Georgia State University
ScholarWorks @ Georgia State University

Philosophy Theses | Department of Philosophy

1-7-2010

Subjectivity and Fallibility in the Instrumental and Epistemic Defenses of a "Right to Do Wrong"

Thomas Wright
Georgia State University

Follow this and additional works at: http://scholarworks.gsu.edu/philosophy_theses

Recommended Citation
Wright, Thomas, "Subjectivity and Fallibility in the Instrumental and Epistemic Defenses of a "Right to Do Wrong"." Thesis, Georgia State University, 2010.
http://scholarworks.gsu.edu/philosophy_theses/67

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

SUBJECTIVITY AND FALLIBILITY IN THE INSTRUMENTAL AND EPISTEMIC DEFENSES OF "A RIGHT TO DO WRONG"

by

THOMAS WRIGHT

Under the Direction of William Edmundson

ABSTRACT

An instrumental defense of a right to do wrong is plausible because we cannot directly intervene in an individual's choices so as to effectively promote that individual's moral good, if her moral good is conceived as being some form of individual autonomy. An epistemic defense is also plausible if we reorient J.S. Mill's epistemological argument for his Harm Principle in "On Liberty" to center on the agent's knowledge, rather than on the interfering observer's knowledge. Restrictions on harmless acts that are imposed because the acts are wrong are only justifiable to that individual if *she herself knows* that her acts are wrong. Both approaches depend upon the limited subjectivity and fallibility of the agent or interfering observer. Moreover, both approaches make the justification for a right to knowingly do wrong problematic.

INDEX WORDS: Instrumental defense of rights, Epistemic, Epistemological defense, Fallibility, On liberty, Moral rights, Right to do wrong, Rule utilitarianism, Waldron, Mill

SUBJECTIVITY AND FALLIBILITY IN THE INSTRUMENTAL AND EPISTEMIC DEFENSES OF "A RIGHT TO DO WRONG"

by

THOMAS WRIGHT

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2009

Copyright by

Thomas Wright

2009

SUBJECTIVITY AND FALLIBILITY IN THE INSTRUMENTAL AND EPISTEMIC DEFENSES OF "A RIGHT TO DO WRONG"

by

THOMAS WRIGHT

Committee Chair: William Edmundson

Committee: William Edmundson

George Rainbolt

Peter Lindsay

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
December 2009

TABLE OF CONTENTS

1 INTRODUCTION 1

2 AN INSTRUMENTAL JUSTIFICATION FOR A RIGHT TO DO WRONG 3

2.1 WALDRON'S JUSTIFICATION FOR A RIGHT TO DO WRONG 3

2.2 AN INSTRUMENTAL JUSTIFICATION FOR THE GENERALITY OF MORAL RIGHTS 16

2.3 SOME DISTINCTIONS FROM RULE-UTILITARIAN ACCOUNTS OF MORAL RIGHTS 27

2.4 A RIGHT TO KNOWINGLY DO WRONG 32

2.4.1 CONCEPTUAL ANALYSIS 33

2.4.2 INSTRUMENTAL JUSTIFICATION 37

3 AN EPISTEMIC JUSTIFICATION FOR A RIGHT TO DO WRONG 41

3.1 TRUTH AND FALLIBILITY 46

3.2 JUSTIFICATION AND PROGRESS 51

3.3 THE AGENT'S KNOWLEDGE 57

3.4 FREEDOM 60

3.5 A RIGHT TO KNOWINGLY DO WRONG ON THE EPISTEMIC ACCOUNT 64

ENDNOTES 67

1 INTRODUCTION

Jeremy Waldron[1] identifies an intuition within the domain of moral rights that poses a puzzle for philosophers who believe that there is a deep connection between rights and what is morally right, or that moral rights should protect only what is morally right. When confronted with what seems to be a wrong act, observers of the act will sometimes judge that "act X is wrong..." but then immediately qualify that judgment with a statement like "but you [the agent] have a right to do it." That conjunction leads to what appears to be an intuition of a moral right to do moral wrong, which seems illogical and paradoxical. One also often hears puzzling statements like "Your speech is wrong and offensive, but I will defend your right to say it." Waldron adduces several examples that raise the intuition of this apparent "right to do wrong" (RTDW): a lottery winner gives nothing to charity; an individual works in support of an organization with racist leanings; a political campaign worker manipulates simpleminded voters; an athlete participates in a race-segregated competition; a man is rude in a railway car. In each case, and in many others like them, the individual agent does something that seems to be wrong while also appearing to have a right to do the act free of anyone's interference.

I will consider two ways in which such a RTDW might be justified. First, I will critically examine Waldron's specific efforts to provide a rationale for a RTDW. Waldron's arguments in defense of a RTDW fail, but his observations on how rights-claims are justified in general suggest an instrumental defense of a RTDW that may succeed. Second, I will examine J.S.Mill's epistemic argument in <u>On Liberty</u> for his Harm Principle, which by its terms may

generate a RTDW. Again, while Mill's specific epistemic argument for the Harm Principle largely fails, the failure suggests an epistemic rationale for a RTDW that may succeed. Both the instrumental defense and the epistemic defense ultimately rest on a respect for the individual's agency and capacity for reflection and self-determination, and both emphasize the individual's freedom to make her own ethically progressive choices. The instrumental approach seeks to protect and promote the individual's freedom to direct her own life; the epistemic approach insists that any interferences with her freedom be ones that she would, at least in theory, rationally accept. Both yield a limited RTDW: the individual has a RTDW, so long as she does not knowingly commit wrong acts. I will apply each rationale to the case of knowingly doing wrong.

The instrumental defense of a RTDW begins with the realization that we cannot directly intervene in an individual's choices so as to effectively promote that individual's moral good, if her moral good is conceived as being some form of individual autonomy. We are fallible and biased, and our direct interventions in an individual's choices are as likely to frustrate as they are to promote the individual's moral autonomy. To mitigate our fallibility and bias, therefore, we should instead set up a regime of general moral rights that identifies general classes of choices that are important to individual autonomy, and in this way indirectly promote individual autonomy. But since this approach to promoting individual autonomy is indirect, there will often be cases where a general moral right protects choices to commit acts that are morally wrong for the individual to do. Individuals would then sometimes have a moral right to choose to commit wrong acts, which is a RTDW. A RTDW, then, is a necessary and perhaps unfortunate result of the best instrumental strategy for promoting individual autonomy.

The epistemic defense of a RTDW may take two forms. Firstly, if our justification for interfering with an individual's (harmlessly) wrong act is that the act is wrong, then we should indeed *know* that the act is wrong. If we do not actually know that the act is wrong, either because our belief that the act is wrong is not properly justified, or is not true, or is held out of piety or rote habit rather than because we believe it is true, then our patent justification (and motivation) for interfering with the individual's putatively wrong act fails. This first epistemic strategy for defending a RTDW, however, is less promising that it seems, since in general we may be relatively confident that we do have some genuine knowledge of what is morally right or wrong to do, even for other people. The second epistemic strategy is more persuasive: We invoke the general and fundamentally liberal principle that any restriction on the freedom of an individual must be justifiable to that individual. That principle requires that if we interfere with an individual's act because it is wrong, then in order to justify our interference to that individual, the individual must herself know that the act is wrong. If the individual does not *know* that her act is wrong, either because she does not understand why it is wrong, or because it is not in fact wrong, or because she does not care if it is wrong or not, or for some other reason, then our interference with her act on the basis of the act's wrongness is not justified *to her*. The individual has a RTDW, then, when she chooses to do wrong acts, so long as she herself does not know they are wrong.

2 AN INSTRUMENTAL JUSTIFICATION FOR A RIGHT TO DO WRONG

2.1 WALDRON'S JUSTIFICATION FOR A RIGHT TO DO WRONG

Waldron's justification for "a right to do wrong" (a RTDW) is not a directly instrumental one. Waldron does not argue that an agent should have the right to perform wrong acts because performing them spurs the agent's moral growth or best leads the agent towards moral autonomy, perhaps on the rationale that the agent learns best from making mistakes.[2] People may or may not learn best by making mistakes, and protecting them in making mistakes may lead them to autonomy, but that is not the structure of Waldron's justification for a RTDW. Waldron argues, instead, that 1) rights are characteristically "general" in the sense that they cover broad ranges of protected action that include acts that are morally right, acts without moral import, and acts that are morally wrong; and 2) rights are justified *at this general level* by how they promote an individual's personal integrity or self-definition, or what I will henceforth term individual "autonomy." Waldron does therefore make an instrumental argument to justify *moral rights* at the general level - although alternative ways of justifying moral rights may remain open to him - but he makes no direct instrumental argument to justify committing *wrong acts*.

A RTDW is simply the paradox that seems to result when wrong acts sometimes fall under the protections of moral rights, or the apparent paradox in saying that "A is wrong ... but you have a right to do it."[3] A RTDW is therefore not a general moral right; indeed, a RTDW construed as a moral right, rather than as a paradoxical-seeming consequence of the generality of moral rights, is probably logically incoherent, as well as unjustifiable. For such a RTDW would then seem to mean "the moral right to do acts that are wrong," and if having a moral right implies that interference with its exercise is morally wrong[4], then because that interference is itself

morally wrong, that interference would be protected under the general moral right to do wrong, and "the" RTDW dissolves in self-contradiction. If we adjust this general RTDW as we must to make it coherent, by restricting it to mean "the moral right to do some but not all acts that are morally wrong," and specifically exclude moral wrongs that result from interfering with moral rights from the set of wrong acts protected by the general RTDW, then we have solved the logical problem at the expense of the problem of justifying the RTDW. For to justify such a general RTDW, we would then have to specify the set of wrong acts protected by the RTDW, and determine why it protects that set of wrong acts but not others, and how the protected wrong acts connect to the integrity/self-definition/autonomy rationale usually invoked to justify moral rights. It is possible that a plausible way to make that connection is to argue that sometimes committing wrong acts serves moral autonomy when agents learn from their mistakes. But how are to we to uniquely identify the wrong acts that lead to moral learning, so as to pick out and protect only those wrong acts in the RTDW? It would seem that almost any wrong act may lead to moral learning, or may on the other hand just as likely lead to moral heteronomy and confusion. "The right to do wrong acts that lead to moral learning" seems much too ambiguous to function properly as a general moral right. Moreover, learning from moral mistakes doesn't seem necessary to achieving moral autonomy, as coercion into doing right acts may in fact be the best way to guide an agent towards moral autonomy understood as doing the right thing for the right reasons.[5] Regardless, Waldron never defends or needs to defend such an unlikely general moral right as "the" RTDW.

Waldron does defend a liberal ideal of bold self-determination, self-definition, and personal

integrity,[6] and he defends moral rights by showing how the protections they afford important life-choices serve that liberal ideal. A particular wrong act may disserve that ideal but nonetheless fall under the protection of a general moral right that as a rule does serve it. For example, the traditional liberal moral right of free expression may protect people from interference in freely communicating their thoughts and ideas with others because such expressive communication is an important way for people to discover and realize truth, which allows them to better direct their own lives. When a person wrongly advocates a racist political doctrine, for example, and then perhaps joins a racist organization and shuts out other points of view, then probably these acts do not help her achieve individual moral autonomy. Nonetheless her speech is communicative and free and violates no one's rights;[7] and therefore, her speech is probably a valid instance of the right to free expression, which *does* help people to realize truth and autonomy. Her racist speech and associations are probably wrong and even damaging to her moral autonomy, but because the right of free expression as a general rule serves her autonomy, it would be morally wrong to interfere with her speech, a RTDW.

The form of Waldron's argument recalls that of rule or "indirect" utilitarianism, which may be no accident, as Waldron may be extracting elements of a rule-utilitarian argument for moral rights. Utilitarians are often called upon to defend the principle of utility in the face of notorious counterexamples to its apparent application, such as murdering or torturing one person in order to save many others. Rule-utilitarians characteristically respond to such counterintuitive outcomes by restricting the utilitarian calculus to a set of utility-maximizing rules insulated from the particular acts proposed in the counterexamples. Rule or indirect

utilitarianism then holds that following these rules - which are presumed to largely converge upon our familiar moral intuitions - rather than trying to determine in each case how best to directly maximize utility, is in fact the best way to maximize utility in the long run because of various considerations related to human limitations in acting perfectly morally-rationally. Humans cannot possibly calculate every potential effect of their acts (or perhaps such effects are intrinsically impossible to calculate because of their complexity) and so they need general rules to guide them; and moreover humans are not perfectly morally disciplined; they are weak; and without bright, clear rules, fewer people overall will do what generally tends to promote utility. The utility-maximizing moral rules are learned over time and embedded in culture, and taught in the nursery, and so do not need to be discovered and tested by each individual, as they are already known to us as our moral intuitions.

But because the rules derive their moral authority ultimately from the principle of utility, they are therefore necessarily theoretically defeasible in exceptional cases, and the moral permissibility of rejecting, revising or respecifying them - indeed the moral requirement to do so - is always a possibility. Hence an intransigent critic may insist that a particular (counterintuitive) case calls for a change or modification to the rules by appealing directly to the principle of utility. Such a critic may observe, for example, that perhaps we should make an exception to the rule against torture or murder when that torture or murder would prevent the death of a large number of people, since the unrefined rule against murder as applied in such cases would not maximize utility. The point of such criticism is more to object to the repulsive callousness of the utilitarian procedure for determining what is moral than it is to make any particular moral claim as a counterexample. To consider trading one life off for

another in the utilitarian a way is inherently immoral and corrupting to many such critics;[8] others dispute utilitarian monism and quantitative value commensurability, believing that morality is instead pluralistic in such a way as to sometimes or often generate intractable moral conflicts in which one does wrong no matter what one chooses.[9] Indeed much of the motivation for rights stems from objections to the way the utilitarian procedure countenances weighing and trading off important interests of individuals as a way of resolving moral conflict,[10] a problem that rule utilitarianism ameliorates but cannot completely avoid.[11]

Just as rule utilitarianism depends upon insulating its general rules from a direct application of the principle of utility, so Waldron's instrumental defense of a RTDW depends upon insulating general moral rights from a direct consideration of whether individual autonomy (or integrity or self-definition) is served or not in a particular instance. Rather than subjecting each individual act to a consideration of whether choosing that particular act serves individual autonomy, we identify certain classes of action in which individual autonomy is critically at issue and protect those classes of action, which are moral rights. Within these general classes will inevitably exist certain acts that are morally wrong. But Waldron tends only to *describe* how justification by moral rights in particular cases "usually proceeds,"[12] rather than offering an explicit argument to defend this partitioning of the justification of moral rights from the justification of the acts themselves. "Characteristically," he says, when considering whether a particular act is protected by a moral right, the act is "usually supported by indicating that [the act] A is a member of a certain set of actions any of which [a person] P has a right to perform in the circumstances."[13] And: "the general claim is not inferred ... from the more specific claims," but instead "we establish the general claim and then derive the more specific propositions from

it."[14] Waldron does observe that when disputing whether a moral right exists at the general level, "what is defended or contested ... is the claim that choice within a certain range is not to be interfered with,"[15] and that what is at issue is whether choice in that range is important to individual autonomy. Rights have a "crucial link with the notions of choice and alternatives."[16] But this observation is not an argument for why we should not ask at the level of the particular choice whether this particular chosen act serves individual autonomy or not, and so bypass the question as to whether the choice is an instance of a general moral right, if it is true that moral rights exist to promote autonomy. If general moral rights are ultimately justified by reference to their connection to individual autonomy, then why shouldn't we appeal directly to that value when determining whether to interfere in a particular case, rather than engage in this roundabout reasoning? We would then simply ask ourselves, for example: does choosing to advocate a racist political doctrine help a person become morally autonomous? - no, it probably does not - therefore we may (or even should) interfere.

Waldron offers no explicit answer to this question, though he may implicitly assume that the link he sees between choice and autonomy provides a justification for the generality feature of rights, and for the insulation he describes between the instrumental justification of moral rights by autonomy and the justification of particular choices as instances of those moral rights. But this assumption is ill-conceived. It may be true that autonomy requires choice.[17] Moreover, there is a sense in which any choice is "general." A choice by definition must offer at least one option and an alternative; for example, one may join a racist group, or not join it. If one could not choose to join it, or could not choose *not* to join it, then one has no *choice* whether to join the racist group or not. And if there are other groups one may or may not join, then

one has a broader or in a sense more general choice of which groups to join and whether to join them. One might think, then, to build up a right like that of free expression or association by additively summing together more and more options: the choice to say this or that or the other, or not, or to join this group or that group or another group, or not, and so on. So then the generality of the right to free expression or association would be no different from the generality of a choice between at least two options (like joining a racist organization or not), except that the right of free expression or association would be a choice between a massive number of options.[18] Since the ideal of individual autonomy requires choice, autonomy therefore requires the generality that choices have, and that is what rights protect.

But this form of additive or iterative generality does not capture the generality of a moral right like that of free expression or association, or how such a right's generality connects to the ideal of individual autonomy. For the generality of a right that matters for purposes of autonomy, it is not primarily the breadth of a choice that matters, but the significance of the options available. A choice made between a massive range of banal, uncontroversial political remarks, though more general in the additive sense, does not matter to individual autonomy as much as a single choice to criticize the ruling power or not. Waldron observes in a slightly different context that morally neutral choices, which are often trivial in this way, are those "least likely to be regarded as an appropriate subject matter for rights."[19] Moreover, the iterative generality of a choice does not give any reason for why rights should protect wrong options. While it is true that adding wrong options to a choice would add options and therefore make it more general, there is no obvious explanation for why increasing a choice's iterative generality in this way *proportionally* increases the way in which a choice serves

autonomy. Again, it is possible to add many trivial options to a choice and so make it more general without making it any more significant to individual autonomy. Perhaps reaching a threshold number of options is good enough, or even best, for autonomy. We do know the minimum additive generality a choice requires to be a choice, and if autonomy requires choice, then autonomy requires this minimum generality, but after that there seems little necessary connection between the additive "generality" of choice, and autonomy, and the generality feature of a moral right that protects choices between significant (and possibly wrong) options.

Waldron does make an argument that if we eliminate all wrong acts from the domain of choices protected by moral rights, then rights would protect only a field of enervated choices with options insufficient to support individual autonomy. What would be left, he claims, is either a field of (non-)choices to do solely what is right, on the one hand, or on the other hand, a similarly fallow field of morally indifferent choices composed of "the banalities and trivia of human life,"[20] and neither alternative leaves the kind of robust, significant choices that rights characteristically protect. He reaches this impasse by claiming as a premise that whenever there is a choice between a morally "called-for" action and an alternative that is morally "indifferent," then the alternative indifferent act by that opposition would become an act that is *not* called-for, and therefore, he claims, morally "impermissible."[21] For example, if being respectful at a funeral is called-for, then being disrespectful is by that opposition a wrong act, whereas if being respectful were not called-for, then perhaps being disrespectful would be a morally indifferent act. So if we tried to eliminate all wrong acts from the protections of moral rights, then the called-for acts would systematically dominate all morally indifferent

alternatives by converting them into wrong ones (because they are *not* called-for), leaving a field of (non-)choices to do solely what is right. And if we tried to remedy this problem by also eliminating all called-for acts from the protections of moral rights, then there would remain a field of only choices between morally indifferent trivialities. Since neither alternative fairly represents the significant moral choices that moral rights protect, rights must of necessity protect some wrong acts, Waldron claims.

Let us try to sort out this argument. Waldron asserts as a premise that there are only three types of moral acts: acts that are morally "required", acts that are morally "prohibited" and acts that are morally "indifferent." Morally "required" acts are acts that are "*called for* from the moral point of view", while acts that are morally "prohibited" are acts that are "*subject to moral criticism,*" and finally acts that are morally "indifferent" are acts "on which *morality has nothing of importance to say*."[22] But this division is vague and misleading. He at first appears to seek to divide all morality straightforwardly between morally right, wrong and neutral types of acts. But within the category of morally right acts, besides acts that are morally "required" (Waldron's "called for"), such as keeping a promise to a friend or repairing a harm one caused, there are also acts that we are not required to do but that we have *some moral reason* to do, such as being polite to strangers, or respectful at funerals; and within the category of wrong acts, besides acts that we have some moral reason not to do, or that are "subject to moral criticism" (also Waldron's "prohibited"), such as wasting resources, or being rude, there are also acts that we are *morally required not to do*, such as stealing, or lying, or murdering. There are, therefore, acts that we are *morally required* to do or avoid, acts we have *some moral reason* to do or avoid, and acts that we are *neither morally required nor*

have any moral reason to do or avoid.[23] Waldron's classification obscures these important differences. For example, while Waldron's category of right acts that are "called for" are all morally required, his category of wrong acts that are "prohibited" includes both acts we are morally required not to do and acts that we have some moral reason not to do (ambiguously, all such "prohibited" acts are "subject to moral criticism"). Waldron entirely omits acts that we have some moral reason to do but are not required to do. This omission allows Waldron to argue that right acts dominate neutral acts whenever there is a choice between them; if a right act is required, then we are required to avoid an otherwise morally neutral alternative. But if a right act were not required but is only right because we have some moral reason to do it, then we would have some moral reason, but no requirement, to avoid an otherwise morally neutral alternative, and the right act would not automatically eliminate its neutral alternative.

For Waldron, then, it would seem that it is specifically choices between wrong acts that we have some moral reason but no moral requirement to avoid, and neutral acts about which morality is silent, that support the robust autonomy that rights protect.[24] But this outcome is odd and partial. Is individual autonomy really constituted specifically and only by our choices between morally neutral acts and wrong ones we have some moral reason to avoid? On Waldron's account, it is not choices like being charitable or not (since we have moral reason to give to charity, to give to charity is therefore "called for" and required, and not a genuine choice because we must do it), or choices like driving a car versus walking to work (a choice between morally indifferent options that are therefore insignificant), but instead *only* choices between being rude or not, wasting one's money or not, or misleading people or not, etc. (choices between what we have some moral reason but no requirement to avoid, and what is

arguably indifferently moral) that Waldron claims support the ideal of individual autonomy. But this division seems entirely arbitrary, since it would seem that choices to do what is right but not required should be at least as important to establishing one's moral autonomy as choices to avoid what is wrong but not morally barred, and Waldron gives no reason for why we should close off the possibility of the existence of right acts that we have some moral reason but no moral requirement to do. It may be that morality is strenuous in a way that leaves no room to act morally indifferently or morally wrongly and so always yields a definite moral requirement, as in some forms of Godwin's act-utilitarianism, but this issue is much broader than the one Waldron is discussing here, and makes moral rights that leave any discretion in the hands of the individual problematic, not merely rights that leave open the possibility of doing wrong.[25] Regardless, it seems unlikely that morality would be strenuous on only one side, as it were. If all acts we have moral reason to do are in fact acts we are morally required to do (when opposed to indifferent alternatives), then it would seem equally that all acts we have moral reason to avoid are in fact acts we are morally required to avoid (with indifferent alternatives, again). Waldron tries to inject just enough strenuousness into his argument (by making all right acts required but not making all wrong acts impermissible) to get the RTDW outcome he wants. Moreover, although Waldron argues that doing what is morally required is an insignificant (non-)choice because it is required, and choosing between morally indifferent options is always a trivial choice, these conclusions don't seem warranted. If I choose not to steal when I could easily have done so, such a choice does not seem without moral merit, or like a non-choice, despite that I am morally required not to steal. And if I choose to violate expectations and be a philosopher instead of entering the family business, it is likely that I would not consider that choice trivial or banal despite it being a choice between morally neutral alternatives. Waldron celebrates how, because of their

"connections with the ideas of choice and of the importance of certain areas of decision," moral rights protect both acts that are "stupid, cowardly, tasteless, [and] inconsiderate," as well as acts that are "wise, courageous, cultured, compassionate, creative, honest and good,"[26] and although he involves himself in no contradiction here because on his account rights do protect (non-)choices to do the (required) right thing, nonetheless the congratulatory tone misleads because it suggests that our choices to do right, which Waldron discounts because all are required, are as expressive of the ideal of individual autonomy as are our choices to avoid doing what is wrong but not impermissible in the sense of requirement. For Waldron, only the latter kinds of choices matter to the liberal ideal.

So if general moral rights do sometimes protect wrong acts because they protect choices, and choices that serve autonomy must include wrong options, then no argument Waldron has offered or intimated about choice has established exactly why or how that may be so. He has argued that eliminating choices to do wrong leaves only choices to do right or, instead, only morally indifferent choices, but even if that were the case (and it seems not to be the case), such choices might nonetheless sometimes promote individual autonomy. Nor do choices that are more additively general because they include more options than others necessarily support autonomy any better than choices with fewer options. Autonomy may seem to require a range of significant choices, which may generate a reason for rights to protect such choices, but "significant" is too vague a term to suggest any obvious rationale for including wrong acts within one's choices. Why must significant choices include wrong options? There does not seem to be any identifiable mix of morally required, "called-for", or indifferent or wrong options, or few or many options, that make some choices significant to individual

autonomy and others insignificant to it. So what is meant by the generality of the range of choices that a moral right protects? Why are rights general in this way?

2.2 AN INSTRUMENTAL JUSTIFICATION FOR THE GENERALITY OF MORAL RIGHTS

To be clear on the "generality" of rights again: the structure of Waldron's overall instrumental argument for a RTDW is not that certain choices that include wrong acts (or right acts) in some way uniquely serve autonomy and so rights protect these choices to commit wrong acts; it is instead that certain wrong acts sometimes fall under the general protections that rights provide so as to serve individual autonomy. We do not ask of a particular choice whether this choice serves the individual agent's autonomy in this case; instead, we determine whether this particular choice falls under the range of choices protected by the moral right at issue. There are always direct instrumental arguments as to why or why not a particular choice - of any quality or breadth - might be thought to serve an individual's autonomy. For example, in a particular case - as suggested a few times before - choosing a wrong act may be the best way for the agent to grow morally because the agent will best learn from making a mistake. (Marching in a parade with a racist organization, perhaps the individual encounters fearful or anxious minority-race children, and comes to understand a common humanity.) Or perhaps instead in a particular case it is better to coerce the agent into choosing to do right in the hope that it is by choosing rightly that an agent comes to a better understanding of the reasons behind right choices. (When a lottery winner who has never given to charity is forced to do so instead of wasting it on self-indulgences, she sees how grateful the people she helped are, and she experiences an outpouring of compassion

for them.) Or perhaps an agent may best learn moral autonomy by being tested with a difficult decision to choose right over a tempting wrong. (A sage goes into the desert and is tempted but resists evil.) Each of these outcomes seems to promote the individual's autonomy, yet in one the agent chooses wrong, in another the agent is forced to choose right, and in another the agent chooses right over tempting wrong. Many additional such variations are possible, as well. Yet such instrumental considerations do not seem to matter when we are determining whether a particular choice is protected by a general moral right or not, and therefore whether it is morally wrong to interfere or not. Instead, we determine whether the choice is protected by asking whether the choice is or is not an instance of the class of choices protected by the moral right in question, and if it is, then the choice is protected. Indeed, the procedure for protecting a particular choice as an instance of a moral right may seem to operate as if the instrumental value of the particular choice to the individual's autonomy didn't matter at all. If we determine that a particular choice is an instance of a moral right, then it is protected *regardless* of whether making the choice appears to serve the agent's autonomy or not, and regardless of whether the choice may include a wrong act, right acts, indifferent acts, or whatever. Why might this be so?

The examples of how particular choices may or may not lead to a realization of individual autonomy for the agent demonstrate how difficult it is to determine whether and how making a particular choice may figure in an individual's moral development. It seems difficult to tell whether an agent who marches with a racist organization would emerge with more fervor for the group's racist doctrine or would instead give it up, or whether coercing an agent to give to charity would create a seething resentment or instead an enlightened view, or whether

offering a particular agent difficult choices between right and easy, highly rewarded wrong options would result in good choices or tragic ones that damage the agent's very capacity for moral autonomy. (The marching demonstrator might find her racist prejudices confirmed; the charitable donor might feel bitter if she see her beneficiaries waste her donation; or the sage might lose her faith in the desert.) Even from the agent's own perspective, deciding what is right or wrong in a particular case is not easy because one's own choices and motives are rarely perfectly clear even to oneself. Making such determinations for another person would seem fraught with many times more uncertainty.

Might such uncertainty provide some moral basis for not interfering with others' choices and so for the generality of a moral right? Recall the rule-utilitarian response to objections rooted in some counterintuitive outcomes that arise when directly applying the principle of utility. Critics of utilitarianism often point out that some grievous moral wrongs, such as torturing one person to save many others, seem sanctioned by the principle of utility. Rule utilitarians respond that 1) limited human beings cannot ever fully or sufficiently know all the effects of their actions upon overall utility, and moreover that 2) when given discretion to make such determinations, humans are more likely to merely rationalize their own moral weaknesses than act so as to promote utility; and therefore, 3) a regime of rules is the best way to maximize utility in the long run. Hence when confronted with a moral decision, an individual should determine what rule applies to the case at hand, rather than try to determine how one's action would directly affect overall utility, a calculation laden with error and possible bias. These rules do ultimately derive their justification from their tendency to maximize collective utility and hence are defeasible in extraordinary cases, but the rules enjoy a strong

presumption against their violation, a presumption that is in fact enshrined in moral intuitions.

Critics of the argument that moral rights instrumentally serve autonomy often proceed in a similar way, though their aim is not to raise counterintuitive counterexamples but instead to expose internal inconsistencies in the argument that moral rights exist to promote autonomy. Such critics often point out that moral rights protect many acts, like demonstrating for a noxious doctrine, or hoarding one's money, that are wrong and do not seem to serve an individual's moral autonomy at all. Since most liberals would agree that moral rights protect some such activities, however, the point of this objection is to argue that individual autonomy must not provide a sufficient basis for an instrumental justification for moral rights. But a response resembling that made by rule-utilitarians is nonetheless available. A defender of moral rights as serving autonomy might argue that 1) human beings are fallible, and therefore cannot ever fully know what effects interfering with an agent's acts would have upon that agent's individual autonomy; and, moreover, that 2) human beings are morally limited or biased as well, and therefore more likely to rationalize interfering with an agent's acts so as to reinforce their own moral prejudices than they are to disinterestedly consider how their interference might affect or frustrate the agent's autonomy. Rarely is interfering with another person's choices "for her own good" truly for that purpose; instead it is usually a rationalization for unwarranted interference with acts a majority finds offensive or distasteful.[27] Such considerations of human fallibility and moral limitation or bias lead to the conclusion that therefore 3) a regime of general moral rights that protects harmless moral choices from interference better serves individual autonomy than would constantly exposing the particular acts of individual agents to the many-sided instrumental tests that fallible, biased observers

might impose as to whether that agent's acts serve or disserve that agent's autonomy in each case. Perhaps an omniscient, benevolent God would know precisely when and how to interfere with an individual's acts in such a way so as to never frustrate that individual's autonomy, but only such a God could fully know precisely what the agent's choices are, what they mean, what the agent intends, and who the agent is and how she will react, and therefore what the effects of His interference would be upon the agent's developing moral autonomy, as well as be certain that His motives were entirely benevolent and His reasons for interfering not mere rationalizations. But human beings are not such gods, and so they are better off with moral rights that mark out general classes of choices that are presumed to serve individual autonomy, than they are trying to directly evaluate whether a choice serves an individual's autonomy in each case in an ad-hoc way.

Note that this defense of moral rights that protect a RTDW does not imply that moral judgment is impossible because human beings can never perfectly know the effects of their actions or perfectly understand their own moral motives. Despite such human limitations, people can and often do make morally autonomous choices for themselves. Nor is the argument primarily that an observer cannot possibly know what the agent knows about the agent's situation, and that therefore the agent is in the best cognitive position to make a decision, although that will usually be true.[28] Instead, the argument is that because one cannot know these things for certain about oneself and especially about others in a particular case, or completely know one's own biases, one therefore cannot reliably know the *effects of interference* upon another individual's autonomy. Hence a system of moral rights that picks out in advance what sorts of acts are important to individual autonomy is necessary in order

to effectively protect and promote it. Human fallibility and bias magnify error when one tries to make and enforce moral choices for others.[29] While it may be possible to know that a wrong act is wrong for an individual, and moreover that our interference would prevent the wrong act, we cannot know how our interference would directly affect (or not affect) that individual's autonomy to a degree sufficient to justify our interference instrumentally. By contrast, in the case of the agent "interfering" with herself so as to make a choice, there is no such moral friction between one's intervention and its effects. A RTDW therefore depends upon the moral distinction between self and other. In the case of oneself, autonomous moral action is possible because, despite not fully knowing the nature of one's own choices or motives, or precisely how one's choices will play out, nonetheless there is sufficient knowledge of one's choices and motives and the direct effects of one's action on oneself to justify the hope that making a certain choice will serve or at least not damage one's own moral autonomy. If the effects of one's moral choices upon one's own life were entirely unknown, then moral choice could have no connection with autonomy or self-determination. Autonomy depends at least in part on the expected consequences of one's choices playing out.[30] In the case of interfering with another individual, the interfering observer's fallibility and bias as to the nature of a particular choice, were the observer to make it herself, is multiplied by a further fallibility and bias as to the nature of the agent's perspective on the choice, and that further uncertainty will almost always make the effects of interference upon an agent's autonomy uncertain. The observer must in a sense model the agent's mind within her own mind, asking not only 1) what the observer should do were the observer in the agent's position, and if that is known to a fair degree of certainty, then also 2) whether doing that act would be best for the (modeled) agent, not the observer, to do in the agent's position, and then further calculating 3) how interference would influence the observer were she in the agent's position, and then 4) what

influence that interference would have on the (modeled) agent, not the observer, in the agent's position, with all of these calculations directed towards determining the potential direct effects of her acts of interference upon that particular agent's individual autonomy.[31]

Considerations of human fallibility and bias do not imply that interference is always wrong in every case, of course. The argument here applies only when an observer argues that her interference is for the agent's own good, or would have no adverse effects upon the agent's *autonomy.* In cases where autonomy is not at issue, human fallibility and bias are often superable. In easy cases of justifying interference, interference is usually justified by its reliable connection to the wrongness of the act and the absence or weakness of competing moral values (like individual autonomy, though obviously there are many others). If an agent is doing something wrong, then that always provides a moral reason, and sometimes a moral requirement, to interfere with the agent's act. There may always be some uncertainty as to whether an act is wrong for an agent to do, of course, but in easy cases there is little or no uncertainty, as for example when an agent attempts to murder an innocent person. There is also always uncertainty as to whether the observer interprets the act correctly, but again in easy cases one's interpretation is reliable: perhaps the observer had previously overheard the agent planning the murder and then sees the agent attempt it as planned. Finally, there are often competing moral reasons (or requirements) not to interfere, but once again, in easy cases there are few such reasons: the observer may prevent the murder by anonymously alerting a nearby police officer, for example, rather than endangering herself or anyone else. In such easy-interference cases, the observer's fallibility and bias are minimal, and competing moral reasons not to interfere, such as endangering other lives, don't exist, and so

interference with the attempted crime is both morally justified and here required. By contrast, in more difficult cases of interference, while interference may still justified to the extent it can be by the perceived wrong avoided, either the wrongness of the act for the agent is uncertain, because the observer is limited or biased in some way, or preventing the perceived wrong competes with other important values in such a way as to make interfering morally dubious. In the attempted murder example, the agent may in fact be defending herself from the one she kills, or perhaps the agent and her "victim" are in fact actors in a street performance-art production. Or perhaps the observer is racially biased and assumes that an agent of a certain race must be attempting a murder rather than defending herself. And even if the agent's perception of the situation is accurate and unbiased, no one would blame the observer for not risking her own life so as to prevent an armed violent agent from attempting a murder. Doing so would be morally praiseworthy, perhaps, but such supererogatory acts are not morally required. And it would be wrong for the observer to interfere by risking the life of some other innocent person so as to prevent the agent's crime. So while considerations of fallibility and bias, and competing moral values, must always be considered in determining whether interference is justified, they certainly do not make interference unjustified in every case - only in difficult cases of interference might they do so. The instrumental argument for the generality of moral rights presented here is scoped to apply only when an observer interferes and argues that her interference would not adversely affect the agent's *autonomy*, or that such interference would be good for the agent's autonomy. Fallibility and bias in the heat of a particular case make such a determination too uncertain to be trusted.

For example, suppose an agent chooses to join a public march with a racist organization, and

an observer considers stopping the agent. The observer believes it would be wrong for she herself to march, and moreover, that participating in such a march could never serve her own autonomy. These beliefs provide the observer with a moral reason not to herself join the march, and to the extent that others are similarly situated, also provide her with a reason to interfere with others who choose to march. The observer values individual freedom and autonomy, however, and does not want to interfere with other people if her interference would damage their autonomy, so long as the individual agent is not harming anyone.[32] To the extent that the observer knows the agent here, the observer asks whether it is also wrong for that agent to march, or whether the agent may have some special reason to march that is unknown to the observer, and whether the agent understands the group's noxious message, and then, given the agent's situation, whether marching might nonetheless serve that agent's individual autonomy or not.[33] Does the agent herself sincerely believe that joining the march is morally right, or morally right for her? If so, then that may be a clue to the observer that the agent's autonomy is at issue in a way the observer does not fully understand. The observer must also consider whether her own repugnance at the idea of marching for such a cause may influence her conclusions as to whether marching is wrong for the agent, and moreover is not expressive of the agent's personal autonomy. Suppose the observer concludes that marching is wrong for the agent, perhaps regardless of what the agent thinks about it, and that the observer's own motives in reaching that conclusion are unbiased, and moreover that choosing to march will not serve the agent's autonomy. The observer's conclusion, then, is that the agent does not promote her own autonomy by choosing to march, and since the march is wrong for the agent, the agent should choose not to march. But that is not all. Next the observer must consider whether, if she the observer had thought marching was right for herself and someone who disagreed had interfered with her marching, then how would she

react to that interference? Would such interference result in a reaction in her that would be good or bad for her individual autonomy? Would she the observer tend to react with more strident efforts to march at another time, or in some violent way, or, alternatively, give up on political expression forever? If the observer concludes that such outside-party interference would be good for (or not damage) her own autonomy if she were in the agent's place, then next she must consider, knowing the agent's personality and life circumstances, whether interference would also be good for the agent, or whether she would react in a way differently from how the observer imagines she herself would react in the agent's place, such that interference would make things worse. How would the agent react to being stopped? Would she react in the cool, grateful way that the observer imagines that she herself would react? Or is the agent intransigent and obstreperous in her views, and so would react in some unpredictable negative way?[34] The observer must also consider whether her predictions as to the agent's reaction upon being stopped are not prejudiced with naive ideas of the agent's gratitude at being saved from immorality, or on the other hand, poisoned with thoughts of wreaking revenge or just deserts upon those who do what she herself thinks wrong. If she finally concludes that stopping the agent would not adversely affect the agent's own good of autonomy, and is correct in all her conclusions, and everything works out as she predicts, then perhaps the observer may defend her interference from charges that it impinges upon the agent's autonomy by appealing to its direct (non-) effects upon that agent's individual autonomy.[35] But if the observer is wrong at any step in her chain of inferences and conclusions, then her interference with the agent's act is not defensible by such an appeal.

Contrast how efforts to justify interference proceed under a regime of moral rights. In the

case of the marcher again, the observer believes the agent's act is wrong and considers whether to interfere. The observer does not ask whether marching is wrong for the agent, or serves or disserves the agent's autonomy; or whether interfering might cause the agent to react violently or to become more entrenched in her wrong views. Instead, the observer considers whether the agent's act here is protected under a moral right. She knows that the moral right of free expression protects the free exchange of ideas, and connects with other important political purposes in a liberal democracy such as fair elections and rational, transparent exercises of public power. Free expression is moreover closely allied with a concomitant right of free association that may also be at issue here. While marching is not literally speech, nonetheless marching as a group to proclaim support for racist policies is a potent symbolic act of endorsement for a political doctrine, and is therefore communicative. The observer believes the doctrine here is wrong and odious, and thinks marching in its support is wrong, but nonetheless she concludes that the march is intended to communicate support for a political doctrine and so falls under the protection of the moral right of free expression. Speech protected by a moral right presumptively serves the individual agent's autonomy, and the observer concludes that she would violate the individual's autonomy by interfering.[36] Now suppose that instead of merely marching with the racist group, the agent intends to incite the group to burn down targeted churches or synagogues, and suppose also that it is likely that if the agent urges angry marchers to burn down her targets, then they likely will. The observer knows that the moral right of free expression does not protect crimes, or speech closely related to crimes. Nor does the right protect speech that incites imminent unlawful action, because such speech has little or no connection to truth and the free exchange of ideas.[37] The observer reasons that while the agent's speech here is indeed speech and occurs in the context of a political march, nonetheless the agent's speech seems

closely tied to the commission of a crime, and in fact if the agent is not stopped, then the crime will likely occur, and moreover more speech would not remedy the situation in time to prevent it. The observer concludes that the agent's speech is therefore not protected, and so does not bear on that individual's autonomy. The observer may therefore morally interfere so as to prevent what she believes is a wrong act, without damaging the agent's autonomy. While it is possible that inciting the commission of the crime here could serve the agent's autonomy in some extraordinary way, and so the observer's interference might ultimately frustrate the agent's autonomy, the observer may nonetheless interfere in the confidence that the areas in which interference is detrimental to the agent's autonomy have already been picked out by moral rights, and moreover that the regime of moral rights protects individual autonomy better than would one of ad-hoc determinations. So long as the observer follows the rules laid out by general moral rights, she may be reasonably certain that her acts of interference respect the individual agent's autonomy, and moreover are generally undistorted by her own prejudices.

2.3 SOME DISTINCTIONS FROM RULE-UTILITARIAN ACCOUNTS OF MORAL RIGHTS

The partitioned instrumental justification for moral rights described here is not a rule-utilitarian account of moral rights, despite the structural resemblances. It is especially important to keep in mind one critical way in which they differ: While utility-maximizing rules derive their instrumental justification from the overall effect that following the rules has on maximizing utility collectively, moral rights derive their justification from how they protect and promote the autonomy of *a single individual*. If moral rights were instead justified primarily by their

tendency to promote autonomy overall, in the collective sense that more people would live autonomous lives under a regime of moral rights than they would otherwise, then the instrumental defense of moral rights offered here would not be a theory of rights, but instead a variation on the rule-utilitarian characterization of moral rights as utility-maximizing rules, but with autonomy replacing utility. The spectre of trading off one individual's interest (here, in autonomy) for the collective interest (in autonomy), which so plagues rule-utilitarian accounts of moral rights, would then arise here. On such an account, one might plausibly argue that sacrificing one individual's autonomy in a particular case so as to promote the autonomy of many others would be a good trade overall for autonomy. Avoiding such outcomes is a primary motivation for theories of rights as distinct from any form of utilitarianism. While it may be true that a regime of moral rights would result in more autonomy for more people, we must insist that it is the importance of a single individual's autonomy in a particular case that is by itself sufficient to instrumentally justify respecting that individual's moral rights. This insistence bars trading-off.

Making that insistence may seem to weaken the rationale for the partitioning of the instrumental justification of moral rights, however. It may seem as though it would be easier to know the effects of one's interference upon a single individual's autonomy as opposed to predicting those effects upon a complex collective, as in rule utilitarianism. Perhaps if it were possible to know for certain and without bias that the direct effects of a particular intervention would not damage or would in fact promote an individual's autonomy in a specific case, then that interference may be justified even if it violated a moral right. Such an outcome may be an inevitable result of the partitioning strategy, as no such strategy can be entirely immune to

attacks upon the rationale for the partitioning itself. But if so, then that outcome is not primarily a result of the individualism of the instrumental justification of rights. While maximizing utility in a particular individual's case may seem relatively easier to control despite human limitations when compared with maximizing utility across an entire collective of people, maximizing individual *autonomy* in a particular case is by contrast not so easy to control, even when compared with maximizing autonomy across a collective. It may in fact be easier to make general assumptions about how autonomy works for people abstracted from their circumstances than it would be to make such assumptions about any specific individual as enmeshed in her peculiar life circumstances. Indeed, moral rights that in general promote autonomy seem easier to identify than what specific intervention in a particular idiosyncratic case would best serve an individual agent's autonomy, or whether that specific intervention would damage that individual's autonomy. It seems easier to see how the right of free expression in the abstract helps people lead autonomous lives than it does to see specifically how an individual should express herself, or what the individual should express, towards the same end. Regardless, however, identifying what choices and acts would best lead to individual autonomy seems considerably more difficult than identifying what choices and acts would maximize individual utility,[38] and so problems of fallibility and bias in directly promoting individual autonomy are not avoidable simply because we are concerned with the individual rather than a collective.

The usual rule-utilitarian response to rule-violating counterexamples in which the critic argues that the direct effects of a particular act upon collective utility are fully and easily determined, is that the critic ignores the negative effects that violating the rule will have upon utility. Since

the rule generally tends to increase overall utility, weakening it will tend generally to decrease utility, and violating the rule will weaken its habitual, intuitive hold upon people's behavior. Telling a lie, for example, may maximize utility within the scope of a particular extraordinary case, but if lying becomes so widespread as to no longer feel intuitively wrong to most people, then utility will greatly decrease overall. So the question is then whether telling a lie in a particular case will indeed generate such widespread effects. This rule-utilitarian response is also available in response to a similar attack upon the partitioned instrumental justification of rights, with some qualifications. Although the partitioned instrumental defense of moral rights is rooted in the value of each individual's autonomy taken one at a time, rather than in the collective value of the autonomy of many individuals, nonetheless habits and intuitions that respect moral rights serve to protect a particular individual's autonomy in a particular case equally as well as such habits will serve to promote autonomy in the collective. If the observer has accurate intuitions as to when her interference might tread upon an individual's moral rights, and habits of avoiding interference in the areas that moral rights mark out, then she is less likely to indulge biased desires to interfere with an individual's autonomy in a particular case, or to engage in fallible, biased chains of reasoning as to whether her interference truly affects a particular individual's autonomy or not. Hence protecting intuitions and habits of respecting moral rights is as important for the instrumental theory of moral rights rooted in a single individual's autonomy as protecting the intuitions supporting the rules is for a rule-utilitarian theory rooted in collective utility. A critic who argues that the direct effects of interference upon a particular individual's autonomy are fully known and innocuous must still then also argue that violating that individual's moral rights would not weaken the moral intuitions and habits that protect moral rights in such a way so as to endanger another particular individual's autonomy in the future; in other words, she must argue that taking a

direct instrumental approach is superior to taking an indirect approach that prizes general moral rights. Moreover the instrumental defense of rights may avail itself of the collective good rationale in a subsidiary way, so long as the interest of individuals taken one at a time is understood to take lexical priority: there will be fewer violations of individual autonomy and therefore more autonomous people under a regime of moral rights than there would be under a regime of ad-hoc determinations. That is not an insignificant collective good, though it is a subsidiary rationale for respecting moral rights.

Critics of rule-utilitarianism typically rejoin the attack from there, however, and argue that rule-utilitarianism should therefore endorse *covert* violations of the rules, since such violations would not broadly affect moral intuitions or habits, an outcome that may seem even more morally outrageous.[39] Such critics seek to broaden the attack from one on the counterintuitive outcomes of rule-utilitarianism to an attack on the counterintuitiveness of the rule-utilitarian procedure itself for determining what is moral. A utilitarian procedure that would not, for example, reject outright the morality of harvesting one person's organs to save many others, but instead must consider whether harvesting organs behind closed doors would not weaken the moral rules overall, is not a moral procedure, argues the critic. The partitioned instrumental account of moral rights is not open to the analogous broadened attack, however, because of its insistence on the primacy of the *individual's* autonomy as the source of value. If a direct intervention would negatively impact an individual agent's autonomy, then that intervention can never be justified by an appeal to its subsidiary effects or non-effects upon the intuitions that support moral rights because those intuitions are ultimately rooted in the unique value of that single individual's autonomy. If on the other hand the critic argues that in

a particular hypothetical case we can be absolutely certain that interference will serve or will not damage an individual's autonomy, despite the fact that the interference violates an important general moral right, and moreover that the effects of violating the moral right would not weaken its supporting intuitions, then the instrumentalist defender of moral rights must concede that the interference is justified. But such a concession does not lead to any sense of a morally outrageous outcome such as when one person's interests are traded-off for the interests of many others. In such a case the individual's autonomy is stipulated to be protected and promoted, and that is the ultimate goal of the instrumental account of moral rights. Indeed such a hypothetical case, if it is not entirely unique, may in fact be a legitimate one for a respecification of the relevant moral right.

2.4 A RIGHT TO KNOWINGLY DO WRONG

Consider the following amplifications of some of the wrong acts that Waldron originally adduces:

1. a lottery winner conspicuously wastes her winnings in front of desperately needy people

2. an individual works in support of an overtly racist organization

3. a highly-paid political lobbyist manipulates simpleminded voters to support a cause the lobbyist herself thinks is immoral

4. an athlete takes part in competition with representatives of a racist state because many other athletes have refused to participate, which increases her chances of

winning

5. a man jeers at someone who is depressed or recently bereaved

In Waldron's original examples, whether the agent herself knows her act is wrong is uncertain; but in these amplifications, the agent appears to know her act is wrong and does it anyway, either for reasons not directly related to the act's wrongness, or for the very reasons that make the act wrong, such as because it is selfish or cruel. The moral intuition of a RTDW begins to break down in such cases. When confronted with the question as to whether agents in these amplified cases have a moral right to do their wrong acts, we are tempted to ask for further clarification: does the agent indeed know that her act is wrong, as the cases seem to suggest? For example, does the lottery winner realize that the people who see her indulging herself are needy? Does the insensitive man know that the person he mocks is depressed or bereaved? If the agent does know her act is morally wrong, then her demand to be free to perform a wrong act because she has a moral right seems somehow hypocritical or disingenuous, and the intuition of the agent's moral right to do wrong seems to break down. But why should the agent's knowledge of the wrongness of her act matter? If the agent has a moral right to perform an act, then why does the agent's intention or belief as to whether that act is wrong matter? If the agent's belief does matter, then why does it matter?

2.4.1 CONCEPTUAL ANALYSIS

J.S.Waldron sometimes speaks as if an agent could never logically do what she believed to be morally wrong: "there is something odd about having the intention to do A and sincerely

believing that A is wrong."[40] But people do sometimes intentionally do what is morally wrong, usually for some nonmoral reason related to profit or pleasure; indeed, several of Waldron's own examples of a RTDW seem squarely cases in which the agent probably believes her act is wrong: a lottery winner gambles and drinks all her winnings rather than donating some portion to charity, no doubt because she would rather indulge her pleasures than do what she believes is morally right; an athlete participates in a sports competition with a racist state because the athlete has invested time in training and wants to win an award, despite knowing her participation is wrong.[41] In both cases, the agent would probably admit that she was doing something wrong but that her nonmoral reasons for action (pleasure, status, etc) are more important to her: "I know my act is wrong, but it's within my rights, and I'm doing it anyway." And outside the domain of wrongs that may be protected by rights, the existence of explicitly criminal intent is mundane: lying to sell products, embezzling or stealing for gain, murdering out of hate or fear, etc. There are perhaps libertine philosopher-criminals like the Marquis de Sade who coolly commit crimes while maintaining a lively skepticism as to moral knowledge and truth, but most criminals probably don't consider such issues and instead do what they believe but don't care is morally wrong as the opportunity for personal gain arises.

Waldron's discussion often seems to flow from a paradigm case in which the agent is presumed to mistakenly and innocently believe that her act is not wrong, while the observer knows the act is wrong yet is morally bound to avoid interfering, saying to the agent, "Action A is wrong ...but you have a right to do it."[42] Yet Waldron's conceptual analysis, which demonstrates that an action's wrongness does not by itself entail that the agent has no right to do it, leaves open analytical space for the possibility of the agent making the observer's

assertion on her own behalf: "Action A is wrong ... but *I* have a right to do it." Waldron never explicitly discusses whether an agent's malicious intention to do wrong matters as to whether that agent has or can have a RTDW or not. Waldron does observe that, pragmatically, when someone asserts that they have a right to do A, they often intend to do A, and if it seems semantically odd to intend to do something one believes is wrong, then asserting a right to do A while believing A is wrong - a RTDW - also seems odd in its turn.[43] It may be thought that since to assert that something is wrong entails that one should not do it, or that the agent should in a sense "interfere" with her own doing of the wrong act, while to assert a moral right to do the act is to assert that interference with the agent's doing of the act is morally impermissible; then one could not both assert that one should oneself interfere with the act *and* that interference with the act is morally impermissible. But the agent's moral right to do A entails only the moral impermissiblity of *others'* interfering with A; while the agent's belief that A is wrong entails only that the *agent* should not do (should "interfere" with) her act A:

(1) If I sincerely believe that A is morally wrong, then (2) I should not do A.

If I should not do A, then (3) I should "interfere" with my own doing of A.

(4) If I have a moral right to do A, then (5) others should not interfere with my doing of A.

"I should interfere with my own doing of A" (3) is not logically incompatible with "others should not interfere with my doing of A" (5). The fact that *I* should restrict my own doing of A does not imply that *others* morally may interfere with my action. A moral right to do wrong therefore is not conceptually incompatible with the agent's belief that her act is wrong, unless there is no significant moral difference between the agent and observers of her action.

If I suffered from a mild Freudian schizophrenia of moral conscience and were speaking to a Superego observer of my immoral act within my own mind, then I could not coherently sustain both a belief in the moral wrongness of A and a moral right to do A as against my Superego self. To assert a moral right to do my iniquitous act A as against myself-as-Superego is to assert that I-as-Superego *should not* interfere with my act. I believe my act is wrong; I am Superego; I-as-Superego believe my act is wrong. But if I-as-Superego believe my act is wrong, then that implies that Superego should avoid the act herself when situated similarly to me, and since Superego is situated similarly (indeed, is identically situated) because she is me, then I-as-Superego *should* interfere with my act. I therefore both should and should not interfere with my act, a contradiction. (A similar contradiction results if we start with Ego believing the act is wrong: My act is wrong; I should not do the act; I should interfere with my doing of the act. I have a moral right to do the act; Superego therefore should not interfere with the act. Since I am Superego, therefore I should and should not interfere with my act.) This analysis is mildly risible, of course; however, it does sharpen an important premise behind any moral theory of rights that protects a RTDW: the moral difference between self and other, i.e. that the morality one enforces against oneself, even if thought to apply generally, is not thereby automatically enforceable as against other people. Waldron notes the importance of this premise to a RTDW as well, though he does not observe how a RTDW becomes actually conceptually incoherent were we to abandon it. By contrast, note that it is conceptually possible to assert a moral right against myself so long as I do not think that my act is wrong in either (or both) of my Ego or Superego roles: I-as-Ego assert a moral right; I-as-Superego therefore should not interfere with my act. I believe my act is not wrong; I-as-

Superego believe my act is not wrong. Superego morally may allow herself to do the act when situated similarly to me; Superego is situated identically; therefore I-as-Superego morally may allow myself to do the act. That Superego morally may allow the act does not contradict that Superego morally should not interfere with it. While moral rights to do acts that are not wrong do not conceptually depend upon there being significant moral differences between self and other, RTDWs do. This difference between self and other creates a moral friction associated with interference that exists independently of other moral considerations.

A RTDW is therefore not logically paradoxical or conceptually incoherent even when the agent herself believes her act is wrong, but is a RTDW where the agent does wrong intentionally unjustifiable? As noted earlier, Waldron's discussion seems to flow from a paradigm case in which the agent *mistakenly* does what an observer knows is wrong, though some of his leading examples of a RTDW strongly suggest the agent's awareness of the wrongness of her act. Does Waldron's own general justification of a RTDW preclude cases where the agent acts while knowing[44] that her act is wrong?

2.4.2 INSTRUMENTAL JUSTIFICATION

The partitioned instrumental justification for moral rights may seem to justify a RTDW even in cases in which the wrong act is committed knowingly, because the justification for the generality of moral rights is rooted in the uncertainty of the direct effects of interference, and we cannot know whether an act committed even with awareness of its wrongness may ultimately serve an individual's autonomy or not. But this conclusion is premature. The

reason for partitioning the instrumental defense of a RTDW is that human fallibility and bias make the direct effects of interfering with an agent's wrong act upon that agent's autonomy too difficult to determine. The most difficult determinations to make for an observer seeking to calculate the effects of interference upon an agent's autonomy lie in how the agent's perspective on her own good may significantly differ from what the observer imagines that good to be, because the agent's circumstances and perspective differ from the observer's perspective. They are different people, and the observer's perspective clouds her vision of what she imagines is the agent's good, and therefore how interference would affect the agent's autonomy. If the agent thinks that what she does is right, or at least not wrong, and moreover thinks doing it is important to her own individual autonomy, then it is more likely (though of course not certain) that the observer is mistaken or biased as to whether the act is indeed wrong and contrary to the agent's own good, either because the observer discerns an independent standard of morality only darkly, or because the observer has overlooked something significant in the agent's peculiar circumstances. Even if the observer is correct and unbiased as to the act's wrongness for the agent, then for the same reasons, the observer may still nonetheless be mistaken or biased as to how the act may figure in the agent's autonomous development. If the agent insists that doing the (wrong) act is important to her autonomy, then that is also evidence that the observer may be missing something important. Finally, the agent's reactions to the observer's interference are much more likely to be indignant and counterproductive if that agent believes her act is right (or not wrong) and important to her autonomy, and indeed that the interfering observer has morally wronged her. For example, suppose an observer thinks homosexual relations are morally wrong and considers interfering with an agent's marriage to a same-sex partner. The agent believes that being gay is not wrong and that her marriage is important to her autonomy. Such agent

responses are clues that the observer may not understand the agent's own good properly, perhaps because the observer could not imagine same-sex attractions, or because the observer has no experience with gay people, or because the observer finds homosexual relations repugnant. So the observer may either be in error about the agent's own good, or she may herself be biased in her own experience. Moreover, even if the observer's interference were justified, and she did interfere with the marriage, the agent would likely react in an indignant way that would cut against the observer's intentions or expectations.

But if on the other hand, the agent herself believes her act is wrong, and even that doing the act would damage her individual autonomy, or at least not serve it, then these facts, if known to the observer, provide evidence that there are no significant errors or bias in how the observer perceives the agent's own good and her perspective on it. Of course the agent may herself be wrong about her own good, and indeed that is the usual rationalization for interfering with an agent, but in the case where the agent herself believes her act bad for her autonomy and wrong, then error or bias seems less likely. Wrong acts are usually tempting when they are right from the point of view of self-interest: one is either a fool or a knave, as in Glaucon's challenge to Socrates in the Republic.[45] In such cases, agents have a strong motive to deceive themselves and others into believing that their acts are in fact right, or at least not wrong. But when the agent herself believes her act is wrong and not good for her autonomy, then it seems less likely that she is deceiving herself, and so the agent's admission of her act's wrongness and wrongness for her autonomy constitutes strong evidence that the observer's belief that the agent's act would not serve the agent's autonomy is correct. Agent and observer are here in accord, while it is the differences between them that generate

mistake and bias in the observer. Finally, when an agent confesses to wrongness in this way, then the observer's interference may have more predictable effects upon the agent's autonomy. The agent may in fact be confessing her weakness and indirectly asking the observer for help. For example, consider the case of an agent who is a recovering alcoholic who sits down at a bar and orders a drink. An observer, who is the agent's friend and knows of a public commitment the agent made to avoid drinking, considers interfering. The observer knows the agent believes drinking is wrong and bad for her, and it is the fact that the agent herself believes it wrong that makes the observer's desire to interfere unlikely to rest upon errors as to the agent's own good, or on the observer's own moral bias against drinking if it exists, and moreover her interference is much more likely to be good for the agent, and ultimately (eventually) greeted with the agent's gratitude. But if the observer does not know the agent and merely suspects that she is an alcoholic, and moreover the observer has strong biases against drinking perhaps because the observer is herself a recovering alcoholic, while the agent by contrast is unconcerned about it, then the effects of interference upon the agent's autonomy seem quite unpredictable. Or suppose a secret homosexual relationship is ongoing between two priests who had previously sworn oaths of chastity, and they each believe it wrong, and are in fact members of a support group intended to help priests avoid sexual temptations; then an observer who was also a priest and fellow member of the support group would be less likely to damage the agents' autonomy by discreetly interfering. It is possible, of course, that in such cases both agent and observer are mistaken or biased as to whether the act is bad for the agent's autonomy; and therefore it would be wise for the observer interested in advancing individual autonomy to interfere only when interference seems necessary to avoid serious harm to the agent's autonomy, even in cases when the agent herself admits wrongness and weakness. The agent's admissions here are not

conclusive evidence that interference would not damage her autonomy; they are merely strong evidence for that conclusion.

Hence in general the better the observer knows the agent and the agent's understanding of the good and of her own good, the better are the observer's chances of being correct in her judgments about what is good for the agent's autonomy, and specifically whether interference in a particular case would be good for an agent's autonomy. Moreover the more in accord are the agent and observer in their general moral outlook, the less of a chance there is that the observer's bias will influence her judgment as to the effects of interference upon the agent's autonomy. In general when the agent herself believes her act is wrong and moreover admits it is not good for her own autonomy, then the observer who agrees with these beliefs is in a much better position to justify interfering in a way that promotes or avoids damaging the agent's autonomy.

3 AN EPISTEMIC JUSTIFICATION FOR A RIGHT TO DO WRONG

> "[The Harm Principle is] that the only purpose for which power can be rightfully exercised over any member of a civilised community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right. These are good reasons for remonstrating with him, or reasoning with him, or persuading him, or entreating him, but not for compelling him, or visiting

> him with any evil in case he do otherwise. To justify that, the conduct from which it is desired to deter him must be calculated to produce evil to some one else." (On Liberty, 48).

Because Mill's Harm Principle forbids interfering with an agent's act even when the observer believes that the agent's act is wrong, the Harm Principle may assert a RTDW. I will frame Mill's arguments in On Liberty for the Harm Principle in the context of a RTDW and moral rights, rather than as adjuncts of an ethical theory such as Mill's rule-utilitarianism. I do not believe that this way of framing Mill's argument, however, distorts either his explicit text or the argument's conceptual structure. It is of course possible to interpret Mill's argument in the more straightforward utilitarian way: a right to liberty is instrumentally valuable because freedom of thought and action leads to the discovery of truth, which is valuable to all humanity as a collective. There is support for this interpretation in Mill's text; for example, when Mill remarks that "the peculiar evil of silencing the expression of an opinion is, that it is robbing the human race... of the opportunity of exchanging error for truth..."[46] The argument is not convincing, however, since it is doubtful empirically whether freedom of thought and action generally or usually (or ever) leads to the discovery of moral truth. Moreover, if a truth is discovered, then this utilitarian argument anyway does not seem to provide any reason not to enforce that truth on everyone. It would seem that if liberty is a means to truth, then if the truth is known, then the means to it would become no longer necessary. (And Mill does not skeptically claim that we know nothing.) Yet the various restatements of the Harm Principle[47] and Mill's defenses make it clear that Mill believes the right to liberty survives the unearthing of truth, even to the point of a RTDW. So if Mill does espouse the collectivist utilitarian view, something else must be doing the work that yields the Harm Principle and a possible RTDW.

Mill divides his main thesis in On Liberty into three sections[48] by noting that, when an observer interferes with an agent, either 1) the agent's act is not morally wrong (and possibly is morally right) and the observer is mistaken, or 2) the agent's act is morally wrong and the observer is correct, or 3) the agent's act is partly wrong and partly right and the observer is similarly partly right and partly wrong. Mill here implicitly assumes that the agent herself believes her act is not wrong and, moreover, that the observer's motive in interfering with the agent's act is the wrongness of the act. Mill argues in the first division[49] - where the interfering observer is mistaken that the agent's act is wrong - that while observers typically presume that they know that an agent's act is wrong, that assumption cannot be certain, and has in salient history[50] often proved false, and therefore there is always this possibility that the observer is mistaken and the agent's act is in fact not wrong. The interfering observer is fallible because she is human and may therefore be incorrect, but more importantly for Mill, the observer's belief that the agent's act is wrong is not genuine knowledge unless the observer's belief is held with an acute awareness of its fallibility. That awareness of fallibility requires *inter alia* that the observer allow the agent the liberty to propose and experiment with what seems morally wrong to the observer.[51] If the observer instead suppresses the agent's (harmlessly) wrong act, then the observer implicitly asserts that her own knowledge of the good is infallible and that she need not submit it to any challenge.[52] Such assertions are insufficient justification for interfering. Mill argues in the second division[53] - where the interfering observer is correct that the agent's act is wrong - that if opposing ways of life are never proposed and tested but are instead suppressed, then the observer's knowledge of the moral good, even if correct, lacks the distinctive liveliness and clarity of knowledge, and the

observer's belief is in fact not like knowledge at all but instead more like dogma or prejudice. If the observer does not know why or how an act or way of life that she believes is morally wrong is wrong, then even if she is serendipitously correct in her belief, that belief is not genuine knowledge[54] of the moral good for the agent, and her interference with the agent on the basis of knowing the agent's good is unjustified. Finally, in the third division[55] - where agent and observer are each partially right and partially wrong, though each believes herself wholly correct - Mill argues that when the observer interferes with the agent's act, the observer disrupts the dialectical process essential to acquiring fallible human knowledge of the moral good, and so both cuts herself off from knowing the full truth while also poisoning what partial knowledge she has. The observer should instead allow the agent's (partially) wrong act because it is only through an interplay between contesting ideas that a full knowledge of what is morally good may then emerge. (It is possible to interpret this third division of Mill's argument as a summary statement of his underlying epistemology, rather than merely another analytical possibility.) Hence whether the observer is mistaken, correct, or partially correct that the agent's act is morally wrong, that observer may not justifiably interfere with the agent's harmless act on the basis of her fallible knowledge that the agent's act is morally wrong. To interfere with an agent's harmless act because the observer believes it to be wrong, even in cases where the observer is correct (Division II), vitiates the observer's knowledge of its moral wrongness and invalidates the justificatory basis of her interference. The agent, therefore, has a right to do wrong.

Although Mill here makes an explicitly epistemological argument for moral rights in On Liberty for the Harm Principle, the more usual way to interpret his defense of the Principle in a

context of moral rights is that it establishes a general moral right to liberty that is instrumentally valuable because it protects and promotes individual autonomy, which is uniquely valuable in maximizing overall utility. Such an approach has the advantage of harmonizing Mill's thought in On Liberty with that in Utilitarianism, as well as avoiding the problem of explaining how the epistemic grounds for a right to free expression also extend to support a right to autonomy,[56] since the approach instead puts autonomy on its own footing as a higher-order value. Moreover, there is considerable textual support, especially in Mill's Chapter 3, "Of Individuality, as One of the Elements of Well-Being," for considering autonomy to be of special worth to Mill, as for example when Mill bemoans that "individuality," by which he means individual autonomy and not mere eccentricity or difference, is rarely recognized as having any "intrinsic worth, or deserving any regard on its own account."[57] The argument is that protecting individual autonomy is critical to maximizing utility either or both because each individual is empirically the best judge of her own good[58] or because the individual's interest in her own autonomy is a "higher-order" utility that "gathers feelings around it so much more intense than those concerned in any of the more common cases of utility, that the difference in degree... becomes a real difference in kind."[59] Mill is also sometimes thought to defend the idea that diversity is intrinsically more valuable than uniformity, or that liberty is a precondition for the emergence of the man of genius romantically idealized in the 19th C.[60] These arguments are no doubt present in Mill's thought, but I will for the most part avoid them so as to sharpen his epistemological argument, which it seems to me has been given short-shrift. Where Mill's epistemological argument has been addressed, Mill is sometimes presumed to offer a skeptical account of the possibility of knowledge of moral truth, or to offer a relativist theory of truth.[61] Mill's empiricism is well-known, but his theory of knowledge is mostly imputed from his conclusions, rather than explicit. While it is possible that Mill had no

defensible theory of knowledge guiding his thesis, I will consider some of the possibilities for such an argument that Mill at least suggested.

Mill's epistemological argument seems to be that if the observer's justification for interfering with an agent's act is that the observer knows the act is wrong, then the observer should indeed *know* that the agent's act is wrong; otherwise the observer's justification (and probably her motivation) for interfering fails. The observer's justification for interfering may fail at succeeding points, as well, of course - for example, if her interference would result in a net harm, or for the instrumental reasons discussed earlier - but as an initial matter, the observer must at least know the act is wrong, if that is her justification for interfering. Mill then argues that when the observer interferes with an agent, that very interference vitiates the observer's knowledge in various ways. Since Mill is making an epistemic argument, a good way of thinking about his three divisions may be to think of them as various ways that the interfering observer fails to establish crucial elements of her knowledge of the agent's act's wrongness. The classical formula for knowledge is "justified true belief," which in the light of Gettier problems is better thought to specify some necessary conditions for knowledge rather than jointly sufficient conditions.[62] I will consider these necessary elements as well as discuss some other important epistemic concerns.

3.1 TRUTH AND FALLIBILITY

Mill's first division, then - where the observer is mistaken that the agent's act is wrong - seems concerned primarily with fallibility, or with how an observer's belief may fail to be knowledge

because the belief is not *true*. If the observer's belief that an agent's act is wrong is not true, then the observer does not *know* that the agent's act is wrong, and since the observer is fallible, there is always a chance that the observer's belief may not be true. If I believe that you are breaking and entering a home at night and I stop you, but in fact you own the home and have merely misplaced your keys, then I don't know that your breaking into the house is wrong, because my belief is not true, and my interference to prevent your breaking into the house is therefore not justified. (If you proved to me that you owned the home, then both my justification and my motivation to interfere would disappear.)

But the mere fallibility of our knowledge does not imply that we know nothing. While there is always a chance that whatever we take to be true is in fact false, as for example in the perennial hypothetical case of the Cartesian evil demon who deceives us at every turn, only an epistemological skeptic would deny that we have knowledge on the basis of this bare theoretical possibility of error.[63] Mill is not an epistemological skeptic, and he responds to this obvious point by conceding that while we must assume our opinion to be true despite its fallibility for the guidance of our own conduct, nevertheless when we suppress others' incorrect opinions and wrong acts, we invalidate our assumption of (fallibly known) truth and instead assume our own infallibility. "Complete liberty of contradicting and disproving our opinion is the very condition which justifies us in assuming its truth for purposes of action; and on no other terms can a being with human faculties have any rational assurance of being right."[64] When we don't allow others the complete liberty to disagree and act accordingly, then, Mill argues, we are implicitly assuming that no challenge to our belief can ever be successful, and therefore that we are infallible with regard to that belief. Being open to such

challenges, no matter how much evidence has already been accumulated, or how sure one is of the truth of one's belief, is what it means to admit to the fallibility of one's knowledge. Even Newton's physics remains open to challenge in this way.[65] "It is not the feeling sure of a doctrine which I call an assumption of infallibility... it is the undertaking to decide that question for others, without allowing them to hear what can be said on the contrary side."[66]

But even if the fallibility of knowledge requires that a knower always be open to new evidence telling against one's view, Mill's argument here does not explain why the observer's fallibility entails that she tolerate the agent's various wrong views and actions if the observer has already considered and rejected the views the agent proposes. The observer's fallibility may entail that the observer remain open to the possibility that her own belief may be revised, according to her own best judgment, but that does not mean that the observer must tolerate views that she already knows are wrong, or listen to arguments that she has already evaluated and dismissed. If the agent asserts and acts on a doctrine that the observer (fallibly) knows is wrong, then the observer may still fallibly know the doctrine is wrong even as she interferes with it. Mill may argue that the observer can never truly evaluate whether a doctrine or act that she has rejected for herself is also wrong for the agent, because the observer cannot ever be completely aware of all agent-relative reasons for action.[67] But this argument seems to depend on the contentious empirical assumption that the agent always knows her own circumstances best, and moreover does not seem to be an argument about the fallibility of the observer's knowledge so much as whether the observer may fruitfully apply her knowledge of what is morally good to others, an argument considered in Part II *supra*.

Perhaps Mill is making a more subtle epistemological point here about truth, bearing more on what sort of "belief" is proper to knowledge, rather than on the banal possibility that a belief may be in error. Knowledge does not require simple belief, but specifically the sort of belief that is accepted for the purpose of acquiring truth and avoiding error.[68] If someone believes that Neil Armstrong walked on the moon only because she likes the fanciful idea of it, and not because she cares whether it is true or not, then she (arguably[69]) does not know that Armstrong walked on the moon because she does not accept it for the purpose of acquiring truth and avoiding error, even if her belief is justified (perhaps because she saw a video of the moon-landing on television). She accepts the information because it suits her fancy, and so she is not sure whether her belief is true or not, nor does she care, although she does believe the moon landing happened. For this reason, she may hesitate to communicate her belief to others, or to rely on her belief "for purposes of action."[70] She may for example hesitate to challenge someone who falsely asserts that the moon landing was a hoax. Her belief is therefore not the sort of belief necessary for knowledge. Is Mill arguing that when the observer refuses to tolerate dissenting views that the observer must therefore be holding her belief for some reason other than because the belief is true? For the observer to know the agent's act is wrong "for purposes of action," the observer must believe that her belief (that the agent's act is wrong) is true, and moreover believe it because she believes it is true, and not because the belief is important to society, or useful, or salutary.[71] Perhaps such truth-purposed belief requires that the holder of the belief always be on the lookout for any indication that her belief is not true, or in error, and that would require her to consider and tolerate the agent's assertions of contradictory incorrect beliefs.

But this argument does not reach Mill's conclusion, either, and for the same basic reason. While the observer, in order to have knowledge, must believe what she believes because she believes it is true, and not for some other reason such as because it is useful or pious, this fact does not entail the "complete liberty" for agents to disagree that Mill asserts. The goal of truth in belief-acceptance, coupled with fallibility, may require that the observer remain sensitive to new evidence concerning the truth of her belief, and require that she come to her true belief through a concern for truth, but it does not require her to tolerate views that she has already judged wrong. She need not reconsider over and again views that she thinks are incorrect in order to reasonably hold her own views. Mill's various remarks about the fallibility of knowledge fail to address this basic point: "if the lists are kept open, we may hope that if there be a better truth, it will be found when the human mind is capable of receiving it; and in the meantime we may rely on having attained such approach to truth as is possible in our own day. That is the amount of certainty attainable by a fallible being, and this the sole way of attaining it."[72] But the observer need only keep *her own* lists open for new arguments or proof; she need not tolerate the *agent's* wrong arguments and spurious proof in order to establish her knowledge. While it may be true that "the only way in which a human being can make some approach to knowing the whole of a subject, is by hearing what can be said about it by persons of every variety of opinion, and studying all modes in which it can be looked at by every character of mind,"[73] that is relevant only to the observer's own truth-seeking beliefs, not to the agent's liberty to think and propose alternatives that the observer has already considered and knows are wrong. The fallibility and revisability of the observer's knowledge does not seem to require that she avoid interfering with an agent's wrong views or actions, nor does the observer's interference imply that she interferes for any reason other than because she believes that her belief that the agent's act is wrong is correct.

3.2 JUSTIFICATION AND PROGRESS

Mill's second division - where the observer is correct that the agent's act is wrong - is more concerned with whether the observer's true belief may fail to be knowledge because the observer's belief that the agent's act is wrong is not *justified*, and it is here that the case for a wide liberty for contesting points of view may seem stronger. A belief, even if true, may fail to be knowledge because it is not completely justified. If I correctly believe that you have ten dollars in your pocket, then despite the fact that my belief is true, I do not know you have ten dollars because I am merely lucky in my guess. I have no reason to believe that you have ten dollars in your pocket. Moreover, a slight justification for belief is not enough; one's justification must be robust enough to support knowledge. If I believe you are now at home because you are usually there at this time of day, then even if you are indeed at home right now, I do not know you are at home because I have not verified that you are, and so I cannot rule out the possibility that you are out on an errand, or somewhere else.[74] Hence even if an observer's belief that an agent's view or act is wrong is correct, if the observer is not also completely justified in believing the agent's act is wrong, then the observer does not know the act is wrong, and her interference is unwarranted. If I correctly believe that you are illegally breaking into a house, but I am unaware that only yesterday you lost ownership of the house in a divorce settlement, then my belief is true merely by happenstance, and though true, is not knowledge. If I interfere, then I do so without the knowledge that would properly warrant my interference, despite the fact that you are doing something that is indeed wrong.

These initial epistemological observations about justification as a necessary condition for

knowledge are relatively uncontroversial. But Mill extends the idea of what sort of justification a true belief requires in order to constitute knowledge considerably further. Mill argues that not only must the observer's true belief be positively justified, but that justification also requires that the observer must be able to defend her belief in debate; she must be able to make "a tenable defense" against at least the superficial objections of an opponent who disagrees. If she cannot defend her belief against such an opponent, then her "true opinion abides in the mind ... as a prejudice, a belief independent of, and proof against, argument - this is not the way in which truth ought to be held by a rational being. This is not knowing the truth."[75] Even amongst educated people who are familiar with the grounds of their belief, and who could probably offer a passable argument in support of their belief, if "they have never thrown themselves into the mental position of those who think differently from them, and considered what such persons may have to say; [then] consequently they do not, in any proper sense of the word, know the doctrine which they themselves profess."[76] Mill grants that some truths, such as the mathematical truths of geometry, may be known without considering the arguments against them, since difference of opinion regarding such truths is not possible, but he asserts that even knowing natural scientific truths requires the consideration of alternative views, and to a much greater degree do moral and political truths require such consideration of alternatives.[77] To the objection that most people would not be able to defend their true beliefs in this way, and so they should be able to rely on authority when confronted with the arguments of an ingenious opponent, Mill responds that the very meaning of a doctrine lies in its grounds, and that if a person does not understand the reasons for and against a doctrine, then that person does not really know it, even if she is otherwise justified in believing it is true on reliable authority.[78]

On coherence theories of knowledge, the process of justification takes up a fundamentally disputational pattern of challenge-response in which justification continues until an agreement is reached between disputing parties. Relying on the pragmatic, social context of knowledge, coherence theories thereby resolve the problem of infinite regress in which one belief justifies another, which is justified by another, and another, without end, because as a practical matter justification is only required to proceed to the point of agreement.[79] So if Mill is thinking in terms of a coherence theory of justification, he may see a need for the observer claiming knowledge to be able to answer any challenge that an agent may make to the observer's knowledge claim; otherwise, the justification of the observer's claim is not complete. But even if justification does require that the observer meet all challenges and, therefore (possibly), must fully comprehend the grounds of her belief, as Mill asserts, nonetheless this process need not recur for the observer every single time she encounters an agent holding a wrong view. The observer must *herself* understand the grounds for her knowledge, and if an agent asserts a challenge that the observer has not already considered, then she may be required to properly process that challenge with a suitable response, but if the agent asserts a challenge that the observer has already considered and dismissed in her best judgment, then whether the observer ignores or suppresses the agent's assertion is irrelevant to the justification of the observer's knowledge. The requirement that the observer understand the grounds of her belief does not therefore lead to the conclusion that agents acting on wrong views should be free of interference in asserting and acting upon them. The observer may interfere with an agent doing something foolish and wrong, while fully retaining her understanding of the grounds of her true belief.

Perhaps Mill believes that if the observer were justified in believing that the agent's act is wrong, then the observer should be able to transfer that knowledge to the agent, and so interference would be unnecessary because the agent would then avoid the act on her own volition. If the observer could not convince the agent that her act is wrong, then that by itself is some evidence that the observer's knowledge is not completely justified. Suppose the observer hears a daimonic voice in her head that imparts true information about the good of other people but never explains why that is their good.[80] The observer would not be able to convince the agent that her act is wrong, despite being correct, because the observer does not understand the grounds for her belief; her belief lacks justification. But it is not always possible to communicate one's justified knowledge to another person. Moral knowledge is not scientific in the testable, falsifiable sense, although for Mill it is empirical like all other forms of knowledge. If justification is a matter of coherence and personal acceptance, then the observer may often be unable to communicate all the various experiences and insights she has had that led her to the knowledge of the wrongness of the agent's act.[81] The observer may be able to teach the agent why her act is wrong, but it is likely that for many wrong acts, the agent will have to learn on her own why they are wrong. Most people have already learned a tremendous amount about morality from their families and their culture, and that is quite possibly all that can be learned without first-hand experience. Moreover, it is odd to think that one's justification for one's knowledge depends so critically upon how adept one is at convincing other people that one knows. The observer may simply be a bad teacher or a poor communicator.

Mill argues in his third division - where observer and agent are each partly right - that knowledge of truth progresses in a dialectical fashion when it proceeds at all, though in practice "progress, which ought to superadd, for the most part only substitutes one partial and incomplete truth for another."[82] If truth proceeds dialectically, then that may support Mill's disputational model of justification because any moral view will always be lacking some aspect of the truth that a close consideration of a contending view will reveal. "On every subject on which difference of opinion is possible, the truth depends on a balance to be struck between two sets of conflicting reasons."[83] But again, even if that is the case, the observer need only engage in dialectical debate at whatever point of progress *she herself* has reached in her knowledge; she need not listen to or tolerate wrong views or acts which she has already previously encountered and from which she has already gleaned whatever limited truth they contain. She may suppress and ignore those views, while remaining open to new views, without threatening the justification for her knowledge, even given that truth can only be approached dialectically. If she does not expressly justify her view to the agent, then *the agent* may fail to know that her doctrine or act is wrong or why or what about it is wrong, but the observer's knowledge is no less justified, even given Mill's disputational, dialectical model. If, on the other hand, we took Mill's occasional asides about how difficult truth is to achieve to mean that he is covertly a relativist who holds that there is no way to judge between partial and incomplete truths,[84] and so "progress" indeed always merely "substitutes" one wrong view for another without getting any closer to truth, then while it is true that the observer would then have no knowledge that the agent's act is wrong and so no justification on that basis for interfering, the observer would then also seem to have no basis for *not* interfering with the agent's act, since moral rights themselves would seem to have no basis.[85]

We have seen the same fundamental objection to Mill's epistemic argument reappear in each of these three divisions, and throughout Mill's discussion. The validity of the *observer's* knowledge does not depend upon its justification *to the agent*. The observer must remain open to new views in the interest of obtaining truth in order to be properly cognizant of her own fallibility, but that does not mean that she must be open to any opposing view that any agent proposes, if the observer has already considered and rejected the agent's view as false. The observer may suppress such wrong views and interfere with agents who act upon them without assuming her own infallibility. Moreover, the observer must understand and respond to reasonable challenges to her knowledge in order to justify it, but that does not mean that she must listen and respond to any view of any agent who asserts views that the observer has already considered and rejected. Even on a strongly disputational model of justification, the observer is not required to debate any agent with any view; the observer's knowledge is sufficiently justified if she has met only legitimate competitors to her view. Finally, even supposing a dialectical model of truth, the observer need only consider those truths that she has not previously digested in order to advance her understanding of the full truth; she need not reconsider any truth that any agent proposes. Hence the epistemological requirements that the observer recognize her own fallibility, believe what she believes because it is true, and properly justify her belief, even together with a dialectical account of progress towards truth, do not yield anything like the wide-ranging liberty for agents to propose and act upon their individual views that Mill hopes to defend.

Given how rather obvious this fundamental objection is, it is somewhat surprising that Mill

couches his argument so expressly in epistemological terms. Indeed commentators for the most part have ignored the explicitly epistemological cast of Mill's argument. I have already described one common approach, which regards Mill's essay as a defense of liberty for the purpose of promoting individual autonomy. Waldron, in a slight variation on that approach, interprets Mill to propose a moral ideal of "committed open-mindedness," in which ethical confrontation between differing points of view spurs new ideas and sharpens old ones, resulting in autonomous lives lived with a "lively and committed spirit."[86] On Waldron's interpretation, Mill argues that suppressing and interfering with even wrong views dampens ethical confrontation, and leads to a stagnant society full of people who are not meaningful autonomous. But Mill explicitly discusses "fallibility" and "certainty," and the importance of understanding the "grounds" of one's opinions, of what a "rational" being who is "corrigible" can know, and of what "knowing the truth" means, and the nature of truth. Is there some way to adjust Mill's epistemic argument so as to cure the fundamental objection that the observer's knowledge does not really seem to depend on the agent's liberty to disagree?

3.3 THE AGENT'S KNOWLEDGE

Mill's discussion does suggest an alternative way to epistemologically justify the right of individual liberty: interference with an agent's act because the act is wrong is not justifiable unless *the agent* knows her act is wrong. If the agent does not know that her wrong act is wrong, then the observer's interference for the purpose of preventing the act's wrongness is not warranted. Most of Mill's observations about knowledge would be true if he were talking about a single person's knowledge, that is, the agent's knowledge. For example, Mill asserts that "complete liberty of contradicting and disproving our opinion is the very condition which

justifies us in assuming its truth for purposes of action,"[87] which is a relatively good description of how any fallible knower should *herself* treat her own beliefs: she should not assume any of her beliefs to be infallibly true but, instead, grant *herself* the complete liberty to reexamine them in the light of new countervailing or contradicting evidence. If she does not regard herself as free to potentially revise or discard her views when new evidence appears, then she holds her views on faith, or she assumes her own infallibility. The ambiguity of Mill's statement gives rise to this interpretation, and makes it sound more convincing than it is; if Mill had explicitly said that granting *others* the liberty to contradict our opinion is what justifies us in assuming our opinion is true, then the statement immediately raises the question as to why or how hearing others' (possibly obtuse or inept) assertions should make so much difference to the truth of our own opinions. It is our *own* freedom and willingness to revise our opinions that gives us what limited assurance we have that our opinions are true, not others' freedom to express opposing opinions. Moreover, sometimes Mill seems to be talking about how the observer's interference might vitiate *the agent's knowledge*, which makes more sense, rather than how the observer's interference with the agent would vitiate the observer's knowledge, which seems somewhat mysterious. Mill asserts for example that "it is not the feeling sure of a doctrine which I call an assumption of infallibility... it is the undertaking to decide that question for others, without allowing them to hear what can be said on the contrary side."[88] While it is possible that an interfering observer would assume her own infallibility when preventing an agent from hearing a contrary view, it would make more sense to say that the observer assumes her own infallibility were she to refuse or prevent *herself* from hearing a contrary view. An infallible observer would have no need or desire *herself* to hear contrary views, but whether she allows the *agent* to hear contrary views or not seems irrelevant to whether the *observer's* knowledge is fallible or not. Hearing contrary views is

quite relevant, however, to whether the agent can establish her own knowledge of her view. And anyway an observer who assumed she only fallibly knew what is good may be just as willing to enforce her view as an infallible observer without bothering to explain it to the agent. Indeed, it is the infallible observer who would seem more motivated to let the agent hear all sides, so that the agent might more readily comply with the observer's view.

Finally, Mill sometimes slips into language about the observer's "authority" to interfere with the agent's efforts to acquire knowledge and understanding, implying that the observer's authority to interfere depends upon that agent's knowledge, rather than on the observer's own knowledge. For example, Mill asserts that "those who desire to suppress [an opinion], of course deny its truth; but they are not infallible. They have no authority to decide the question for all mankind, and exclude every other person from the means of judging."[89] But if the observer's authority to interfere depends on the *observer's* knowledge of what is good, then there is no need to grant agents the means of judging for themselves when those agents judge wrongly. If agents are deprived of the opportunity to consider and judge an opinion for themselves, then they would not be able to establish their own knowledge of what the observer already knows - for the reasons Mill so ably outlines: if they believed it, they would do so blindly, and without a suitable justification in the light of alternatives, and deprived of part of the full truth - but the observer would still nonetheless know what *she* knows. And again: "If all mankind minus one were of one opinion, and only one person were of the contrary opinion, mankind would be no more justified in silencing that one person, than he, if he had the power, would be justified in silencing mankind."[90] Mill goes on to explain that the reason humankind must not silence the lone dissenting individual is that the individual may

very well be correct in her opinion. But, as we have seen, and as Mill ultimately concedes, the mere theoretical possibility of error by itself is no reason to avoid interfering with opinions or acts that we (admittedly fallibly) know are wrong; if it is our knowledge that the individual's opinion and acts are wrong that grants us that authority, then its mere fallibility does not undercut that authority. Mill's statement makes sense, however, if our authority for interfering depends upon the *one dissenting individual's knowledge* that her opinion is wrong. If the individual sincerely and with some justification dissents, then presumably she does not herself know that her opinion is wrong, even if we do know it is, and then our interference would not be justified, if our authority to interfere depends upon the agent's knowledge that her opinion is wrong.

3.4 FREEDOM

"The only freedom which deserves the name, is that of pursuing our own good in our own way..." J.S. Mill, "On Liberty," p. 50.

These considerations suggest a novel way of reading Mill's Harm Principle, and indeed of understanding Mill's entire essay in On Liberty. Freedom is choosing how one shall live according to one's own progressive knowledge of what is morally good, so long as one avoids harming others.[91] Since the individual's interest in freedom is of sufficient weight and importance to hold others to be under moral duties to protect and promote it, the individual possesses a right to freedom.[92] Any restrictions on an individual's freedom must be justifiable *to that individual*. Restrictions that are imposed to prevent harm to others are justifiable to that individual.[93] Restrictions on harmless acts that are imposed because the acts are wrong, however, are only justifiable to that individual if *she herself knows* that her acts are wrong.

Hence if an individual does not know that a harmless act she performs is morally wrong despite her best efforts to understand why, and in fact justifiably believes that her act is not wrong, then interference with that act because it is wrong is not justified even if the act is indeed wrong.

The Harm Principle, then, may be understood as a specification or instantiation of the more general principle that any restrictions on an individual's freedom must be justifiable to that individual, an idea that has deep roots in the liberal tradition of thought.[94] The Harm Principle simply specifies a set of restrictions that are presumptively justifiable to the restricted individual: those that prevent harm to others. Restrictions on harmless acts are more difficult to justify to the individual. The general principle does not by itself, therefore, establish that an agent should be free to commit harmlessly wrong acts, or a RTDW, because the wrongness of the act may be thought a sufficient justification to the agent to impose restrictions on her freedom to commit the act. Indeed, that is why a RTDW is a paradoxical intuition, because the wrongness of the agent's act seems to provide a *prima facie* moral reason and justification for interfering that would seem good even to the agent. To establish a RTDW under the general principle, one needs some argument showing that the wrongness of an agent's act does not imply that interference with the act is justifiable to that agent. Mill's reoriented epistemic argument fills this need. Even if an act is wrong and an interfering observer knows that the act is wrong, if the agent does not herself know that her act is wrong, then interference with the act is not justified *to her*. The agent does not understand why she should not be free to commit what she herself does not believe or see any justification for believing is a wrong act. She has the right to pursue what she believes is her "own good,"

according to her own understanding of that good, that is, in her "own way." If we reoriented Mill's epistemic argument in this way, then Mill's real innovation here may lie in thinking of the traditional liberal requirement that restrictions on individual freedom must be "justified" to the restricted agent as laying out an epistemic standard for that justification, at least where such restrictions are imposed because the agent's act is wrong, and the act does no harm.

If an observer interferes with an agent's act because that act is wrong, then in order to justify that interference to the agent, the observer must justify the wrongness of the act to the agent such that the agent would then herself know that the act is wrong. Such an effort would probably require the observer to convincingly explain to the agent why her act is wrong, so that the agent might understand the reasons that her act is wrong. This makes some sense of Mill's arguments in Division II, where he insisted that the justification of knowledge requires that a person be able to articulate the grounds for her opinion in the light of the arguments of able, adversarial advocates for contrary views. Such a requirement seems unnecessary to justify one's knowledge to oneself, and moreover would not require a person to debate any argument proposed by anyone, especially if one had heard the argument before. But if it is the agent's knowledge of the act's wrongness that matters in justifying interference, then the observer would have to respond to any contrary argument any agent submits in order to justify interference, even if the observer had already done so a thousand times before, so that the agent could herself come to an understanding of the act's wrongness. And the observer knowing the full grounds of her opinion in the way that Mill insists would better enable the observer to justify the wrongness of the agent's act to the agent. To justify the wrongness of the agent's act to the agent, the observer must be able to answer each and every reasonable

objection that the agent makes.

Hence the epistemic cast of Mill's main arguments in Chapter Two, "Of the Liberty of Thought and Discussion," begins to make sense as a response to the objection that if an act is wrong, then that by itself should be sufficient justification to the agent for the observer to intervene. This objection flows naturally from a consideration of the general principle that implicitly underlies the Harm Principle presented in the introductory Chapter One, that restrictions on an agent's freedom should be justifiable to that agent. Mill argues that restrictions on (harmless) wrong acts are only justifiable to the agent if the agent knows they are wrong, and describes various ways that knowledge of an act's wrongness may fail. Chapter Three, then, "Of Individuality, As One of the Elements of Well-Being," which commentators have often read as making out an entirely different argument for liberty from that made in Chapter Two, is in fact the next logical step for Mill, then, because Chapter Three describes how individuals come to possess knowledge of what is morally good. Mill argues there that people are ethically progressive beings: we choose different ways of life as we come to a progressively better understanding of what is morally good, and what is morally right or wrong. We come to that knowledge by exercising our "human faculties of perception, judgment, discriminative feeling, mental activity, and ... moral preference" in making autonomous choices as to how to live, and then progressively learning from evaluating those choices. If an individual never makes such autonomous choices, then she "gains no practice either in discerning or in desiring what is best."[95] Individuals should engage in "experiments of living" such that "the worth of different modes of life should be proved practically."[96] As children people should be taught what is morally right and wrong, so as to "know and benefit by the ascertained results

of human experience," but upon adulthood "it is the privilege and proper condition of a human being... to use and interpret experience in his own way."[97]

In a favorite metaphor, Mill regards human nature as "a tree, which requires to grow and develop itself on all sides, according to the tendency of the inward forces which make it a living thing."[98] The inward forces that help an individual achieve her own moral good are her efforts to understand that good. Mill thinks of a human being's moral good in a way that recalls how pragmatists think of truth as an idealization of the conditions of inquiry: what is good is what arises spontaneously in an ideally limpid state of self-understanding, just as what is true is what is revealed in ideal conditions of inquiry. While it is possible to conceive of a true self that is known better to an expert psychologist or guru than it is known to the individual, the individual cannot be freer expressing the truth of the psychologist or guru than that self of which she is aware. If the guru's truth is truer for her, she must first understand that truth *herself* for any expression of it to be free. If the guru or psychologist undertakes to enforce a moral truth upon the individual that she does not understand, then that enforcement violates her freedom, even if it is indeed the truth.

3.5 A RIGHT TO KNOWINGLY DO WRONG ON THE EPISTEMIC ACCOUNT

Mill rejects a moral right to free expression where "want of candour, or malignity, bigotry, or intolerance of feeling manifest themselves."[99] But since "to argue sophistically, to suppress facts or arguments, to misstate the elements of the case, or misrepresent the opposite

opinion" is common even when arguments are made in good faith, "it is rarely possible on adequate grounds, conscientiously to stamp the misrepresentation as morally culpable."[100] Mill asserts that "intemperate" advocacy, advocacy that is for example sarcastic or rude or *ad hominem*, is not immoral and may in fact be useful to clarifying one's commitments, so long as the intemperance does not rise to the level of lies, hate, or bigotry. Mill's remarks come primarily in the context of a concern for how an individual might misrepresent an opponent's view in the advocacy of her own, which Mill for the most part presumes that the individual holds in good faith; however, his remarks apply just as well when the individual misrepresents her own view for the purpose of defending immoral behavior that she knows is wrong.

Hence Mill does not believe that individuals have a right to intentionally espouse or do what they know is wrong. They have a right to be wrong in good faith, and to ignorantly assert and experiment with a wrong doctrine, but they do not have a right to lie or deceive. This conclusion accords with Mill's epistemic argument that restrictions on an individual's freedom to prevent her from committing harmlessly wrong acts are justified if the individual herself knows the acts are wrong. If the individual knows that her act is wrong, as is the case when the individual lies or deceives, then interference with her act is justified to her. Mill also marks out views and acts committed out of bigotry or hate as unprotected. Such motives for acting are not truth-seeking, and an individual who was primarily motivated by them in committing her wrong acts is probably aware that her acts are wrong, or would be aware they are wrong if she considered them in good faith. Individuals have the freedom to pursue their own good in their own way, even if they are mistaken about that good and so commit wrongful acts; they are not free to maliciously do wrong out of hate and bigotry without any regard for whether

their acts are morally wrong or not.

The paradoxical moral intuition of a RTDW flows from a keen appreciation of the individual's interest in living according to *her own understanding* of what is morally good, even if her understanding is flawed. Restrictions on the liberty of the individual to live her life according to her own understanding of the moral good must be justifiable *to her*. If an individual agent does not herself know that her act is wrong, or understand the reasons her act is wrong, therefore, then interference with that act against the agent's will is not justified, even if the act is indeed wrong. This conception therefore deprives cases where an agent exercises her freedom with no regard for what is morally good from the protections of moral rights. In such cases, the agent is not interested in pursuing the good to the best of her knowledge, but is instead malicious or selfish or otherwise indifferent to what is morally good while at the same time claiming a moral right to freely commit immoral acts that she knows are wrong. Such cases are an abuse of freedom as J.S. Mill conceived of it, as that of "pursuing our own good in our own way." Such agents do not pursue their own good; instead, they pursue what they know is wrong.

ENDNOTES

1 See Waldron, "A right to do wrong," in Liberal Rights, Jeremy Waldron (New York: Cambridge University Press 1993), pp.63 - 64.

2 Cf. Al Cohen, "Virtues, Opportunities, and the Right to Do Wrong," Journal of Social Philosophy 28 (1997. Cohen proposes this justification, which he draws from some observations George Sher makes about autonomy, Ibid., p. 53. See George Sher, "Liberal Neutrality and the Value of Autonomy," Social Philosophy and Policy 12, (Winter 1995).

3 A RTDW does not consist in the right of "choosing and doing wrong acts," as Cohen defines it. Ibid. p. 44. Waldron is careful to always refer to "a" RTDW, not "the" RTDW, as Cohen often labels it. Ibid., et seq.

4 In a legal context, Hohfeld distinguishes several analytical senses of what a right is: claim-right, privilege, power, immunity. See Wesley Newcomb Hohfeld, Fundamental Legal Conceptions as Applied in Judicial Reasoning (1919). Here, I discuss primarily what he would term a claim-right, which imposes a correlative duty upon others (or one other), though obviously I am discussing moral, not legal, rights. Whether Hohfeld's distinctions extend straightforwardly or at all to the domain of moral rights, however, is beyond my scope here.

5 George Sher suggests that autonomy may be defined as doing the right thing for the right reasons, which does not require that the agent be able to choose to do wrong because an agent may just as well be coerced into learning to do the right thing for the right reasons. Sher, op. cit.

6 Waldron envisions a liberalism of "men and women of high spirit and high ideals, each living life on his or her own terms..." Waldron, "Liberal Rights: Two sides of the coin," in Liberal Rights, p. 2. I will refer to Waldron's ideal as "autonomy," though I am aware that the term is contentious. I don't think the term's contended meaning will matter to the argument, however.

51 Ibid., p. 55

52 "All silencing of discussion is an assumption of infallibility." Ibid., p. 54

53 Ibid., p. 68.

54 Ibid., p. 76

55 Ibid., p. 76

56 Mill himself notes that the grounds for a right of free expression may not entirely extend to a right of autonomy: "no one pretends that actions should be as free as opinions. [...] Acts, of whatever kind, which without justifiable cause, do harm to others, may be and in the more important cases absolutely require to be, controlled by...the active interference of mankind." 84 But for harmless acts, Mill argues that the same reasons that protect free expression also apply to protect free self-direction.

57 Mill, op. cit., p. 85.

58 "With respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by any one else." Ibid., p. 100 Mill's overall argument does not depend on this assumption, however.

59 From Mill, Utilitarianism, ed. George Sher (Indianapolis: Hackett: 1979), p. 53. I took the quotation from J.L. Mackie's argument in "Can there be a right-based moral theory?" in Theories of Rights, p. 176. See also Edmundson, op. cit., p. 71, who outlines the same argument.

60 J.F.Stephen, imputes this argument to Mill, and responds that genius arises as readily in conditions of oppression as in conditions of freedom. James Fitzjames Stephen, "Mill's Fallacies" in Mill, p. 247.

61 See for example Berlin, "John Stuart Mill and the Ends of Life" in Mill, p. 265. Berlin presumes that Mill's argument fails unless Mill assumes that "human knowledge is in principle

never complete, and always fallible; that there is no single, universally visible truth; that each man, each nation, each civilization might take its own road towards its own goal, not necessarily harmonious with those of others" and that there is no basic universal human nature. But I think Mill neither makes this assumption nor is he required to do so. Some of Mill's explicit remarks in fact conflict with Berlin's view: "As mankind improve, the number of doctrines which are no longer disputed or doubted will be constantly on the increase: and the well-being of mankind may almost be measured by the number and gravity of the truths which have reached the point of being uncontested. ... this gradual narrowing of the bounds of diversity of opinion is necessary in both sense of the term, being at once inevitable and indispensable..." Mill, op. cit., p. 75. Mill does remark, however, that this convergence of known doctrines upon truth is "a stage of intellectual advancement that at present seems at an incalculable distance." Ibid., p. 76 It is possible to reasonably interpret these remarks as an endorsement of something like Pearse's pragmatic Ideal Limit theory of truth, but I don't think it is reasonable to interpret them as a form of relativism, as Berlin does. Berlin may be right, however, that Mill's argument is incompatible with any non-perspectivist theory of truth.

62 See Keith Lehrer, Theory of Knowledge, p. 16, for a general discussion of this point.

63 We must at some point posit the basic trustworthiness of our corpus of beliefs. See Lehrer, Chapter 9, "Skepticism" for a more extensive discussion of this point.

64 Mill, op. cit., p. 55.

65 Ibid., p. 57

66 Ibid., p. 59.

67 "With respect to his own feelings and circumstances, the most ordinary man or woman has means of knowledge immeasurably surpassing those that can be possessed by any one else." Ibid., p. 100.

68 For a fuller discussion, see Lehrer, op. cit., p. 35, in the subsection "Knowledge and the Functional Role of Acceptance."

69 On some "externalist" accounts of knowledge, she may know despite not being aware that her belief is true. A clever epistemological "externalist" response to the skeptical Cartesian circle of ideas is that I do know that what I see exists if what is true properly generates what I see, even if I don't know that it is the truth rather than a demon that generates my perceived world. Theoretical knowledge is not necessary for knowing on such externalist accounts. If the truth causes the world I perceive, then I know it, even if I don't know that I know it. See Lehrer's Chapter 8, ibid., for futher discussion of externalism.

70 "To accept the information that p implies a readiness in the appropriate circumstances to think, infer, and act on the assumption that the information is correct." Ibid., p. 35.

71 "There are, it is alleged, certain beliefs so useful, not to say indispensable, to well-being that it is as much the duty of governments to uphold those beliefs, as to protect any other of the interests of society." Mill, op. cit., p. 57 Mill says this contention merely shifts the assumption of infallibility to the truth of the usefulness of a doctrine.

72 Ibid., p. 57.

73 Ibid., p. 56.

74 See Lehrer, op. cit., p. 12, for similar examples.

75 Mill, op. cit., 68.

76 Ibid., p. 69. Moreover, if there is no such actual opponent, then one must invent one, an *advocatus diaboli*, because "...no one's opinions deserve the name of knowledge, except so far as he has either had forced upon him by others, or gone through of himself, the same mental process which would have been required of him in carrying on an active controversy..." Ibid.

77 Ibid., p. 69.

78 When people don't know the grounds of a true doctrine, "the words which convey it cease to suggest ideas, or suggest only a small portion of those they were originally employed to communicate." Mill, op. cit., p. 71.

79 Lehrer, op. cit., p. 14, and p. 89.

80 If the observer knew that the daimonic voice was always correct, however, then she would not need to justify her interference because she would be infallible. Mill discussion in Division I implicitly assumes that an infallible observer would possess sufficient warrant to enforce her view upon agents with wrong views. If the observer is infallible and knows she is, then she may completely trust her reason to interfere with the agent's act, which is that she knows it is wrong.

81 See Lehrer, op. cit., p. 89, for a discussion of how the relation of a belief to a system of other beliefs may account for its justification, without that relation being an inferential one. A belief may be justified in this way even when no further evidence for it can be articulated.

82 Mill, op. cit., p. 77.

83 Ibid., p. 69.

84 Note, however, that the mere fact that all truths are partial and incomplete does not necessarily imply that there is no way to judge between them. Relativism requires the addition of the latter claim, that rational judgment between truths is not possible, which Mill suggests obliquely when he says that progress merely "substitutes" one for another. Weak relativism see such truths as complementary; strong relativism would see them as incompatible. Mill may be suggesting a strong relativism, then, with his term "substitution."

85 Again, it is important to distinguish this relativism from perspectivism, which does not hold that the observer is unable to judge one perspective better than another. Such a view can

become quite sophisticated. See Alexander Nehamas, Nietzsche: Life as Literature, (Harvard University Press 1987).

86 Waldron, "Mill and moral distress," in Liberal Rights, p. 122.

87 Mill, op. cit., p. 55.

88 Ibid., p. 59.

89 Ibid., p. 54.

90 Ibid., p. 53.

91 We should note that Mill is careful to distinguish acts that affect others' *interests* from acts that simply affect others, so as avoid the objection that it is not possible to meaningfully distinguish harmful acts from harmless ones. Only the former acts qualify as acts that cause harm. See, Rees, op. cit., p. 300.

92 This is why and how rights may generate duties on an Interest Theory approach, such as that Joseph Raz proposes. See Joseph Raz, The Morality of Freedom (Oxford, 1986), p. 166.

93 Mill takes it as more or less self-evident that there is no freedom to harm other people's interests. The vast bulk of his discussion in "On Liberty" is directed towards rebutting the argument that interfering with harmless acts that are wrong is not justified simply because the acts are wrong, which is interesting since he says that "the object of this Essay is to assert one very simple principle," which is, positively, that interference with harmful acts is justifiable. Mill, op. cit., p. 48. Perhaps it is wrong to violate the rights of others because one's own right to freedom implies an equal freedom for everyone else. See Hart, "Are there any natural rights?" in Theories of Rights, pp. 77 - 90.

94 See Waldron, "Theoretical foundations of liberalism" in Liberal Rights, pp. 35 - 62. Waldron asserts that the "*fundamentally liberal*" thesis is that individuals must (at least

theoretically) consent to the social and political order under which they live. Ibid., p. 50. Presumably such consent would only occur when the social order is *justified* to the individuals living under it.

95 Mill, op. cit., p. 86.

96 Ibid., p. 84.

97 Ibid., p. 86.

98 Ibid., p. 87. Mill also frequently invokes this metaphor in his The Subjection of Women.

99 Ibid., p. 83.

100 Ibid., p. 82.