

JOHN STUART MILL
AND THE ART OF LIFE

Edited by Ben Eggleston, Dale E. Miller, and David Weinstein

OXFORD
UNIVERSITY PRESS

2011

OXFORD

UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further Oxford University's objective of excellence in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Copyright © 2011 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data
John Stuart Mill and the art of life / edited by Ben Eggleston, Dale E. Miller, and David Weinstein.

p. cm.

ISBN 978-0-19-538124-5

1. Mill, John Stuart, 1806–1873. I. Eggleston, Ben, 1971– II. Miller, Dale E., 1966–

III. Weinstein, D. (David), 1949–

BR1607.J63 2010

192—dc22 2009052293

9 8 7 6 5 4 3 2 1
Printed in the United States of America
on acid-free paper

Rules and Their Reasons

Mill on Morality and Instrumental Rationality

BEN EGGLESTON

1. Introduction

One of the major unresolved questions about John Stuart Mill's theory of morality is what role Mill regards rules as playing in the determination of morally permissible action. This question is implicated, for example, in the long-standing controversy over whether Mill is better understood as an act utilitarian or as a rule utilitarian. In this paper, I approach this question about Mill's view of rules from a different perspective, one based on Mill's remarks about instrumentally rational action in the final chapter of *A System of Logic*. To provide context for interpreting those remarks, I'll devote section 2 of this paper to an overview of consequentialist normative theories, and I'll devote section 3 to a discussion of an important objection to rule-based consequentialist theories. In section 4, I'll summarize the considerable textual evidence suggesting that Mill's moral theory is, in fact, a rule-based consequentialist one. Section 5 focuses on passages in the final chapter of the *Logic* that, I argue, suggest that Mill anticipates and endorses the objection to rule-based consequentialist theories described in section 3. In section 6, I'll conclude by exploring some ways in which this tension in Mill's thought might be resolved.

2. Consequentialist Normative Theories

Normative theories specify principles that imply judgments in particular cases. The kind of judgments implied by a given normative theory depends on its normative domain. One important normative domain is that of the morality of actions: theories in this domain offer judgments of the rightness and wrongness of actions. Another important normative domain is that of the instrumental rationality of actions: theories in this domain characteristically identify some actions as instrumentally rational, and others as falling short of this standard. In principle, it makes sense to allow that there are as many normative domains for the judging of actions as there are evaluative dimensions applicable to actions (such as politeness, cleverness, and generosity). But we can restrict our inquiry to theories of two classes: theories of the morality of actions and theories of the instrumental rationality of actions—or, more briefly, theories of morality and theories of instrumental rationality.

Consequentialist normative theories determine the normative status of actions with reference to consequences of certain kinds.¹ Accordingly, consequentialist theories of morality determine the rightness and wrongness of actions with reference to consequences of certain kinds, with the designation of the consequences that matter being the main source of the divergences among various consequentialist moral theories. For our purposes, the divergence between act-consequentialist moral theories and rule-consequentialist ones is central.

Act-consequentialist moral theories say that the consequences that matter are the consequences of the particular act that is being evaluated. Thus, the defining principle of act consequentialism is typically formulated as follows:

An act is right if and only if its consequences are at least as good as the consequences of any other act that the agent could have performed instead.²

In contrast, rule-consequentialist theories say that the consequences that matter, for the determination of the rightness and wrongness of actions, are the consequences of the general acceptance (or internalization or inculcation) of a system of rules allowing the action in question. Thus, a standard formulation of rule consequentialism is as follows:

An act is right if and only if it would be allowed by a system of rules the consequences of whose general acceptance would be at least as good as the consequences of the general acceptance of any other system of rules.³

So whereas act consequentialism can be understood as identifying the right act by ranking the possible acts in terms of the consequences of their being

performed and selecting the highest-ranked act (or a highest-ranked act, if there is a tie), rule consequentialism can be understood as identifying the right act by ranking the possible systems of rules in terms of the consequences of their being generally accepted, selecting the highest-ranked system of rules, and selecting an act that that system of rules would allow in the circumstances in question.

Although morality is the normative domain to which Mill has made the most well-known contributions, the chapter of the *Logic* that will interest us later actually bears more directly on instrumental rationality. Theories of morality normally require the agent to respect the interests of others in ways that go beyond the concern for others that may already be incorporated into the agent's own interests and ends. In contrast, theories of instrumental rationality are essentially concerned only with an agent's own interests and ends. Thus, such theories may regard, as rational, an act that any plausible moral theory would regard as immoral. Additionally, the qualifier "instrumental" is important: in contrast to theories of rationality in some strong (for example, Kantian) sense, theories of instrumental rationality evaluate actions simply on the basis of their being appropriately related to the maximal advancement of the agent's self-interest, or to the agent's life's going as well as possible. This is the nature of instrumental rationality, whatever else may be said about rationality in some deeper sense.

The divergence between act-consequentialist and rule-consequentialist moral theories is paralleled by the divergence between egoistic and rule-egoistic theories of instrumental rationality. Egoism is typically formulated as follows:

An act is instrumentally rational if and only if its consequences are at least as good for the agent as the consequences of any other act that the agent could have performed instead.⁴

As noted above, act consequentialism and egoism pertain to different normative domains. But they also have an obvious structural similarity: they both require agents to choose acts that maximally advance certain interests. On the basis of this similarity, we can refer to both of these theories as *straightforwardly maximizing* normative theories.

Rule egoism, in turn, may be formulated as follows:

An act is instrumentally rational if and only if it would be allowed by a system of rules the consequences of whose acceptance by the agent would be at least as good for him or her as the consequences of his or her acceptance of any other system of rules.⁵

Like act consequentialism and egoism, rule consequentialism and rule egoism have an obvious structural similarity: they both require agents to choose

acts that comply with rules that are optimal in a certain respect. On the basis of this similarity, we can refer to both of these theories as *rule-based* normative theories.

So, we have the following simple taxonomy of the normative theories with which we shall be concerned:

<i>normative domain</i>			
		<i>morality</i>	<i>instrumental rationality</i>
structure	straightforwardly maximizing	act consequentialism	egoism
	rule-based	rule consequentialism	rule egoism

In the domains of both morality and instrumental rationality, the straightforwardly maximizing theories are, in a sense, the default options, at least for consequentialists: if one believes that morality or instrumental rationality is best theorized along consequentialist lines, then one is likely to be drawn, at least as a first option, to the idea that the normatively privileged act (the right one, or the instrumentally rational one) is simply the one with the outcome that is best (in whatever way the goodness of outcomes is assessed). But such theories are also subject to certain criticisms that have motivated the development of rule-based alternatives.

Many of these criticisms are based on the idea that, despite the simplicity and initial appeal of straightforwardly maximizing theories, they actually tend to conflict significantly with our considered judgments about the normative domains to which they apply. Act consequentialism, it is commonly said, clashes with our considered judgments by allowing that any kind of act—lying, breaking a promise, betraying a friend, punishing an innocent person, violating the solemn rights of a few in order to generate trivial benefits for the many—is permissible whenever the consequences of doing so are even slightly better than the consequences of acting otherwise. In fact, act consequentialism entails not only that such acts are permissible in such circumstances but also that they are obligatory. It is one thing, advocates of this criticism say, for a theory to allow that any kind of act might be permissible or even obligatory if committing such an act is the only way to avoid a true catastrophe. What renders act consequentialism objectionable, on this way of thinking, is that it requires such acts even in relatively ordinary contexts—as long as their overall consequences are better (by no matter how trifling a margin) than the consequences of acting otherwise.⁶

Another criticism of act consequentialism has also been influential in motivating the development of forms of rule consequentialism. This criticism is that when agents regard it as their moral duty to bring about the best possible consequences every time they act, the consequences tend to be worse than when they construe their moral duty differently. One consideration supporting this conclusion is that people tend to misapply the straightforwardly maximizing requirement of acting in whatever way will produce the best consequences. Due to obvious limitations of human powers of reasoning and also more insidious factors such as the temptation to view one's preferred option unduly favorably, people tend to both incorrectly estimate the probabilities of the possible consequences of their acts and incorrectly compare the values of those possible consequences. A second, and more subtle, consideration is that when people do not subscribe to general rules prohibiting acts of certain kinds, the result is the decay and sometimes the disappearance of valuable personal traits and social customs such as honesty, promise-keeping, friendship, and respect for individual rights. And when these background conditions wither, everyone tends to be worse off. Act consequentialism, then, does not fare well when assessed *pragmatically*—that is, in terms of whether the general acceptance of it would have good consequences. Since act consequentialism would presumably counsel people not to subscribe to any moral theory—including it—that did not lead to the best available consequences, act consequentialism is often criticized on this basis for being self-defeating.⁷

These difficulties for act consequentialism motivate the turn to, and correspond to strengths of, rule consequentialism. Since rule consequentialism eschews any case-by-case approach to evaluating acts and instead evaluates acts with reference to rules that allow or prohibit them, an act of a kind that is generally immoral is not deemed moral, from the perspective of rule consequentialism, simply in virtue of having the best consequences available in a particular situation. Instead, the act must be of a kind that it is generally desirable for people to regard as permissible. Of course, this will not be enough to keep rule consequentialism from allowing occasional acts of lying, promise-breaking, and so on; but it will be enough to make rule consequentialism agree much more closely than act consequentialism does with most people's considered judgments about when such acts are permissible and when they are not. Moreover, by evaluating acts with reference to rules rather than on a case-by-case basis, rule consequentialism does not encourage agents to enter into the uncertain activity of comparing the likely consequences of individual acts, and also encourages the maintenance of the sorts of valuable personal traits and social customs discussed above. Thus, the general acceptance of rule consequentialism is likely to have better consequences than the general acceptance

of act consequentialism, and so rule consequentialism is decidedly less vulnerable to the charge of being self-defeating.

Of the two considerations just characterized as motivating the move from act consequentialism to rule consequentialism—intuitive plausibility and pragmatic effectiveness—the second has been far more influential than the first in motivating consequentialists about instrumental rationality to move from egoism to rule egoism. Since instrumental rationality pertains to the agent's own interests, it is rare for cases to arise in which it seems counterintuitive for a theory of instrumental rationality to direct the agent to maximally advance his or her own interests, as egoism does. Furthermore, even restricting our focus to pragmatic effectiveness, egoism is rarely criticized on the grounds that one is likely to misapply the straightforwardly maximizing requirement of choosing the act that will maximally advance one's interests.⁸ Instead, egoism's main vulnerability has to do with structural issues that are essentially analogous to the concerns about personal traits and social customs discussed above.

These structural issues arise in a variety of cases, but it will suffice, for our purposes, to look closely at the workings of just one kind of case. To construct an example of this kind of case, let us suppose that an egoist would like to gain the assistance of another person by offering to reciprocate when the time comes. For example, let us imagine a farmer who would like to persuade another farmer to help him harvest his crops.⁹ He points out to the other farmer that the latter's crops will be ready for harvest just a couple of weeks after his own and tells him that if he (the second farmer) gives him a few days' help, then he'll return the favor a couple of weeks later. Let us suppose, however, that for the first farmer, helping the other farmer then would not maximally advance his interests: regardless of whether the second farmer had, in fact, shown up and helped him with his crops, setting aside his tasks would not advance his interests as much as staying home and tending to his own matters. Knowing this, and knowing that he is an egoist, the first farmer is unable to sincerely assure the second farmer that he will, in fact, return his assistance. As a result, the first farmer is unable to make an offer to the second farmer that makes it worthwhile for the latter to help the former with his harvest. As a result, the first farmer's interests do not end up being advanced as much as they would have been advanced if the first farmer had been able to enlist the cooperation of the second farmer.

To appreciate the force of this pragmatic indictment of egoism, note that the frustration of the first farmer's interests does not depend on any desire on his part to cheat the second farmer: obviously getting the second farmer's cooperation and then not having to reciprocate would advance his interests the most, but his interests would still be advanced (compared to

the default outcome of harvesting his crops himself) if he got the second farmer's assistance and had to reciprocate afterward. The problem is that, as an egoist, he is denied even this outcome: since he can't count on himself to act against his interests, he can't sincerely assure anyone else that he will, either. Of course, other options at his disposal might obviate the need for him to offer that assurance in the present case: he might be able to deceive the second farmer about his ability to act against his interests; he might be able to assure the second farmer that he'll be motivated to reciprocate for moral reasons, if not for reasons of instrumental rationality; he might be able to identify something that the second farmer needs help with beforehand (as opposed to afterward); he might be able to provide the second farmer with some kind of collateral to ensure his own cooperation; or he might be able to tell enough people that he'll reciprocate that the reputation costs of doing otherwise would be prohibitive. But these other possibilities are really irrelevant, according to this pragmatic indictment of egoism. For a truly instrumentally rational person should not have to depend on such extraneous possibilities in order to enter in to something as simple as a cooperative arrangement that requires him to act against his interests, as long as the cooperative scheme is beneficial to him overall.

Such cooperative arrangements are central to the pragmatic indictment of egoism. But they form only one class of cases in which the egoist is disadvantaged by his inability to intend to act against his interests. A second class of cases concerns threats: it can be advantageous to be able to sincerely threaten to do certain things (hit back, vote for the other side, take one's business elsewhere) that would, if the threat were to fail, be disadvantageous to oneself.¹⁰ Third, Kavka's toxin puzzle¹¹ is a far-fetched but philosophically intriguing situation in which egoists would be unable to secure the prize that is, in some sense, there for the taking. Finally, it has been argued that, if one were in the vicinity of a predictor of the kind found in Newcomb's problem, being an egoist would make one fare badly.¹² A variety of cases, then, underwrite the pragmatic indictment of egoism. On this basis, egoism, like act consequentialism, is often regarded as self-defeating.¹³

Just as rule consequentialism can be understood as reacting to and remedying the two main problems associated with act consequentialism (its counterintuitive implications and its pragmatic ineffectiveness), so likewise rule egoism can be understood as a response to the main problem associated with egoism (its pragmatic ineffectiveness). The exact requirements of rule egoism depend on what rules it would be maximally interest-advancing for an agent to adopt, and thus will vary from agent to agent. As a result, here they can be characterized only roughly. But it seems plausible to suppose, based on the desirability of being able to make sincere assurances and threats, that rule egoism will endorse, for any agent, a system of rules allow-

ing him to follow through on a sincerely made assurance or threat, even if doing so would be a net loss to his interests. For if an agent accepts such a system of rules, then presumably he will know that about himself and thus will not anticipate that he will be unable to regard following through on an assurance as rational in the way that the first farmer anticipates that about himself in the example presented above.¹⁴ To be precise, we should note that rule egoism is not likely to endorse a system of rules allowing an agent to follow through on any assurance or threat at all, for it would surely not maximally advance one's interests to be allowed, by the norms of instrumental rationality one accepts, to follow through on assurances or threats to incur grave sacrifices for the sake of trivial benefits.¹⁵ But rule egoism is likely to endorse a system of rules allowing (or possibly requiring) an agent to follow through on certain assurances and threats, and this alone is sufficient to enable rule egoism to blunt the force of the pragmatic indictment to which egoism is vulnerable. In this way, rule egoism, like rule consequentialism, largely avoids the charge of being self-defeating.

3. The Incoherence Objection

Although rule consequentialism and rule egoism are motivated by legitimate concerns associated with act consequentialism and egoism, they face a serious objection of their own (one that, I'll argue in section 5, Mill seems to have lodged against rule-based ways of thinking about instrumental rationality, even while advocating a rule-based theory of morality). This objection is that such theories are guilty of "rule worship," or are incoherent, because they require agents to comply with certain rules even when violating those rules would have better consequences, in the sense of more effectively promoting whatever objectives are the ones in terms of which those rules are identified as optimal or ideal in the first place. For example, suppose rule consequentialism prohibits the punishment of innocent people. (Suppose, in other words, that a rule prohibiting the punishment of innocent people is part of the optimal system of rules.) But suppose, furthermore, that a case were to arise in which an agent (such as a judge) could either punish an innocent person or not and knew that better consequences would actually result from doing so than from refraining. In such a case, the fact that punishing an innocent person would be optimal would not change the fact that rule consequentialism would require the judge to refrain from doing so—for the insistence that optimal acts are not always right is what keeps rule consequentialism from collapsing into act consequentialism. But the rule consequentialist's insistence that the rule must be followed, despite the consequences, lacks a clear rationale. After all, the

production of the best available consequences is the whole point of the rules that the rule consequentialist endorses. Cases can arise in which that objective—the production of the best consequences—can be served better by breaking a rule than by complying with it, and rule consequentialism seems to have no resources for justifying compliance with the rule in such cases. As a result, the term “rule worship” has struck many critics as an apt denigration of this aspect of the rule-consequentialist position.¹⁶

The same objection applies, with equal force, to rule egoism. In a case in which an agent can advance her interests more by deviating from her optimal system of rules than by acting in compliance with it, the rule-egoist perspective contains no apparent resources for explaining why she should, in fact, comply with it. For example, to return to the case of the farmer discussed above, suppose he has been able to sincerely assure the second farmer that he will reciprocate if he helps him harvest his crops. (He might have been able to do this because he accepted rule egoism as a theory of instrumental rationality and judged that the optimal system of rules for him would allow him to follow through on that sort of assurance. Or he might have been able to make that sincere assurance because, although he was an egoist, he didn't think things through enough to realize that, when the time came, he would regard the balance of reasons as favoring reneging on his assurance. How he was able to make the sincere assurance does not actually matter.) Now suppose, furthermore, that the first farmer has gotten the second farmer's assistance and is deciding whether to reciprocate. Of course, if he is (still, or newly) a rule egoist, he might not think about this very carefully; he might just apply rule egoism to the case at hand and do what it says. And it may well require him to reciprocate. But if it does, and if that would not advance his interests as much as declining to reciprocate would, then he would be entitled to wonder exactly why it is instrumentally rational to reciprocate. If his being a rule egoist enabled him to get the second farmer's assistance, he needn't wonder what good it did him to be a rule egoist; he might well appreciate that beneficial effect of his having previously been a rule egoist. But that would not dispel his present puzzlement, concerning why he should comply with the rules now. Like that of rule consequentialism, rule egoism's insistence on compliance with rules, even when better consequences would result from the violation of them, seems explicable only in terms of some sort of irrational, blindly worshipful attitude toward rules.

The rule-worship objection is also known as the incoherence objection, and I'll refer to it that way here. Because this objection will be of central importance in what follows, it is worth pausing now to state it rigorously and emphasize some of its essential features. Here is what I regard as the most useful formulation of the incoherence objection for this discussion:

An agent who regards a system of rules *S* as binding for her solely because she believes that it is optimal for the achievement of some aim deliberates and acts irrationally in a particular case if, in that case, she refrains from doing an act *A* that she knows is optimal with respect to that aim (and, so, does some other act) solely because *S* prohibits *A*. Any theory that prescribes such irrationality in deliberation and action is incoherent.

Two features of this objection are worth emphasizing here. First, the scope of this objection is limited to theories that endorse systems of rules solely for consequentialist reasons. Thus, it does not apply to nonconsequentialist rule-based theories, such as a theory that just demands compliance with the Ten Commandments without supplying any further justification than that it is the will of a divine being. Such a theory might be arbitrary or might ultimately be construable as a mere counsel of prudence on the assumption of divine punishment for disobedience, but it is not incoherent in the sense that concerns us here. Furthermore, this objection does not apply to what may currently be the leading rule-consequentialist theory, that of Brad Hooker, since the rules of his moral theory are justified in terms of their coherence with our considered moral judgments rather than in terms of an overarching commitment to the production of the best possible consequences.¹⁷

Second, the scope of this objection is limited to those theories for which there can arise cases in which an agent *knows* that the act required by the theory is suboptimal. Thus, the plausibility of the incoherence objection is entirely compatible with the empirical observation that agents are usually mistaken when they think they have identified cases in which optimal systems of rules require them to perform suboptimal acts. What the incoherence objection says is that if an agent does know that she is in such a situation (according to however rigorous a standard for knowledge one might want to impose), then it is irrational for her to act suboptimally, and incoherent for a theory to require her to do so.

4. Mill's Rule Utilitarianism

As I mentioned above, this paper will explore the tension in Mill's thought between his adherence to a rule-based theory of morality and his endorsement of the incoherence objection against rule-based theories of instrumental rationality. I now want to establish the two elements of Mill's thought just mentioned. To do this, I'll first present, in this section, some evidence in support of the claim that Mill's moral theory is a form of rule utilitarianism; and then, in the next section, I'll turn to Mill's endorsement of the incoherence objection.

Whether Mill's utilitarianism should be interpreted as a form of act utilitarianism or as a form of rule utilitarianism is a question that has been debated in dozens of books and papers for more than fifty years now. The evidence amassed on each side of this debate is extensive and complex, including numerous passages from a wide variety of Mill's writings. Here I do not purport to add anything new to this debate—much less to settle it—but only to briefly mention what I see as three obvious pieces of textual evidence in support of the rule-utilitarian interpretation of Mill. All these passages are in the second chapter of Mill's *Utilitarianism*—arguably the most important chapter of his most important work on morality.

The first passage concerns what Mill says about “the standard of morality.” Assuming this refers to a standard for determining what acts are right and what acts are wrong, the following quotation suggests that Mill regards that standard as involving essential reference to rules:

the standard of morality...may accordingly be defined [as] the rules and precepts for human conduct, by the observance of which an existence such as has been described might be, to the greatest extent possible, secured to all mankind; and not to them only, but, so far as the nature of things admits, to the whole sentient creation.¹⁸

This remark obviously has a strong affinity with the canonical formulation of rule consequentialism given above.

The second passage I want to quote is particularly striking in view of the disagreement between act consequentialism and rule consequentialism in cases of the kind that the incoherence objection refers to: cases in which the agent knows that she can produce better consequences by breaking a rule than by following it. In such cases, act consequentialism requires the agent to break the rule, while rule consequentialism requires the agent to follow it. Here is Mill's remark about such cases:

In the case of abstinences indeed—of things which people forbear to do, from moral considerations, though the consequences in the particular case might be beneficial—it would be unworthy of an intelligent agent not to be consciously aware that the action is of a class which, if practised generally, would be generally injurious, and that this is the ground of the obligation to abstain from it.¹⁹

Obviously, this remark endorses the rule-consequentialist insistence on rule-following even in this contested class of cases.

Finally, in the last paragraph of the second chapter of *Utilitarianism*, Mill again affirms the importance of moral rules. He does so by drawing on his notions of first principles and secondary principles: that morality should be based on happiness is clearly, for Mill, a first principle; ordinary rules of conduct are secondary principles. According to Mill,

We must remember that only in... cases of conflict between secondary principles is it requisite that first principles should be appealed to.²⁰

Like his characterization of the “standard of morality,” this remark suggests that moral requirements are normally dictated by rules, not by the act-utilitarian principle of doing the act that will produce the most happiness—though this latter principle may need to be consulted in certain cases.

To be sure, proponents of interpreting Mill as an act utilitarian have alternative readings of these and other passages. As I mentioned above, my aim here has been not to advance the debate over the interpretation of Mill as an act utilitarian or as a rule utilitarian, but only to briefly mention three familiar passages that are relevant to that debate. I take it that the rule-utilitarian interpretation of Mill is sufficiently well argued elsewhere for Mill’s endorsement of the incoherence objection to be of significant interest.²¹

5. Mill’s Endorsement of the Incoherence Objection

Mill’s endorsement of the incoherence objection occurs in the last chapter of his *A System of Logic*. Early in that chapter, Mill contrasts the proper decision-making modes of judges and legislators in order to illustrate the general point that some activities, when engaged in properly, are more rule-governed than others. A judge—at least in a matter of law in which “there is a definite written code”—“is not called upon to determine what course would be intrinsically the most advisable in the particular case at hand, but only within what rule it falls.”²² Then the judge is just supposed to apply the rule, regardless of whether he or she thinks doing so in that particular case would serve the purpose for which the rule was enacted. The case of the legislator is quite different, Mill writes:

the legislator has rules, and maxims of policy; but it would be a manifest error to suppose that the legislator is bound by these maxims in the same manner as the judge is bound by the laws, and that all he has to do is to argue down from them to the particular case, as the judge does from the laws. The legislator is bound to take into consideration the reasons or grounds of the maxim.²³

Thus, although legislators have rules that rightly influence their deliberations about what laws to make, they should not follow those rules blindly or rigidly. Instead, they should consider the purposes those rules or maxims are supposed to serve and violate them when doing so would serve those purposes better than compliance would. Mill concludes,

the legislator, or other practitioner, who goes by rules rather than by their reasons, like the old-fashioned German tacticians who were vanquished by Napoleon, or the physician who preferred that his patient should die by rule rather than recover contrary to it, is rightly judged to be a mere pedant, and the slave of his formulas.²⁴

Obviously Mill's position resonates strongly with that of the incoherence objection.

Moreover, Mill's references to the military tacticians and the physician can quite straightforwardly be interpreted as references to cases that would trigger the incoherence objection. The military tacticians, we can imagine, regarded a certain system of rules as binding because they believed it to be optimal for the achievement of the aim of military success but continued to follow it even when victory could be achieved only by deviating from it. Similarly, we can imagine that the physician regarded a certain system of rules as binding because he believed it to be medically optimal but refused to deviate from it even when he could see that the path to recovery, in the case of the patient in question, lay elsewhere. Thus, both of Mill's somewhat oblique examples readily lend themselves to being interpreted in ways that align with the incoherence objection.

Two complications might initially seem to stand in the way of interpreting Mill's remarks as an endorsement of the incoherence objection, but further reflection shows both of them to be unproblematic. The first complication concerns Mill's account of the military tacticians. Whereas Mill clearly seems to have in mind a physician who knew himself to be in a situation in which breaking the rule would have better consequences than keeping it, Mill does not explicitly say that the military tacticians were aware that they were in such a situation. Given that the incoherence objection refers only to cases in which the agent knows that she is in such a situation, Mill might seem to be getting at something other than that objection. But two responses should suffice to address this concern. First, Mill evidently intends for the two situations to stand together as examples of the same sort of thing, and if his description of one is silent on a matter on which his description of the other is explicit, it is probably reasonable to interpret the former in light of the latter. Second, and more significant, if Mill's description really is to be understood as neutral on the matter of the tacticians' knowledge, then Mill's condemnation of them commits him to an even bolder claim than the incoherence objection. That objection (as stated earlier) is formulated deliberately cautiously: in order for rule-following to be irrational, it must produce worse consequences than compliance, *and the agent must know this*. If this last condition is omitted, then the revised objection is bolder—criticizing, as irrational, agents whose rule-following is sub-

optimal regardless of whether they know that it is. And if Mill commits himself to this stronger objection, then he is obviously committed to the incoherence objection as well. Moreover, even if we assume that (in Mill's example) the tacticians specifically did *not* know that following their rules would frustrate their rules' purpose, surely Mill would condemn them even more harshly if this fact about their conduct were pointed out to them and they insisted on adhering to their rules anyway. Thus, Mill's silence on the matter of exactly what the tacticians knew about their situation poses no problem for interpreting their situation, like that of the physician, as supporting the claim that Mill endorses the incoherence objection.

The second complication is that Mill's examples do not involve agents who clearly subscribe to a particular theory of morality or a particular theory of instrumental rationality. Instead, the agents are much more naturally seen as regarding, as binding, systems of rules with much narrower scopes (those of military success and medicine, respectively). But this does not mean that Mill's objection to their way of thinking is not essentially an instance of the incoherence objection. Admittedly, there are objections that could arise in such narrower contexts that would not naturally carry over to the broader contexts of morality and instrumental rationality. For example, if Mill had condemned the military tacticians and the physician for pursuing the aims of their professions at the expense of other goods such as economic justice or the preservation of the environment, then his objection would have depended specifically on the narrowness of the tacticians' and physician's respective systems of rules, and his objection would not naturally carry over to other, broader contexts. But his objection does not depend on the narrowness of their systems of rules. Mill's objection is not that the tacticians and the physician were ignoring considerations that fell outside of their too-narrow systems of rules (considerations that might have been adequately captured by a theory of morality or a theory of instrumental rationality). Rather, his objection is that the tacticians and physician were following their rules even in circumstances in which those rules' purposes would have been better served by breaking them. (And then it would have been a separate objection, had Mill been inclined to make it, to say that the tacticians and the physician should have been mindful of other things than success within their perhaps narrowly defined fields.) Given this understanding of Mill's remarks, they are clearly an endorsement of the incoherence objection, and it is hard to see how he could decline to regard parallel considerations as applying equally forcefully to rule-based theories of morality and instrumental morality.

Mill's discussion of legislating is not the only passage in the last chapter of the *Logic* that implies his endorsement of the incoherence objection. On the next page, he abstracts from the specific activities of judging and legis-

lating and reflects more generally on the role of rules in goal-directed activity. He writes that rules of art are “always imperfect,” since they are designed to handle ordinary cases and complicating them so that they can also handle extraordinary cases would render them “too cumbrous.”²⁵ Then, after noting that such imperfect rules may safely be acted on in “the manual arts,” he writes that the rules of more complicated arts must be acted on more cautiously:

[I]n the complicated affairs of life, and still more in those of states and societies, rules cannot be relied on, without constantly referring back to the scientific laws on which they are founded....

By a wise practitioner, therefore, rules of conduct will only be considered as provisional.... [T]hey do not at all supersede the propriety of going through (when circumstances permit) the scientific process requisite for framing a rule from the data of the particular case before us.²⁶

If rules cannot be relied on in complicated affairs without making sure they are appropriate to the case at hand, and if wise practitioners do indeed make sure they are appropriate to the case at hand when circumstances permit (notice Mill’s reference to “the data of the particular case before us”), then presumably it would be nothing short of irrational for an agent to follow a rule when he or she could see that doing so would serve the purpose for which it was framed less effectively than deviating from it would. These remarks, then, provide further support for the claim that Mill endorses the incoherence objection.²⁷

6. Possible Resolutions of This Conflict

In section 4, I reviewed some of the evidence that suggests that Mill’s theory of morality is a form of rule utilitarianism. In section 5, I argued that Mill endorses the incoherence objection. These results obviously appear to be in conflict (barring the attribution to Mill of a Hooker-style rule consequentialism—a possibility I will discuss below), and there must be more to say if this conflict is to be explained or resolved. If we take it as given that Mill endorses the incoherence objection, there must be more to say about his theory of morality, to explain why it is not vulnerable to the incoherence objection. Or, if we take it as given that Mill espouses rule utilitarianism, there must be more to say about his apparent endorsement of the incoherence objection, so that it can be interpreted narrowly enough not to impinge on his view about morality. Or, finally, perhaps Mill’s views really are in conflict, in which case we might then seek some explanation of Mill’s failure to notice and resolve this inconsistency. In this section, I’ll review

five potential explanations and resolutions of this conflict—four that I think are ultimately unpersuasive and one that I think has some promise.

The first is an explanation that might seem promising but that rests on a historical hypothesis that does not pan out. It is well known that Mill's *Logic* went through many editions, from the first in 1843 to the eighth in 1872. Since this latter year postdates Mill's most important ethical writings (*Utilitarianism*, of course, having been published in 1861), it might be wondered whether the passage from the *Logic* discussed above was added in a late edition. Specifically, it might be wondered whether Mill was a late convert to the church of the incoherence objection and never had occasion to indicate how this might require some revision of his moral views. Unfortunately, the textual evidence confounds this proposal: the passage discussed above was in every edition of the *Logic*. Moreover, while nearby passages underwent emendations and redactions, the one discussed above remained essentially unchanged. Thus, not only did Mill commit himself to the incoherence objection early; he also declined seven further opportunities to soften his stance.

The second response to the conflict is a potential resolution that rests on further characterizing what kind of rule utilitarianism Mill subscribes to. Mill holds that morality is essentially a matter of rules, and he holds that happiness provides the standard for evaluating rules. Might he nevertheless subscribe to a form of rule utilitarianism that is not vulnerable to the incoherence objection, because the rules would ultimately be justified on grounds other than the promotion of happiness? Recall, for example, Hooker's rule consequentialism. In that theory, the promotion of good consequences is the standard for evaluating rules, but that fact about his theory has a deeper rationale—namely, that the rules thus identified cohere best with our considered judgments. But Mill does not say enough about (what we now refer to as) considered judgments and reflective equilibrium to attribute to him a view resembling Hooker's. And such a reliance on existing considered judgments would seem to be incompatible with Mill's strong reformist tendencies and his explicit willingness to criticize moral views that were widely and sincerely held among his contemporaries. So it is unlikely that Mill's theory resembles Hooker's in that specific way. Still, might it resemble Hooker's theory in a more general way, by having some deeper rationale (albeit one other than coherence with considered judgments) for establishing the promotion of happiness as the standard for evaluating rules? No clear possibility presents itself: Mill's appeal to the promotion of happiness, as the standard for evaluating rules, seems to be quite direct and fundamental—in precisely the way that triggers the incoherence objection.

The final three responses to the conflict that I'll consider are all potential resolutions that suggest that the general claim that we are to under-

stand Mill to be endorsing in the final chapter of the *Logic* is not actually the incoherence objection, but something narrower—narrow enough, in fact, not to extend to his rule utilitarianism. The first of these three potential resolutions rests on the claim that when Mill criticizes the legislators, he is not necessarily criticizing them just for their adherence to their system of rules. Instead, he might well be criticizing them for their adherence to a system of rules that they are in a position to know is not good enough to deserve their unswerving adherence. After all, there is no reason to think that Mill regarded the art of legislation as so well developed that the best system of rules available to its practitioners should be regarded as flawless; and presumably he would have thought that this fact would be evident to anyone reasonably familiar with that art. On this hypothesis, our understanding of Mill's remarks about legislators needs to be revised. Instead of interpreting those remarks as endorsing the incoherence objection, we should interpret them as endorsing the claim that when one is in a position to appreciate the serious shortcomings of the best system of rules one has available in a particular context, then it is irrational for one to comply with them when they require suboptimal action.

Unfortunately, this potential resolution of the conflict rests on an implausible interpretation of Mill's remarks about the role of rules in legislation. When Mill contrasts legislation and adjudication, he makes no mention of any difference in the excellence of those activities' respective systems of rules—there is no suggestion that the reason why judges are rightly more rule-bound than legislators is that the rules of their art are in a more advanced state. Moreover, when he likens rule-bound legislators to the feckless German tacticians and the foolish physician, he does not complain of flaws in the systems of rules those agents consult. Instead, the basis of his criticism is his description of them as “go[ing] by their rules rather than by their reasons.” Since this phrase would still be applicable regardless of how good their rules were, the suggestion that Mill's criticisms really stem from the inadequacy of the legislators' rules is, ultimately, unsupported.

I'll turn now to the final unsuccessful response to the conflict that I think merits our attention. Like the response just discussed, this one suggests that the general principle underlying Mill's remarks in the final chapter of the *Logic* is narrower than the incoherence objection. It does so by suggesting that Mill holds that the moral department of the Art of Life differs from that of prudence or instrumental rationality not only in its content but also in its canons of reasoning. That is, Mill might think that a certain kind of rule-governed reasoning that would be fallacious in ordinary goal-oriented contexts is valid in the moral context.

It is hard to see how this might be the case, however. Perhaps the most obvious possibility for fleshing out this vague suggestion is that since

morality is, by definition, a matter of rules, then there is no further thinking to be done except to follow the rules. But even if this appeal to the definition of morality were correct, it would not be very plausible. For if an agent were to find herself in a situation in which the maximization of happiness required breaking a moral rule, and it were pointed out to her that morality is a domain in which one's reasoning is supposed to be rule-governed, she could still reply along the lines of the incoherence objection. That is, she could point out that morality, as Mill conceives of it, is an institution with a purpose—the promotion of happiness—and even if morality (due to its rule-based character) does, indeed, require her to comply with the rule, the underlying purpose of morality itself, and the rules it comprises, would be better served by breaking the rule. Thus, she could conclude, any supporters of the whole institution of morality should, if they support it because of the purpose it serves, applaud any deviation from that institution that serves its underlying purpose more effectively than compliance does. Thus, the impact of the incoherence objection against rule-governed deliberation cannot be evaded by elevating rules into an essential component of the institution of morality.

The fifth and final response to the conflict that I'll consider here is a potential resolution that I think has some promise. It rests on a version, with a Millian theme, of the argument that because of the difficulty of accurately predicting the consequences of one's actions, one is never justified in believing that one needs to deviate from the optimal system of rules in order to produce the best available consequences. In the present context, this argument begins with the simple observation that morality is an exceptionally complex enterprise. Although its overall goal of maximizing happiness is straightforward enough, its rules must encompass and speak to an extremely diverse range of possible actions that agents might undertake. Moreover, the laws of cause and effect that must go into the accurate formulation of its rules are also drawn from an extremely diverse range of disciplines. Consider, by way of contrast, the rules of medicine. If we characterize medicine as the art of promoting health, then its rules can be formulated pretty much without reference to any laws of cause and effect except those that are concerned with the promotion (or impairment) of health. But since morality is concerned with the promotion of happiness, and happiness can be promoted (or impaired) by phenomena that are also the province of medicine or warfare or legislation or building or any number of other activities, morality will turn out to be an exceptionally complex art—even compared to most other "complicated affairs of life" (to use Mill's phrase, quoted above). Given this, Mill might regard morality as exceptional in another respect: Mill might think that it cannot reasonably be expected that any agent will

ever be in a situation in which she is in a position to know that the optimal system of rules requires her to perform an act whose consequences are not as good as those of some other act that she could perform instead. And if Mill thinks this, then he must think that no case of the kind to which the incoherence objection refers can reasonably be expected to arise in the realm of morality. Such a case can be expected to arise in arts of narrower scope, such as legislation, warfare, or medicine, but maximizing happiness is an activity of a different order of complexity and unpredictability.

This, of course, is not yet enough to resolve the conflict, since the incoherence objection does not depend, for its applicability, on the reasonable expectability of cases of the kind to which it refers—it is enough, to trigger the incoherence objection, that cases of that kind could occur. But it would be easy to formulate a slightly weaker variant of the incoherence objection that does depend, for its applicability, on the reasonable expectability of such cases. It could be expressed, for example, by changing the final sentence of the incoherence objection to say, “Any theory that prescribes such irrationality in deliberation and action, *in cases that can reasonably be expected to occur*, is incoherent.” And a principle of this sort would explain Mill’s remarks in the final chapter of the *Logic* without extending as far as his rule utilitarianism. In sum, if we attribute to Mill (1) the argument about the difficulty of accurately predicting the happiness-affecting consequences of our actions described in the previous paragraph, (2) the denial of the incoherence objection, and (3) the acceptance of the weaker variant of the incoherence objection just described, then we can make sense of both (4) his remarks in the final chapter of the *Logic* and (5) his espousal of rule utilitarianism as the correct theory of morality.

As I said, I think this line of thought has some promise. Obviously, though, further investigation would be necessary in order to comprehensively assess the degree of textual support that can be found for it. Furthermore, there is no reason to rule out the possibility that some other, even more promising, potential resolution could be imagined. The problem is a difficult one because of how fully the conflicting views are developed in Mill’s writings. As I indicated in section 4, the evidence in support of Mill’s adherence to rule utilitarianism is strong. And as I indicated in section 5, Mill’s remarks in the final chapter of the *Logic* demonstrate that he was quite sensitive to the concerns about compliance with rules that are expressed by the incoherence objection. Despite the promise of the last of the potential resolutions described above, this conflict remains a puzzling aspect of Mill’s thought that merits further study because of the centrality of the role of rules in Mill’s thinking about morality and goal-directed enterprises more generally.

NOTES

I received helpful comments on earlier versions of this chapter from several people. I would particularly like to thank Dale Miller, David Weinstein, and audience members at the 2006 conference of the International Society for Utilitarian Studies at University College London.

1. Exactly what makes a theory consequentialist or not is a surprisingly vexing question, and one that I shall not try to explore here. For some discussions of it, see Stocker, "Consequentialism and Its Complexities," 276; Parfit, *Reasons and Persons*, 26–27; Slote, "Satisficing Consequentialism," 140–44, and *Common-Sense Morality and Consequentialism*, 35–39; Griffin, "The Human Good and the Ambitions of Consequentialism," 118–20 and 125; and Shaw, *Contemporary Ethics: Taking Account of Utilitarianism*, 12 and 75.

2. See, for example, Scheffler, *Rejection of Consequentialism*, 1; Frey, "Act-Utilitarianism," 165; and Sinnott-Armstrong, "Consequentialism."

3. See, for example, Brandt, "Some Merits of One Form of Rule-Utilitarianism," 217 and 300; Hooker, *Ideal Code, Real World*, 32; and Sinnott-Armstrong, "Consequentialism."

4. Similar principles are formulated by Parfit (*Reasons and Persons*, 4) and Robert Shaver (*Rational Egoism*, 2).

5. Similar principles are formulated by Moore ("Hobbes on Obligation, Moral and Political; Part One: Moral Obligation," 45 and 48), Brandt ("Rationality, Egoism, and Morality," 691), and Kavka (*Hobbesian Moral and Political Theory*, 358–59). But these writers discuss rule-egoistic moral theories, as opposed to rule-egoistic theories of instrumental rationality. Perhaps the most sophisticated theories in the spirit of rule egoism are developed in the work of David Gauthier and Edward F. McClennen. In this paper I'll rely on Gauthier's paper "Assure and Threaten" as showing the rule-egoistic approach to instrumental rationality in its best light.

6. For previous statements of this point, see Rescher, *A Constructive Critique of the Utilitarian Theory of Distribution*, 48; Hare, "Ethical Theory and Utilitarianism," 222; Frey, "Act-Utilitarianism: Sidgwick or Bentham and Smart?" 99; and Allan Gibbard, "Inchoately Utilitarian Common Sense: The Bearing of a Thesis of Sidgwick's on Moral Theory," 76.

7. See, for example, Hodgson, *Consequences of Utilitarianism*, 38–62; Stocker, "The Schizophrenia of Modern Ethical Theories," 461; and Parfit, *Reasons and Persons*, 27–28.

8. But note the recurring theme of the paradox of happiness—according to which the best route to happiness lies, paradoxically, in aiming at things other than happiness. See, for example, Mill, *Autobiography*, 117; Hodgson, *Consequences of Utilitarianism*, 61; Stocker, "Schizophrenia of Modern Ethical Theories," 456; Elster, *Sour Grapes*, 9, and *Ulysses and the Sirens*, 40; and Railton, "Alienation, Consequentialism, and the Demands of Morality," 140–41.

9. This example closely follows one that Gauthier ("Assure and Threaten," 692–93) adapts from Hume.

10. Gauthier, "Assure and Threaten," 710–11.

11. Kavka, "The Toxin Puzzle."

12. Gauthier, "In the Neighbourhood of the Newcomb Predictor."

13. Gauthier, "Assure and Threaten," 702.

14. *Ibid.*, 701.

15. *Ibid.*, 703–7 and 715–19.

16. Authors objecting to rule-consequentialist theories on this basis (though not all of them using the term "rule worship") include Smart ("Extreme and Restricted Utilitarianism," 348–49 and 353), Sprigge ("A Utilitarian Reply to Dr. McCloskey," 286–87, and *The Rational Foundations of Ethics*, 26), Williams (*Morality: An Introduction to Ethics*, 94), Nozick (*Anarchy, State, and Utopia*, 30), Brandt (*A Theory of the Good and the Right*, 296, and *Facts, Values, and Morality*, 151), Lyons ("Utility as a Possible Ground for Rights," 25–26), Scheffler (*Rejection of Consequentialism*, 82), Kagan (*The Limits of Morality*, 33 and 37), Blackburn (*Ruling Passions*, 38), and Hooker (*Ideal Code, Real World*, 99–100). For discussion of this point particularly in regard to the problem of punishing the innocent, see Boonin (*The Problem of Punishment*, 70–75).

17. Hooker, *Ideal Code, Real World*, 101–2 and 188.

18. Mill, *Utilitarianism*, 214.

19. *Ibid.*, 220.

20. *Ibid.*, 226.

21. See, for example, Urmson, "The Interpretation of the Moral Philosophy of J. S. Mill," 35; Brandt, "Some Merits of One Form of Rule-Utilitarianism," 57–58; Ryan, "Introduction," 56–58; MacIntyre, *A Short History of Ethics*, 241; Fuchs, "Mill's Theory of Morally Correct Action," 146; Eggleston and Miller, "India House Utilitarianism: A First Look," 42; and Martin, "Mill's Rule Utilitarianism in Context," 24.

22. Mill, *System of Logic*, 944.

23. *Ibid.*, 944.

24. *Ibid.*, 944.

25. *Ibid.*, 945.

26. *Ibid.*, 945–46.

27. Still further confirmation is suggested by Mill's intriguing remark, in *The Subjection of Women*, that "women are comparatively unlikely to fall into the common error of men, that of sticking to their rules in a case whose specialties either take it out of the class to which the rules are applicable, or require a special adaptation of them" (307).

BIBLIOGRAPHY

- Blackburn, Simon. *Ruling Passions: A Theory of Practical Reasoning*. Oxford: Oxford University Press, 1998.
- Boonin, David. *The Problem of Punishment*. Cambridge: Cambridge University Press, 2008.
- Brandt, Richard B. *Facts, Values, and Morality*. Cambridge: Cambridge University Press, 1996.
- . "Rationality, Egoism, and Morality." *The Journal of Philosophy* 69, no. 20 (November 9, 1972): 681–97.
- . "Some Merits of One Form of Rule-Utilitarianism." In *University of Colorado Studies, Series in Philosophy*, no. 3: *The Concept of Morality*, 39–65. Boulder: University of Colorado Press, 1967.
- . *A Theory of the Good and the Right*. Oxford: Oxford University Press, 1979.
- Eggleston, Ben, and Dale E. Miller. "India House Utilitarianism: A First Look." *Southwest Philosophy Review* 23, no. 1 (January 2007): 39–47.
- Elster, Jon. *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge: Cambridge University Press, 1983.
- . *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Rev. ed. Cambridge: Cambridge University Press, 1984.

- Frey, R. G. "Act-Utilitarianism." In *The Blackwell Guide to Ethical Theory*, edited by Hugh LaFollette, 165–82. Oxford: Blackwell Publishing, 2000.
- . "Act-Utilitarianism: Sidgwick or Bentham and Smart?" *Mind* 86, no. 341 (January 1977): 95–100.
- Fuchs, Alan E. "Mill's Theory of Morally Correct Action." In *The Blackwell Guide to Mill's Utilitarianism*, edited by Henry R. West, 139–58. Oxford: Blackwell Publishing, 2006.
- Gauthier, David. "Assure and Threaten." *Ethics* 104, no. 4 (July 1994): 690–721.
- . "In the Neighbourhood of the Newcomb Predictor (Reflections on Rationality)." *Proceedings of the Aristotelian Society* 89 (1989): 179–94.
- Gibbard, Allan. "Inchoately Utilitarian Common Sense: The Bearing of a Thesis of Sidgwick's on Moral Theory." In *The Limits of Utilitarianism*, edited by Harlan B. Miller and William H. Williams, 71–85. Minneapolis: University of Minnesota Press, 1982.
- Griffin, James. "The Human Good and the Ambitions of Consequentialism." *Social Philosophy & Policy* 9, no. 2 (Summer 1992): 118–32.
- Hare, R. M. "Ethical Theory and Utilitarianism." Chap. 13 in *Essays in Ethical Theory*. Oxford: Oxford University Press, 1989.
- Hodgson, D. H. *Consequences of Utilitarianism: A Study in Normative Ethics and Legal Theory*. Oxford: Oxford University Press, 1967.
- Hooker, Brad. *Ideal Code, Real World: A Rule-consequentialist Theory of Morality*. Oxford: Oxford University Press, 2000.
- Kagan, Shelly. *The Limits of Morality*. Oxford: Oxford University Press, 1989.
- Kavka, Gregory S. *Hobbesian Moral and Political Theory*. Princeton, N.J.: Princeton University Press, 1986.
- . "The Toxin Puzzle." *Analysis* 43, no. 1 (January 1983): 33–36.
- Lyons, David. "Utility as a Possible Ground for Rights." *Noûs* 14, no. 1 (March 1980): 17–28.
- MacIntyre, Alasdair. *A Short History of Ethics: A History of Moral Philosophy from the Homeric Age to the Twentieth Century*. 2nd ed. Notre Dame: University of Notre Dame Press, 1998.
- Martin, Rex. "Mill's Rule Utilitarianism in Context." In *John Stuart Mill and the Art of Life*, edited by Ben Eggleston, Dale E. Miller, and David Weinstein, 21–43. Oxford: Oxford University Press, 2010.
- McClennen, Edward F. "Pragmatic Rationality and Rules." *Philosophy & Public Affairs* 26, no. 3 (Summer 1997): 210–58.
- Mill, John Stuart. *Autobiography*. In *The Collected Works of John Stuart Mill*, vol. 1, edited by John M. Robson and Jack Stillinger, 1–290. Toronto: University of Toronto Press, 1981.
- . *The Subjection of Women*. In *The Collected Works of John Stuart Mill*, vol. 21, edited by John M. Robson, 259–340. Toronto: University of Toronto Press, 1984.
- . *A System of Logic Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. In *The Collected Works of John Stuart Mill*, vols. 7–8, edited by J. M. Robson. Toronto: University of Toronto Press, 1974.
- . *Utilitarianism*. In *The Collected Works of John Stuart Mill*, vol. 10, edited by J. M. Robson, 203–59. Toronto: University of Toronto Press, 1969.
- Moore, Stanley. "Hobbes on Obligation, Moral and Political; Part One: Moral Obligation." *Journal of the History of Philosophy* 9, no. 1 (January 1971): 43–62.
- Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Railton, Peter. "Alienation, Consequentialism, and the Demands of Morality." *Philosophy & Public Affairs* 13, no. 2 (Spring 1984): 134–71.
- Rescher, Nicholas. *A Constructive Critique of the Utilitarian Theory of Distribution*. Indianapolis: The Bobbs-Merrill Company, Inc., 1966.
- Ryan, Alan. "Introduction." In *Utilitarianism and Other Essays*, edited by Alan Ryan, 7–63. London: Penguin Books, 1987.
- Scheffler, Samuel. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. Oxford: Oxford University Press, 1982.
- Shaver, Robert. *Rational Egoism: A Selective and Critical History*. Cambridge: Cambridge University Press, 1998.
- Shaw, William H. *Contemporary Ethics: Taking Account of Utilitarianism*. Oxford: Blackwell Publishers, 1999.
- Sinnott-Armstrong, Walter. "Consequentialism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta (Fall 2008 edition), <<http://plato.stanford.edu/archives/fall2008/entries/consequentialism/>>).
- Slote, Michael. *Common-Sense Morality and Consequentialism*. London: Routledge & Kegan Paul, 1985.
- . "Satisficing Consequentialism." Part I. *The Aristotelian Society* supplementary volume 58 (1984): 139–63.
- Smart, J. J. C. "Extreme and Restricted Utilitarianism." *Philosophical Quarterly* 6, no. 25 (October 1956): 344–54.
- Sprigge, T. L. S. *The Rational Foundations of Ethics*. London: Routledge & Kegan Paul, 1988.
- . "A Utilitarian Reply to Dr. McCloskey." *Inquiry* 8, no. 3 (Autumn 1965): 264–91.
- Stocker, Michael. "Consequentialism and Its Complexities." *American Philosophical Quarterly* 6, no. 4 (October 1969): 276–89.
- . "The Schizophrenia of Modern Ethical Theories." *The Journal of Philosophy* 73, no. 14 (August 12, 1976): 453–66.
- Urmson, J. O. "The Interpretation of the Moral Philosophy of J. S. Mill." *The Philosophical Quarterly* 3, no. 10 (January 1953): 33–39.
- Williams, Bernard. *Morality: An Introduction to Ethics*. Cambridge: Cambridge University Press, 1972.