the case where you must hire either Alice or Betty (but not both) and each wants to be the exclusive hire. Here you must satisfy the Golden Rule toward both. We can formulate a three-party Golden Rule as follows:

Gold-2: Don't act in a given way toward X and Y without consenting to the idea of this act being done when you imagine yourself in the place of X and also consenting to the idea of this act being done when you imagine yourself in the place of Y.

To choose consistently, you must be willing for your choice to be made regardless of whether you're in Alice's or in Betty's place. So if you hire Alice instead of Betty because Alice is more qualified, you must be willing that you'd be hired if you were in Alice's more-qualified place but *not* if you were in Betty's less-qualified place. Gold-3, which resembles Kant's formula of universal law, expresses this idea more elegantly:<sup>15</sup>

Gold-3: Act only as you're willing for anyone to act in the same situation, regardless of where you imagine yourself or others in the situation.

C.I. Lewis expressed a roughly similar idea when he said "Act as if you were to live out in sequence your life and the lives of those affected by your actions."<sup>16</sup>

### III

**How does the Golden Rule work?** Suppose you're thinking about doing something to another and you want to apply the Golden Rule. I suggest these four steps:

1. **KNOWLEDGE:** Try to understand the facts; in particular, try to understand how your action would affect the other person.
2. **IMAGINATION:** Try to imagine yourself, vividly and accurately, in

---

<sup>15</sup> *Formal Ethics* 101–104 delineates 6,480 variations on Gold-1 that are consistency theorems. Theorems more distantly related to Gold-1 include the *self-regard* and *future-regard* principles (see *Formal Ethics* 121–23 and *Ethics* 115–17). I violate self-regard consistency if I treat myself in some way (e.g., by continuing in my drug habit and not seeking treatment) that I'm not willing to have others that I care about treat themselves in similar circumstances. I violate future-regard consistency if I treat myself in some way in the future (e.g., by getting drunk in a way that will cause me a painful hangover in the future) that I'm unwilling to have been treated by myself in the past.

<sup>16</sup> *An Analysis of Knowledge and Valuation* (La Salle, IL: Open Court, 1946), 547. *Formal Ethics* 128–31 lists a dozen other ways to express roughly the same idea, such as "Act only in ways that you'd consent to if you didn't know your place in the situation," which was inspired by John Rawls's *A Theory of Justice* (Cambridge: Harvard University Press, 1971). As Wattles put it after mentioning the love norm (*The Golden Rule* 50), "There is a unity in the moral life, and the Golden Rule is one way of expressing that unity."

the other person's place on the receiving end of the action.

3. DESIRES: Sometimes you may need to rationalize your desires about how you'd be treated in the other person's place (see below).
4. QUESTION: Ask yourself, "Am I willing that if I were in the other person's place in the reversed situation then this be done to me?"

If you act in a given way toward another and yet are unwilling that you be treated this same way in the reversed situation, then you are inconsistent and violate the Golden Rule. Then your action (toward another) is out of harmony with your desires (about how you are to be treated in the reversed-situation) – and one or the other needs to be changed.

When we apply the Golden Rule, usually our desires are fine and we just need to treat others as we desire that we'd be treated in their place. But sometimes our desires are flawed and need to be changed; we can counter flawed desires by appealing to consistency, knowledge, and imagination. Consider prideful Peter, who violates Golden Rule consistency while helping others:

- Prideful Peter helps others who are in need but is unwilling that others help him in similar cases.

We can ask Peter, "Isn't it good for you to help others? Then why wouldn't it be good for others to help you, when you are in need? What would you think about someone else who was too proud to accept your help?" If Peter is religious, we might add, "Don't you think that the reason we are here is to help one another?" A more extreme case of flawed desires is racist Rita, who satisfies Golden Rule consistency while torturing Jews:

- Racist Rita tortures Jews and was taught by Nazis to desire torture for all Jews and even for herself if she were found out to be Jewish.<sup>17</sup>

We can try to get to get Rita to see how her anti-Jewish feelings came from Nazi misinformation (about racial superiority, anti-Jewish stereotypes, etc.) and to broaden her knowledge and experience of Jewish people in an open way. She might also reflect on other socially taught stereotypes and prejudices. All over the world, people in one group are taught to dislike people in another group; we teach young children:

Be suspicious of those other people. They're of a different race (or religion, or ethnic background, or sexual preference, or caste). They aren't our kind. They have strange customs. They're evil and inferior.

When we broaden our knowledge and experience, we conclude, "They're people too, much like us, with many of the same virtues and vices."<sup>18</sup>

---

<sup>17</sup> There's a supposedly true story about a Nazi who hated Jews and put them in concentration camps. One day he discovered that he had Jewish ancestry. Since he hated Jews in a disinterested way, he came to hate himself and his family. So he had himself and his family put into concentration camps and killed. While this Nazi was consistent, his hatred could likely be rationally criticized.

<sup>18</sup> Richard Brandt proposed that a "rational desire" is one that would survive "cognitive psychotherapy," a maximal criticism in terms of logic and a vivid exposure to facts. Suppose you have an aversion to yogurt

The Golden Rule is not a *compass* but a *path*. If the rule were a *moral compass*, we could just take it out and see where it points. But it's not that easy. Instead, the Golden Rule is a *path that we have to walk*; to apply it rationally, we need knowledge, imagination, and rationalized desires. We first need KNOWLEDGE. We have to get to know the people involved, perhaps our children, students, employees, or patients. This takes time and energy, listening and dialoguing, openness and sensitivity. And we need IMAGINATION. We have to imagine ourselves in the other person's place in a vivid and accurate way – a skill we can develop through social interaction and literature.<sup>19</sup> And we may need to RATIONALIZE OUR DESIRES. So we may have to counter the influence of an ignorant society that taught us to hate a certain group, and even to hate ourselves when we imagine ourselves in that group. Only when we have walked a good distance down this path can we apply the Golden Rule wisely, as we struggle about how we ought to live.

Let's now look at a practical example: how the Golden Rule works against segregation and slavery.<sup>20</sup> President John F. Kennedy in 1963 appealed to the Golden Rule in an anti-segregation speech at the time of the first black enrollment at the University of Alabama. He first presented facts about how blacks were treated as second class citizens. He then asked whites to consider what it would be like to be treated thusly because of skin color. Whites were to imagine themselves being black – and being told that they couldn't vote, go to the best public schools, eat at most public restaurants, or sit in the front of the bus. Would whites be content to be treated that way? He was sure they wouldn't; yet this is how they treated others. He said the “heart of the question is whether we are going to treat our fellow Americans as we want to be treated.” In accord with our strategy, he appealed to knowledge, imagination, and a question about our reversed-situation desires.<sup>21</sup>

History gives us impressive examples of people who were brought up to have pro-slavery intuitions but later used the Golden Rule against slavery; these examples include the Stoic philosopher Epictetus in 90, the Pennsylvania Quakers in 1688, and Harriet Beecher Stowe and Abraham Lincoln in the 1800s.<sup>22</sup> Yet the use of the Golden Rule to

---

that comes from social conditioning; although you never tried yogurt, you were told that it has bad germs, tastes awful, and is eaten only by weird people. It may be that your aversion would change if you learned more about yogurt, tried it, and understood the origin of your aversion; then your aversion is irrational. Or suppose you have an aversion to group X because your family hated Xs, called them names, taught false stereotypes about them, and had you meet only a few atypical Xs. It may be that your aversion to Xs would diminish if you understood its origin and broadened your experience and knowledge of Xs in an open way. Cognitive psychotherapy can be a useful tool for conquering prejudices. See Brandt's “Rational Desires,” in the *Proceedings and Addresses of the American Philosophical Association* 43 (1970): 43–64.

<sup>19</sup> Imagining another's perspective is a common and important human experience. A child pretends to be a mother or a cowboy. A chess player asks, “If I were in my opponent's place, how would I respond to this move?” A writer dialogues with an imagined reader who misunderstands and raises objections. A teacher asks, “How would I respond to this assignment if I were a student?” The ability to take another's perspective is especially important for applying the Golden Rule.

<sup>20</sup> This is an apt example, because Green's “Parsing Reciprocity: Questions for the Golden Rule” mentions pro-slavery thinkers who tried to counter Golden Rule objections to slavery.

<sup>21</sup> His speech is reprinted in the 12 June 1963 *New York Times*, 20.

<sup>22</sup> See *The Golden Sayings of Epictetus*, trans. H. Crossley (London: Macmillan, 1935), 90; W.E. Phipps's “The Glittering Rule” (*Theology Today* 39 (1982): 194–98), 195 for the Quaker and Lincoln references; and Stowe's *Uncle Tom's Cabin* (<http://www.gutenberg.org/etext/203> – first published 1851–52 in the *National Era*), chap. 12.

condemn slavery has been disputed. Consider this (slightly edited) dialogue between two women, one anti-slavery and one pro-slavery, from *Uncle Tom's Cabin*:

Anti: The most dreadful part of slavery is its outrages on the feelings – the separating of families, for example.

Pro: That's a bad thing. But it doesn't occur often.

Anti: Oh, it does. I've lived many years in Kentucky and Virginia, and I've seen enough to make one's heart sick. Suppose, ma'am, that your two children be taken from you and sold?

Pro: We can't reason from our feelings to those of this class of persons.

Anti: You know nothing of them! They feel just as keenly as we do.

Pro: Indeed! I think they are better off as slaves.

Put into Gold-1 terms, Anti asks “Are you now willing that in a similar situation your children be sold into slavery?” Pro answers that blacks don't feel the pain that whites would feel in this case; so presumably, yes, she is willing that in a similar situation where she is black (and thus cares little about her children) her children be sold into slavery. To rebut Pro's beliefs about how blacks don't care about their children, the novel later describes how a young slave woman, whose young son was taken from her and sold, was heartbroken and drowned herself. This dialogue reminds us that in applying the Golden Rule we need to get the facts straight about how our actions affect others.

Writing just after the Civil War, the Protestant minister Robert Lewis Dabney criticized the use of the Golden Rule (in its flawed Pyrite-2 form) to condemn slavery and he argued that the rule just as easily supports slavery:

The whole reasoning of the Abolitionists proceeds on the absurd idea, that any caprice or vain desire we might entertain towards our fellowman, if we were in his place, and he in ours, must be the rule of our conduct towards him, whether the desire would be in itself right or not. This absurdity has been illustrated by a thousand instances. On this rule, a parent who, were he a child again, would be wayward and self-indulgent, commits a clear sin in restraining or punishing the waywardness of his child, for this is doing the opposite of what he would wish were he again the child...

The Golden Rule binds the slave just as much as the master... The slave may be very sure that, were he the master, he would naturally desire to retain the services of the slaves who were his lawful property. Therefore, according to this abolition rule, he is morally bound to decline his own liberty; i.e., to act towards his master as he, were he the master, would desire his slave to act.<sup>23</sup>

Dabney appealed to the flawed Pyrite-2 formulation in defending slavery.

---

<sup>23</sup> *A Defense of Virginia* (New York: E.J. Hale, 1867), 196–97. For Dabney, the Golden Rule must be seen in terms of what is *right* to desire; the pro- and anti-slavery sides disagreed about whether it was *right* to desire the abolition of slavery. Dabney's formulation made the rule powerless to promote social change.

How should the Golden Rule have been applied in the mid-1800s to the question of whether to keep or abolish slavery? Again, we need a four-step procedure: *knowledge*, *imagination*, *desires*, and *question*. Since the choice to keep or abolish slavery involves many people, it would be better to apply a consistency condition closer to Gold-3:

Don't believe that slavery should be kept (or abolished) unless you desire that this be done regardless of where you imagine yourself in the situation.

Back in the mid-1800s, the issue depended largely on the facts. Pro- and anti-slavery sides saw the facts very differently:

Pro: Blacks by nature are dull-witted. Most slave owners treat slaves well, almost as if the slaves were their own retarded children. Slavery benefits both slaves (who would do poorly on their own) and slave owners (who profit from the work of slaves).

Anti: Blacks by nature are fairly intelligent (although less than whites<sup>24</sup>) and can live well on their own. Slave owners often treat slaves cruelly, as when they separate slave families for financial gain. Slavery benefits slave owners at the expense of slaves.

The Golden Rule didn't in the mid-1800s give a easy answer to the slavery issue, since so much depended on whether Pro or Anti better described the facts. Was evidence available back then to resolve this question? I'm not enough of an historian to be able to say for sure either way. But I tend to think there *was* such evidence (although it could be ignored), much of it coming from escaped slaves: their ability to thrive on their own as free people and the stories they told about how they were treated.<sup>25</sup> If so, then the Golden Rule argument against slavery back then was strong (although it too could be ignored). The Golden Rule requires facts; once it has these, it becomes a powerful weapon against racism and all related forms of unjust discrimination.<sup>26</sup>

Let me summarize my four suggestions about how to apply the Golden Rule:

1. Imagine yourself in the other person's situation.
2. Ask yourself "Am I now willing that if I were in the other person's

---

<sup>24</sup> Even the anti-slavery side was racist by today's standards. In the mid-1800s, we must remember, views about groups differing in native intelligence (e.g., that the Irish by nature were less intelligent than the English) were widespread. Also, at that time, blacks had had little opportunity for advancement.

<sup>25</sup> Prominent examples include Harriet Jacobs (1813–97), who escaped from slavery and wrote a very articulate book, *Incidents in the Life of a Slave Girl*, about the horrors (including sexual abuse) of her life as a slave, and Frederick Douglass (1818–95), whose *Narrative of the Life of Frederick Douglass, an American Slave* similarly described the horrors of slavery and gained an international reputation. Douglass's owners had taught him to read, even though it was illegal in Maryland to teach one's slave to read.

<sup>26</sup> I've argued (*Ethics* 114–15 and *Formal Ethics* 81–82 and 117–18) that the Golden Rule applies to any being capable of experience and so can also be used against animal cruelty. An African proverb goes "One who is about to poke a baby bird with a pointed stick should first try it on himself to feel how it hurts."

- place in the reversed situation then this same thing be done to me?”
3. If you aren't, and yet you act this way toward the other, then you're inconsistent (and either your action or your desire need to be changed).
  4. Applying the Golden Rule adequately requires knowledge, imagination, and sometimes rationalized desires.

## IV

**How does the Golden Rule matter?** But first, *does* it matter? Is it gold or fool's gold? Most moral philosophers reject the rule as a folk principle that falls apart and leads to absurdities when examined closely. I've argued, however, that the critics have not examined the rule carefully enough. The rule *is* genuine gold, in fact solid gold, and can be reformulated to avoid the absurdities and reveal the lasting luster.<sup>27</sup>

The Golden Rule matters because it captures so much of the *spirit* behind morality. It helps us to see the point behind moral rules. It counters self-centeredness. It engages our own reasoning, instead of imposing an answer from outside. It's psychologically sound and personally motivating. It promotes mutual understanding and cooperation, helping us to reason together about ethical issues and promote just causes. And it concretely applies ideals like fairness and concern. So the Golden Rule makes a good one-sentence summary of what morality is about.

The Golden Rule matters to the lives of many people. If you ask professionals to explain their responsibilities toward others, many will respond using the Golden Rule. So a thoughtful nurse might say, "I try to treat my patients as I'd want to be treated in their place." Many in education, business, or government say similar things. Parents often use the rule when teaching morality; children are encouraged to understand the facts, put themselves in the other person's place, and treat others as they want to be treated themselves. While the Golden Rule is simple enough for children, adults can use it in increasingly more sophisticated ways.

On the negative side, the Golden Rule has some subtle aspects; but we need these to avoid the absurdities and defend the rule. Less sophisticated people can fall back on cruder formulations. Many people, oblivious to instances like "If you want Dr. Davis to remove your appendix, then remove her appendix," apply Pyrite-1 or -2 with an intuitive wisdom that transcends the flawed literal meaning of these principles. Many people are better at living the Golden Rule than at formulating it to avoid absurdities.<sup>28</sup>

The Golden Rule matters even more because it can be based on so many different approaches to ethics. What do you think ethics is based on? *Intuitively self-evident princi-*

---

<sup>27</sup> Wattles's *The Golden Rule* gives three strategies for responding to objections against the Golden Rule: "abandon the rule, reformulate it, or retain it as commonly worded, while taking advantage of objections to clarify its proper interpretation" (8). While he takes the third approach, I think we defend the rule better by reformulating it. Wattles brings up some of the points that I mention; for example, he says "When we say, 'Do not treat others as you do not want others to treat you,' there is the unspoken assumption 'in (essentially) the same situation'" (35). A reformulation makes this more explicit.

<sup>28</sup> This situation isn't unusual in philosophy. For example, most people are fairly good at inductive reasoning (or else humanity wouldn't have survived this long), but yet there are unresolved issues about how to formulate principles of inductive logic (see my *Introduction to Logic*, chap. 13).

ples? Then you can see the Golden Rule (or the more basic consistency axioms from which it can be derived) as self-evident. *A rational procedure?* The Golden Rule uses facts, imagination, rationalized desires, and consistency. *God's will?* Almost every religion teaches the rule.<sup>29</sup> *Cultural conventions?* Almost every society endorses the rule. *A social contract for mutual advantages?* The Golden Rule is a simple idea that promotes cooperation and helps people to resolve social conflicts. *Social usefulness?* The rule is highly useful to society. *Personal feelings?* Most people have feelings that support consistency and the Golden Rule.

We can have higher or lower motives to follow the Golden Rule. We may follow the Golden Rule because we want to act in a rational and morally correct manner, or because we care about other people for their own sake, or because we want to act out of gratitude to God who first loved us and calls upon us to love one another as sisters and brothers. Or we may follow the Golden Rule out of enlightened self-interest, thinking that treating others as we want to be treated makes us feel better about ourselves, encourages others to respect us and treat us better, and helps us avoid the painful self-condemnation and cognitive dissonance that come with inconsistency.<sup>30</sup>

The Golden Rule matters even more because it's so global. For centuries, it has had wide support among the various religions and cultures of the world. Confucius, the Rabbi Hillel, and Jesus Christ all use the rule to summarize their teachings. Buddhism, Hinduism, Islam, Taoism, and Zoroastrianism also support the rule, as do secular thinkers from diverse cultures. Many of these give the rule a central status in moral thinking. The Golden Rule is close to being a global norm common to all peoples of all times.<sup>31</sup>

Nations need to think more in terms of the Golden Rule and less in terms of war and terrorism.<sup>32</sup> The Golden Rule, with roots in a wide range of world cultures, is well suited

---

<sup>29</sup> For a reflection on the added dimension that the Golden Rule takes on when embedded in a religious context, see my "God, Science, and the Golden Rule," in *The Sheed & Ward Anthology of Catholic Philosophy*, ed. James C. Swindal and Harry J. Gensler (Lanham, Md.: Rowman & Littlefield, 2005), 523–31.

<sup>30</sup> Wattles (*The Golden Rule* 3–5 and back cover) asks whether the Golden Rule has the same meaning when used in diverse contexts. For example, we might ask whether it means the same thing for altruists and egoists. I'd answer by distinguishing two senses of "meaning": (a) linguistic meaning and (b) role or function. "A thousand dollars" has the same linguistic meaning for the rich and the poor (since there's no verbal ambiguity if both use the same expression); but the role or function of a thousand dollars differs, since it is little for the rich but much for the poor. I'd say the same about the Golden Rule. While the rule may have the same linguistic meaning (e.g., Gold-1 or Pyrite-2) for altruists and egoists, its role or function differs. Altruists use the rule to operationalize the idea of loving one's neighbor for the sake of the neighbor; egoists use it to promote their self-interest.

<sup>31</sup> The Parliament of the World's Religions in 1993 adopted a "Declaration Toward a Global Ethic" (see [http://www.weltethos.org/pdf\\_decl/Decl\\_english.pdf](http://www.weltethos.org/pdf_decl/Decl_english.pdf)) which emphasizes the Golden Rule: "There is a principle which is found and has persisted in many religious and ethical traditions of humankind for thousands of years: *What you do not wish done to yourself, do not do to others.* Or in positive terms: *What you wish done to yourself, do to others.* This should be the irrevocable, unconditional norm for all areas of life, for families and communities, for races, nations, and religions." This declaration was signed by representatives from Baha'i, Brahma Kumaris, Buddhism, Christianity, Native Religions, Hinduism, Jainism, Judaism, Islam, Neo-Paganism, Sikhism, Taoism, Theosophism, Zoroastrianism, and Interreligious Organizations.

<sup>32</sup> Consider the American response to the terrorist attack of 11 September 2001 on New York in which 3,000 people died. The U.S. did little to understand the perspective of the Muslim extremists, including why they hated the U.S. and whether they might have legitimate complaints. Instead, the U.S. resorted to military force in Afghanistan and later in Iraq, which has done little to solve the root problems. The Golden

to be a standard that different cultures can appeal to in resolving conflicts. As the world becomes more and more a single interacting global community, the need for such a common standard is becoming more urgent.

---

Rule approach would have started by appealing to *facts* and *imagination* (which includes trying to understand the situation from the perspective of both sides) – and by trying to get the other side to do likewise.